Econometrica, Vol. 68, No. 3 (May, 2000), 525–574

OPTIMAL NONPARAMETRIC ESTIMATION OF FIRST-PRICE AUCTIONS

BY EMMANUEL GUERRE, ISABELLE PERRIGNE, AND QUANG VUONG[1]

This paper proposes a general approach and a computationally convenient estimation procedure for the structural analysis of auction data. Considering first-price sealed-bid auction models within the independent private value paradigm, we show that the underlying distribution of bidders' private values is identified from observed bids and the number of actual bidders without any parametric assumptions. Using the theory of minimax, we establish the best rate of uniform convergence at which the latent density of private values can be estimated nonparametrically from available data. We then propose a two-step kernel-based estimator that converges at the optimal rate.

KEYWORDS: First-price auctions, independent private value, nonparametric identification, two-stage nonparametric estimation, kernel estimation, minimax theory.

1. INTRODUCTION

IN RECENT YEARS THE EMPIRICAL ANALYSIS of auction data has attracted a lot of attention (see Porter (1995) and Laffont (1997) for recent surveys). First, auctions are frequently used to exchange commodities and to allocate public projects through procurements. Thus many auction data are now available. Second, the theory of auctions has considerably expanded since Vickrey's (1961) seminal paper due to the development of the Bayesian Nash equilibrium concept by Harsanyi (1967). Third, after many years of intense theoretical developments, modern industrial organization is facing the challenge of its empirical usefulness. As argued by Sutton (1993), auction models, which emphasize asymmetric information and strategic behavior, are the most favorable case for facing this challenge.

Starting with Paarsch (1992), a few empirical papers have adopted a fully structural econometric approach, which has exclusively relied upon a parametric specification of the bidders' private values distribution. As is well known, even for first-price sealed-bid auctions with independent private values, the structural analysis has suffered from the numerical complexity associated with the compu-

[1]We are grateful to J. J. Laffont for introducing us to auction models. We thank D. Brown, W. Hardle, T. Li, H. Paarsch, M. Simioni, J. Wu, three anonymous referees, the co-editor whose comments greatly improved the paper, and H. Raynal for computational assistance. Various versions were presented at the Xth Journées de Microéconomie Appliquée, Sfax, June 1993, the Far Eastern Meeting of the Econometric Society, Taipei, June 1993, the World Congress of the Econometric Society, Tokyo, August 1995, the European Meeting of the Econometric Society, Toulouse, August 1997, the University of California (Santa Barbara, Riverside, Berkeley, Los Angeles), Stanford, USC, Yale, Cornell, GREMAQ-INRA, Weierstrass Institute and Malinvaud econometric seminars. We thank IDEI for its hospitality during the completion of this paper. Financial support from the National Science Foundation under Grants SBR-9409569 and SBR-9631212 is gratefully acknowledged.

tation of the Bayesian Nash equilibrium strategy. As a result, only very simple parametric specifications of the latent distribution of private values have been entertained. More recently, Laffont and Vuong (1993) and Laffont, Ossard, and Vuong (1995) have proposed some computationally convenient estimation methods based on simulations that allow for general parametric specifications.

The main purpose of this paper is to address three fundamental issues at the heart of the structural analysis: (i) Does a theoretical auction model place any restrictions on observable data to be testable? (ii) Does a structural approach require a priori parametric information about the structural elements to identify the model under consideration? (iii) Can an estimation procedure be proposed that does not rely upon parametric assumptions and that is computationally feasible? We answer these issues for the symmetric first-price auction model within the independent private value (IPV) paradigm.

In particular, we propose an *indirect* two-step procedure for estimating the distribution of bidders' private values from observed bids, which requires neither parametric assumptions nor the computations of the Bayesian Nash equilibrium strategy. The crucial idea of our method is that each private value can be expressed as a function of the corresponding bid, the distribution of observed bids, and the corresponding density function. The first step then consists in constructing a sample of *pseudo* private values based on kernel estimates of the distribution and density functions of observed bids. In a second step, this pseudo sample is used to estimate nonparametrically the density of bidders' private values. We establish the uniform consistency of our estimator and show its optimality in the sense that it attains the best uniform convergence rate for estimating the latent density of private values from observed bids.

Our results are important for several reasons. First, on economic grounds, policy conclusions based on our fully nonparametric procedure are necessarily robust to misspecifications of the underlying distribution. Moreover, because economic theory does impose some restrictions on the distribution of observed bids, one can test in principle the validity of the theoretical auction model without making strong parametric assumptions on its structural elements. Second, from a statistical point of view our estimator applies nonparametric techniques in each step of a structural estimation method.[2] Our first main statistical contribution is to derive the best rate of uniform convergence of nonparametric estimates of the density of (unobserved) private values from observed bids. To this end, we apply the minimax theory as developed by Ibragimov and Has'minskii (1981). To our knowledge, such a theory has been seldom applied in econometric work. We show that our two-step nonparametric estimator attains this optimal rate using suitably chosen bandwidths. Third, on computational grounds, our estimation procedure avoids the numerical diffi-

[2]Though in recent years nonparametric techniques have been used in two-step estimation procedures, typically the second step concentrates on the estimation of a finite dimensional parameter. Only a few authors have used nonparametric techniques in both steps of an estimation procedure. For an example in a regression context, see Ahn (1995) who does not characterize the lower bound for his problem and whose two-step nonparametric estimator is suboptimal in the minimax sense.

culties that have plagued the structural analysis of auction data. Indeed, because it is indirect and fully parametric, our procedure requires neither the numerical determination of the Bayesian Nash equilibrium bid function nor iterative optimization algorithms. Such computational advantages are especially crucial when the Bayesian Nash equilibrium strategy is the solution of a differential equation that cannot be solved explicitly. As a result, the structural analysis of many auction models that are known to be untractable become readily accessible through our indirect method.

The paper is organized as follows. In Section 2, we present the first-price sealed-bid auction model with independent private values. To present the basic ideas, we consider first a nonbinding reservation price, and illustrate our procedure with some Monte Carlo experiments. In Section 3, we establish its asymptotic properties allowing for a varying number of bidders and heterogeneity across auctions. Specifically, we derive the optimal uniform convergence rate for estimating the density of private values from observed bids. We prove the uniform convergence of the pseudo sample as well as the optimality of our two-step nonparametric estimator. In Section 4, we consider the case where the reservation price is binding. In Section 5, we stress the generality of our procedure and indicate some future lines of research. Three Appendices contain the proofs of our results.

2. MODEL, IDENTIFICATION, ESTIMATION

2.1. *The First-Price Auction Model*

A single and indivisible object is auctioned. All bids are collected simultaneously. The object is sold to the highest bidder who pays his bid to the seller, provided the bid is at least as high as a reservation price p_0. In such an institutional framework each bidder does not know others' bids when forming his bid. Within the IPV paradigm, each potential buyer $i = 1, \dots, I$ is assumed to have a private value v_i for the auctioned object. Each bidder does not know other bidders' private values but knows that all private values including his own have been drawn independently from a common distribution $F(\cdot)$, which is absolutely continuous with density $f(\cdot)$ and support $[\underline{v}, \bar{v}] \subset \mathbb{R}_+$. The distribution of private values $F(\cdot)$, the number of potential bidders I, and the reservation price p_0 are common knowledge with $p_0 \in [\underline{v}, \bar{v}]$. In particular, all bidders are identical ex ante and the game is said to be symmetric. Each bidder is assumed to be risk neutral.

The (unique) symmetric differentiable Bayesian Nash equilibrium of the corresponding game of incomplete information was characterized by Riley and Samuelson (1981) among others. Specifically, assuming $I \geq 2$, the equilibrium bid b_i of the ith bidder is

$$(1) \qquad b_i = s(v_i, F, I, p_0) \equiv v_i - \frac{1}{(F(v_i))^{I-1}} \int_{p_0}^{v_i} (F(u))^{I-1}\, du$$

if $v_i \geq p_0$. If $v_i < p_0$, then b_i can be any value strictly less than the reservation price p_0. This strategy is obtained by solving the first-order differential equation in $s(\cdot)$:

$$(2) \qquad 1 = (v_i - s(v_i))(I-1)\frac{f(v_i)}{F(v_i)}\frac{1}{s'(v_i)},$$

with boundary condition $s(p_0) = p_0$. The equilibrium strategy (1) is strictly increasing in v_i on $[p_0, \bar{v}]$, and expresses the equilibrium bid as a function of the bidder's private value, the distribution of private values, the number of bidders, and the reservation price.

In general, bids are observed while private values are unobserved. The preceding theoretical model leads to a closely related structural econometric model. Specifically, because b_i is a function of v_i, which is random and distributed as $F(\cdot)$, then b_i is also random with a distribution $G(\cdot)$ (say) that is uniquely determined by (1). This simple observation is the basis of the structural analysis of auction data. It has led to the development of maximum likelihood estimation methods (see Donald and Paarsch (1993, 1996)) and simulation based estimation methods (see Laffont and Vuong (1993), Laffont, Ossard, and Vuong (1995), and Li and Vuong (1997)). In particular, simulation based methods are convenient and computationally advantageous when moments of bids can be simulated without determining numerically the equilibrium strategy $s(\cdot, F, I, p_0)$ or its inverse $s^{-1}(\cdot, F, I, p_0)$.

All preceding methods can be viewed as *direct* methods as they all start from a parametric specification for $F(\cdot)$ so as to derive expressions for the distribution and moments of observed bids. In contrast, the estimation method proposed in this paper is *indirect* and starts from the estimation of the distribution of observed bids so as to construct an estimate of the distribution of bidders' private values. The crucial idea upon which our identification result and estimation procedure rest is to use the differential equation (2) to express each private value as a known function of the corresponding bid, its distribution and density, the number of bidders, and the reservation price. As a result, an important advantage of our procedure is that it requires neither solving the differential equation (2) nor computing the equilibrium strategy (1) within each iteration of an optimization procedure. To clarify both the conceptual and the technical issues, we begin with a nonbinding reservation price, i.e., $p_0 = \underline{v}$ so that $s(\underline{v}) = \underline{v}$ until Section 4.

2.2. *Nonparametric Identification*

A fundamental issue in structural estimation is whether the structural elements of the economic model are identified from available observations. Because the reservation price is nonbinding, the number I of potential bidders is equal to the number of actual bidders. Hence I and b_i, $i = 1, \ldots, I$, are observed. Thus the only unknown structural element of the model is the latent distribution $F(\cdot)$, and the identification problem reduces to whether this distribution is uniquely determined from observed bids.

Though the equilibrium relation that links the observed bid b_i to the underlying private value v_i is strictly monotonic, the identification problem is nontrivial. This is because the distribution $G(\cdot)$ of b_i depends on the underlying distribution $F(\cdot)$ in two ways: directly through v_i, which is distributed as $F(\cdot)$, and indirectly through the equilibrium strategy $s(\cdot)$, which depends on $F(\cdot)$ (see (1)). The next result solves the identification problem by stating that the distribution $F(\cdot)$ is unique whenever it exists. In addition, it gives a necessary and sufficient condition on the distribution $G(\cdot)$ for the existence of a distribution $F(\cdot)$ of bidders' private values that can "rationalize" $G(\cdot)$.

Our result relies upon the fact that the first derivative $s'(\cdot)$ and the distribution $F(\cdot)$ with its density $f(\cdot)$ can be eliminated *simultaneously* from the differential equation (2) by introducing the distribution $G(\cdot)$ of b_i and its density $g(\cdot)$. Specifically, for every $b \in [\underline{b}, \bar{b}] = [\underline{v}, s(\bar{v})]$ we have $G(b) = \Pr(\tilde{b} \le b) = \Pr(\tilde{v} \le s^{-1}(b)) = F(s^{-1}(b)) = F(v)$, where the last equality uses $b = s(v)$. It follows that the distribution $G(\cdot)$ is absolutely continuous with support $[\underline{v}, s(\bar{v})]$ and density $g(b) = f(v)/s'(v)$, where $v = s^{-1}(b)$. Taking the ratio gives $g(b)/G(b) = (1/s'(v))f(v)/F(v)$. Thus the differential equation (2) becomes

$$v_i = \xi(b_i, G, I) \equiv b_i + \frac{1}{I-1}\frac{G(b_i)}{g(b_i)}. \tag{3}$$

Equation (3) now expresses the individual private value v_i as a function of the individual's equilibrium bid b_i, its distribution $G(\cdot)$, its density $g(\cdot)$, and the number of bidders I. Specifically, (3) states that if b_i is the equilibrium bid, as it is assumed in the structural approach, then the bidder's private value v_i corresponding to b_i must satisfy (3).[3]

We define the set $\mathscr{P}$ of probability distributions $P(\cdot)$ on $\mathbb{R}_+$ as

$$\mathscr{P} = \{P(\cdot) \text{ is absolutely continuous with an interval support in } \mathbb{R}_+\}.^4$$

As usual, we restrict ourselves to strictly increasing and differentiable Bayesian Nash equilibrium strategies.[5] Let $\mathbf{G}(\cdot)$ denote the joint distribution of $(b_1, \ldots, b_I)$.

THEOREM 1: *Let $I \ge 2$. Let $\mathbf{G}(\cdot)$ belong to the set $\mathscr{P}^I$ with support $[\underline{b}, \bar{b}]^I$. There exists a distribution of bidders' private values $F(\cdot) \in \mathscr{P}$ such that $\mathbf{G}(\cdot)$ is the distribution of the equilibrium bids in a first-price sealed-bid auction with independent private values and a nonbinding reservation price if and only if*:

*C*1: $\mathbf{G}(b_1, \ldots, b_I) = \prod_{i=1}^I G(b_i)$.

[3]Though equations similar to (3) have appeared in the decision theoretic literature starting with Friedman (1956), a fundamental difference is that in decision theory each bidder plays as if he/she were alone, while (3) is the result of the differential equation (2), where bidders are in a Nash game, and each bidder's expectations about how others bid agree with their actual bidding behavior. Consequently, in decision theory, $F(\cdot)$ and $G(\cdot)$ need not be related by $F(\cdot) = G[s(\cdot)]$, and bidders need not be at the Bayesian Nash equilibrium. See Laffont (1997). See also Section 4, where our idea of expression v_i in terms of observables leads to (25).

[4]The support is defined as the closure of $\{x : p(x) > 0\}$ where $p(\cdot)$ is a density of $P(\cdot)$. As a result, any distribution in $\mathscr{P}$ is strictly increasing on its support.

[5]Throughout we assume that the second-order conditions hold. Thus the Bayesian Nash equilibrium strategy is fully characterized by the first-order condition (2).

C2: The function $\xi(\cdot, G, I)$ defined in (3) is strictly increasing on $[\underline{b}, \bar{b}]$ and its inverse is differentiable on $[\underline{v}, \bar{v}] \equiv [\xi(\underline{b}, G, I), \xi(\bar{b}, G, I)]$.

Moreover, when $F(\cdot)$ exists, it is unique with support $[\underline{v}, \bar{v}]$ and satisfies $F(v) = G(\xi^{-1}(v, G, I))$ for all $[\underline{v}, \bar{v}]$. In addition, $\xi(\cdot, G, I)$ is the quasi inverse of the equilibrium strategy in the sense that $\xi(b, G, I) = s^{-1}(b, F, I)$ for all $b \in [\underline{b}, \bar{b}]$.

Theorem 1 is important for many reasons. First, it shows that the theoretical auction model does impose some restrictions on the distribution of observed bids. These restrictions can constitute the basis of a formal test of the theory (see Section 5 for a discussion). Specifically, Condition C1 says that bids are independent and identically distributed as $G(\cdot)$. Condition C2 says that, given I, the distribution **G**$(\cdot)$ of observed bids can be rationalized by a distribution of private values $F(\cdot)$ only if $\xi(\cdot, G, I)$ is strictly increasing. For instance, any log-concave distribution $G(\cdot)$, i.e., such that $g(b)/G(b)$ is strictly decreasing, satisfies Condition C2 and thus can be rationalized.[6] On the other hand, densities that exhibit deep U-shaped parts can violate Condition C2. An example is the distribution $G(b) = [b/(5-4b)]^{1/5}$ defined on $[0, 1]$ with $I = 2$. Other examples include some highly peaked multimodal densities.

Second, assuming that buyers behave as predicted by the model of Section 2.1, Theorem 1 establishes that the distribution $F(\cdot)$ of bidders' private values is identified from the distribution of observed bids. In particular, it shows that identification of the structural model does not require a priori parametric specifications. Moreover, because it is nonparametric in nature, our identification result applies to parametric identification as well (see Donald and Paarsch (1996) for a recent contribution on parametric identification). As Roehrig (1988) has argued, parametric identification may be achieved through misspecified parametric specifications, and hence can be misleading.[7]

Third, it is useful to note that the function $\xi(\cdot, G, I)$ is completely determined from the knowledge of $G(\cdot)$ and I. Because $\xi(\cdot, G, I)$ is the quasi inverse of $s(\cdot, F, I)$, one has neither to solve the differential equation (2) nor to apply numerical integration in (1) so as to determine the buyer's equilibrium strategy $s(\cdot, F, I)$. This remark is important because it underlies the principle and the computational advantages of our indirect estimation method presented next.

2.3. *Nonparametric Estimation*

The basic idea of our estimation procedure is straightforward. If one knew the distribution $G(\cdot)$ and density $g(\cdot)$, then one could use (3) to recover every bidder's v_i so as to estimate $f(\cdot)$. Unfortunately, $G(\cdot)$ and $g(\cdot)$ are unknown, but

[6]As a matter of fact, by differentiating (3) with respect to b_i, it can be shown that Condition C2 is equivalent to $g(\cdot)/G^I(\cdot)$ strictly decreasing.

[7]Identification of $F(\cdot)$ can be proved from Roehrig's (1988) Condition 3.2 applied to $b - \tilde{s}(v, I) = 0$, where $\tilde{s}(v, I) \equiv s(v, F, I)$, when v and I are independent conditionally upon the exogenous variables x in $F(\cdot|\cdot)$. The latter independence condition, however, is not used in our proof. Moreover, our proof is constructive as it gives $F(\cdot)$. In addition, our identification result is global instead of local.

they can be estimated from observed bids. This suggests the following two-step estimator. In the first step we construct a sample of pseudo private values from (3) using nonparametric estimates of the distribution and density functions of observed bids. In a second step, this pseudo sample is used to estimate nonparametrically the density of bidders' private values.

To clarify ideas, we consider L homogeneous auctions with the same number of bidders I. These assumptions will be relaxed in the next sections. Let l index the lth auction. Our procedure is as follows. In the first step, we use the observations $\{B_{pl},\ p=1,\dots,I,\ l=1,\dots,L\}$ to estimate nonparametrically $G(\cdot)$ and $g(\cdot)$ by the empirical distribution and the kernel density estimator, respectively, i.e., by

$$\tilde{G}(b)=\frac{1}{IL}\sum_{l=1}^{L}\sum_{p=1}^{I}\mathbb{1}(B_{pl}\le b), \tag{4}$$

$$\tilde{g}(b)=\frac{1}{ILh_g}\sum_{l=1}^{L}\sum_{p=1}^{I}K_g\left(\frac{b-B_{pl}}{h_g}\right), \tag{5}$$

where h_g is a bandwidth and $K_g(\cdot)$ is a kernel with a compact support. The kernel density estimator $\tilde{g}(\cdot)$ is, however, biased at the boundaries of the support.[8] Indeed, let $\rho_g<\infty$ be the length of the support of $K_g(\cdot)$. For $b=\bar{b}-\lambda\rho_g h_g/2$ where $\lambda\in[0,1)$, the expectation of (5) gives $E[\tilde{g}(\bar{b}-\lambda\rho_g h_g/2)]=\int_{-\lambda\rho_g/2}^{(\bar{b}-\underline{b})/h_g-\rho_g/2}K_g(u)g(\bar{b}-\lambda\rho_g h_g/2-h_g u)\,du$ using the change of variable $B=\bar{b}-\lambda\rho_g h_g/2-h_g u$. Hence $E[\tilde{g}(\bar{b}-\lambda\rho_g h_g/2)]-g(\bar{b}-\lambda\rho_g h_g/2)\int_{-\lambda\rho_g/2}^{+\infty}K_g(u)\,du$ goes to zero as $L\to\infty$. Because $\int_{-\lambda\rho_g/2}^{+\infty}K_g(u)\,du\neq 1$, the density estimator is asymptotically biased for $b\in(\bar{b}-\rho_g h_g/2,\bar{b}]$, and similarly in $[\underline{b},\underline{b}+\rho_g h_g/2)$. Thus, using (3) to estimate private values corresponding to observed bids close to the boundaries is likely to be problematic.

Let B_{min} and B_{max} be the minimum and maximum of the IL observed bids. Because $\underline{b}\le B_{min}\le B_{max}\le\bar{b}$, $\tilde{g}(\cdot)$ is asymptotically unbiased on $[B_{min}+\rho_g h_g/2, B_{max}-\rho_g h_g/2]$. This leads to defining the pseudo private value $\hat{V}_{pl}$ corresponding to B_{pl} as

$$\hat{V}_{pl}=\begin{cases} B_{pl}+\dfrac{1}{I-1}\tilde{G}(B_{pl})/\tilde{g}(B_{pl}) \\ \qquad \text{if } B_{min}+\rho_g h_g/2\le B_{pl}\le B_{max}-\rho_g h_g/2, \\ +\infty \quad \text{otherwise} \end{cases} \tag{6}$$

for $p=1,\dots,I$ and $l=1,\dots,L$. The *pseudo* sample of private values $\{\hat{V}_{pl},\ p=1,\dots,I,\ l=1,\dots,L\}$ is used to estimate the density of private values by

$$\hat{f}(v)=\frac{1}{ILh_f}\sum_{l=1}^{L}\sum_{p=1}^{I}K_f\left(\frac{v-\hat{V}_{pl}}{h_f}\right), \tag{7}$$

[8]The support of $g(\cdot)$ is always finite. Indeed $0\le\underline{b}\le\bar{b}$. Moreover, from Laffont, Ossard, and Vuong (1995), $\bar{b}=\int_{\underline{v}}^{\bar{v}}y(I-1)f(y)F^{I-2}(y)\,dy\le(I-1)\int_{\underline{v}}^{\bar{v}}yf(y)\,dy<\infty$ because $E[v_i]<\infty$.

where h_f is a bandwidth and $K_f(\cdot)$ is a kernel. Because our kernels have compact supports, (6) in effect trims observed bids that do not belong to $[B_{min}+\rho_g h_g/2, B_{max}-\rho_g h_g/2]$.

The asymptotic properties as $L\to\infty$ and I fixed of such a two-step nonparametric estimator are obtained in Section 3. When private values are observed and when $f(\cdot)$ has R bounded continuous derivatives, the optimal uniform convergence rate for estimating $f(\cdot)$ is $(L/\log L)^{R/(2R+1)}$ (see Stone (1982)). In our case, private values are unobserved while bids are observed. As a result, this rate cannot be attained. In Theorem 2, we show that the optimal rate is $(L/\log L)^{R/(2R+3)}$, which is smaller than the optimal rate when private values are directly observed. By choosing appropriately the vanishing rates of the bandwidths, namely $h_g=c_g(\log L/L)^{1/(2R+3)}$ and $h_f=c_f(\log L/L)^{1/(2R+3)}$, we show in Theorem 3 that this optimal rate can be attained by our two-step estimator.

2.4. *Monte Carlo Experiments*

To illustrate our two-step nonparametric procedure, we conduct a limited Monte Carlo study. We consider $L=200$ auctions, each having $I=5$ bidders, which gives 1000 observed bids. These numbers correspond to realistic sizes of auction data. Our Monte Carlo experiment consists of 1000 replications. The true distribution F of private values is log-normal with parameters zero and one, truncated at 0.055 and 2.5 to satisfy Assumption A2 of Section 3.1, which corresponds to leaving out 20% approximately of the original log-normal distribution. For each replication, we first generate randomly IL private values from this truncated distribution. We then compute numerically the corresponding bids B_{pl} using (1) with $p_0=\underline{v}$.

Next, we apply our estimation procedure for each replication. First, we estimate the distribution function and density of observed bids using (4) and (5). In a second step we compute the pseudo private values $\hat{V}_{pl}$ using (6). From this pseudo data we estimate the private values density function using (7). Specifically, we consider that the latent density $f(\cdot)$ is once-continuously differentiable so that $R=1$. To satisfy Assumption A3 on the kernels in Section 3, we choose the triweight kernel $(35/32)(1-u^2)^3\mathbb{1}(|u|\leq 1)$ for $K_g(\cdot)$ and $K_f(\cdot)$ so that $\rho_g=\rho_f=2$. As indicated above, the orders of the bandwidths are $L^{-1/5}$. They are chosen as $h_g=1.06\hat{\sigma}_b(IL)^{-1/5}$ and $h_f=1.06\hat{\sigma}_v(IL_T)^{-1/5}$, where $\hat{\sigma}_b$ and $\hat{\sigma}_v$ are the estimated standard deviations of observed bids and trimmed pseudo private values, respectively, and L_T is the number of auctions remaining after the trimming (6). The factor 1.06 follows from the so-called rule of thumb (see Hardle (1991)). The use of I arises because we have I bidders per auction.

The program is written in FORTRAN. For each replication, the execution time lasted less than one minute, which reduces by a factor of 1000 the execution time of a parametric method (nonlinear least squares) because the latter requires the numerical integration of (1). Each replication gives us two estimated functions: (i) the estimated inverse $\hat{\xi}(\cdot)$ of the equilibrium strategy

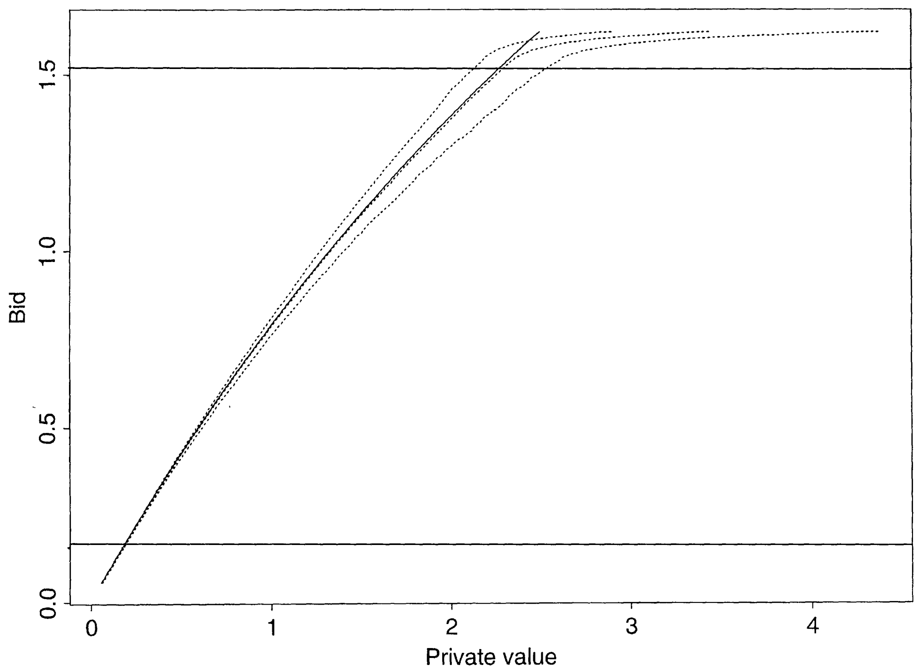

FIGURE 1.—True and estimated equilibrium strategies.

function (see (3)) and (ii) the estimated density function $\hat{f}(\cdot)$, each evaluated at 500 equally spaced points on $[\underline{b}, \bar{b}] = [s(0.055), s(2.5)]$ and $[0.055, 2.5]$, respectively. Our Monte Carlo results are summarized in Figures 1 and 2.

Figure 1 displays the true equilibrium strategy $b = s(v)$ in plain line. We display for each value of $b \in [s(0.055), s(2.5)]$ the mean, the 5% percentile, and the 95% percentile of the 1000 estimates $\hat{\xi}(b)$. This gives the (pointwise) 90% confidence interval for $\xi(b) = s^{-1}(b)$. Figure 2 displays the true density of the truncated log-normal distribution in plain line, and for each value of $v \in [0.055, 2.5]$, the mean, the 5% percentile, and the 95% percentile of the 1000 estimates $\hat{f}(v)$. This gives the (pointwise) 90% confidence interval for $f(v)$. The striking features are that, on the interval bordered by the horizontal/vertical lines, (i) the true curve (the equilibrium strategy or the density) falls within the confidence band and (ii) the mean of the estimates perfectly matches the true curve.[9]

[9]In Figure 1, the horizontal lines are defined by $B_{min} + h_g$ and $B_{max} - h_g$ corresponding to the trimming (6). In Figure 2, the lower (upper) vertical line is defined on average by this trimming to which one h_f is added (subtracted) to eliminate remaining boundary effects. For instance, the lower vertical line corresponds to the average of $\hat{\xi}(B_{min} + h_g) + h_f$.

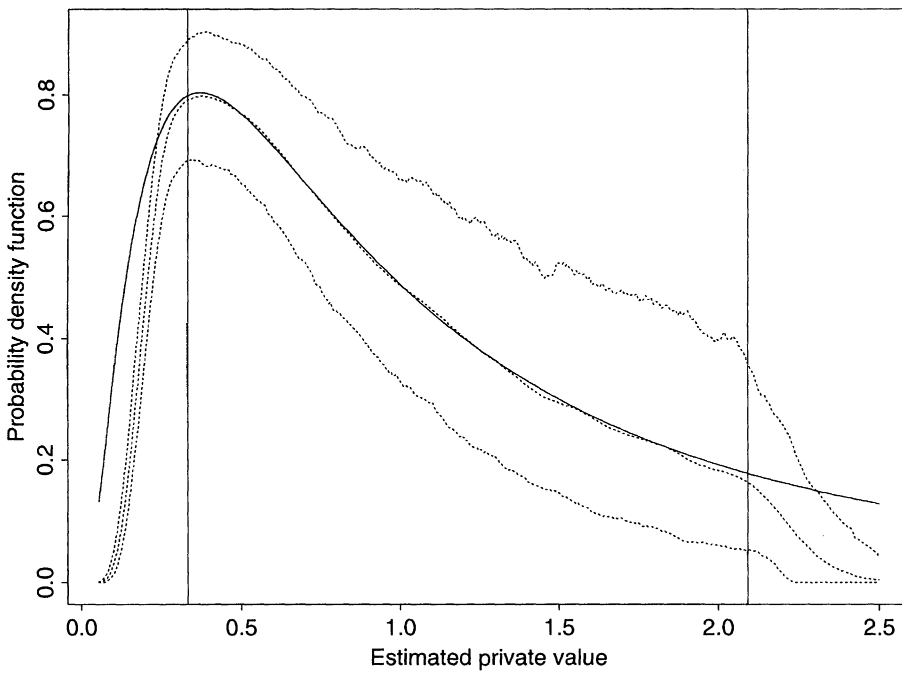

FIGURE 2.—True and estimated densities of private values.

3. ASYMPTOTIC PROPERTIES

In practice, the auctioned objects can be heterogeneous and the number of potential bidders can vary across auctions. These considerations modify our estimator and raise new technical difficulties. In particular, estimation of the boundaries and trimming near the boundaries are not as simple as in (6). To clarify these issues, we still assume in this section that the reservation price is nonbinding.[10]

3.1. *Regularity Assumptions and Key Properties*

Let X_l denote the vector of relevant characteristics for the lth auctioned object, and I_l be the number of bidders in the lth auction.[11] The distribution of bidders' private values V_{pl} for the lth auction is the conditional distribution $F(\cdot|X_l, I_l)$ of private values given (X_l, I_l). Similarly, the distribution of observed

[10] This arises when the reservation price does not constitute an effective screening device such as in Outer Continental Shelf gas and leases auctions (see McAfee and Vincent (1992)).

[11] Following the common knowledge of $F(\cdot)$ in the theoretical model, the vector X_l is common knowledge to all parties. This is frequently justified as the auctioned objects are fully described in a freely available booklet. We then assume that none of X_l is omitted to control heterogeneity across auctioned objects. Thus unobserved heterogeneity comes only from differences in bidders' private values, which are the unobserved random terms in the structural econometric model.

bids in the lth auction is $G(\cdot|X_l, I_l)$. Thus (1) and (3) become

$$(8)\qquad B_{pl} \equiv s(V_{pl}, X_l, I_l) = V_{pl} - \frac{1}{F\left(V_{pl}|X_l, I_l\right)^{I_l-1}} \int_{\underline{v}_l}^{V_{pl}} F(v|X_l, I_l)^{I_l-1}\, dv,$$

$$(9)\qquad V_{pl} \equiv \xi(B_{pl}, X_l, I_l) = B_{pl} + \frac{1}{I_l - 1} \frac{G\left(B_{pl}|X_l, I_l\right)}{g\left(B_{pl}|X_l, I_l\right)},$$

where $\underline{v}_l \equiv \underline{v}(X_l, I_l)$ is the lower bound of the support of $F(\cdot|X_l, I_l)$ and $g(\cdot|\cdot,\cdot)$ is the density of $G(\cdot|\cdot,\cdot)$.

The next assumptions concern the underlying generating process as well as the smoothness of the latent joint distribution of (V_{pl}, X_l, I_l) for any $p = 1, \dots, I_l$.

ASSUMPTION A1:
(i) *The* $(d+1)$*-dimensional vectors* (X_l, I_l), $l = 1, 2, \dots,$ *are independently and identically distributed as* $F_m(\cdot,\cdot)$ *with density* $f_m(\cdot,\cdot)$.
(ii) *For each* l *the variables* V_{pl}, $p = 1, \dots, I_l$, *are independently and identically distributed conditionally upon* (X_l, I_l) *as* $F(\cdot|\cdot,\cdot)$ *with density* $f(\cdot|\cdot,\cdot)$.

In particular, privates values are independent across auctions. As is well known, dependent private values across auctions introduce dynamic considerations that invalidate the Bayesian Nash equilibrium solution (1), and hence are outside the scope of this paper. Note that we do not assume that X_l and I_l are independent from each other. Thus we allow the number of bidders to depend upon the characteristics of the auctioned object.

Let $\mathscr{I}$ be the set of possible values for I_l. Throughout we denote by $S(*)$ the support of $*$, and by $S_i(*)$ the support when the number of bidders is equal to i.

ASSUMPTION A2: $\mathscr{I}$ *is a bounded subset of* $\{2, 3, \dots\}$, *and*:
(i) *for each* $i \in \mathscr{I}$, $S_i(F) = \{(v, x): x \in [\underline{x}, \bar{x}],\ v \in [\underline{v}(x), \bar{v}(x)]\}$, *with* $\underline{x} < \bar{x}$;
(ii) *for* $(v, x, i) \in S(F)$, $f(v|x, i) \geq c_f > 0$, *and for* $(x, i) \in S(F_m)$, $f_m(x, i) \geq c_f > 0$;
(iii) *for each* $i \in \mathscr{I}$, $F(\cdot|\cdot, i)$ *and* $f_m(\cdot, i)$ *admit up to* $R+1$ *continuous bounded partial derivatives on* $S_i(F)$ *and* $S_i(F_m)$, *with* $R \geq 1$.

Without loss of generality, we can assume that $\underline{x}$ and $\bar{x}$ are known.[12] On the other hand, the boundary functions $\underline{v}(\cdot)$ and $\bar{v}(\cdot)$ are typically unknown. Next, because (3) is used to recover private values, it is convenient that $g(\cdot|\cdot,\cdot)$ be bounded away from zero. From Proposition 1 below, this is achieved if the

[12] Because the boundaries $\underline{x}$ and $\bar{x}$ can be estimated at a faster rate than our estimation procedure, our statistical results are not affected. To simplify, we assume that X is a vector of continuous variables. If some X's are discrete, our results still hold with d replaced by the number of continuous variables, and the nonparametric estimators (13), (14), and (20)–(22) modified appropriately following Bierens (1987).

density $f(\cdot|\cdot,\cdot)$ is bounded away from zero, which is the purpose of A2-(ii).[13] Lastly, standard assumptions in the nonparametric literature deal with the smoothness of $f(\cdot|\cdot,\cdot)$. In particular, A2-(iii) implies that $f(\cdot|\cdot,i)$ admits up to R bounded continuous partial derivatives on $S_i(F)$. Consequently, the best uniform convergence rate for estimating $f(\cdot|\cdot,\cdot)$ is $(L/\log L)^{R/(2R+d+1)}$ if private values were observed (see Stone (1982)).

Since private values are unobserved, estimation of the density $f(\cdot|\cdot,\cdot)$ must be based on observed bids. This is called an *inverse* or *indirect estimation* problem; see, e.g., Groeneboom (1996). To determine the best uniform convergence rate for estimating the latent density $f(\cdot|\cdot,\cdot)$ from observed bids, an important step is to study the implied smoothness of the bid density $g(\cdot|\cdot,\cdot)$. This is the purpose of the next proposition.

PROPOSITION 1: *Given A2, the conditional distribution $G(\cdot|\cdot,\cdot)$ satisfies:*
(i) *its support $S(G)$ is such that $S_i(G)=\{(b,x): x\in[\underline{x},\bar{x}],\ b\in[\underline{b}(x,i),\bar{b}(x,i)]\}$, with $\inf_{x\in[\underline{x},\bar{x}]}(\bar{b}(x,i)-\underline{b}(x,i))>0$. Moreover, $(\underline{b}(\cdot,i),\bar{b}(\cdot,i))$ admit up to $R+1$ continuous bounded derivatives on $[\underline{x},\bar{x}]$ for each $i\in\mathscr{I}$, and $\underline{b}(\cdot,i)=\underline{v}(\cdot)$;*
(ii) *for $(b,x,i)\in S(G)$, $g(b|x,i)\geq c_g>0$;*
(iii) *for each $i\in\mathscr{I}$, $G(\cdot|\cdot,i)$ admits up to $R+1$ continuous bounded partial derivatives on $S_i(G)$;*
(iv) *for each $i\in\mathscr{I}$, if $\mathscr{C}_i(B)$ is a closed subset of the interior $S_i^o(G)$ of $S_i(G)$, then $g(\cdot|\cdot,i)$ admits up to $R+1$ continuous bounded partial derivatives on $\mathscr{C}_i(B)$.*

The striking feature of Proposition 1 is item (iv). Specifically, because it has $R+1$ continuous bounded derivatives instead of R, the bid density is smoother than the private value density. The intuition behind this result comes from the equality

$$g(b|x,i)=\frac{G(b|x,i)}{(i-1)(\xi(b,x,i)-b)},$$

which follows from (9). Since $\xi(\cdot,\cdot,i)=s^{-1}(\cdot,\cdot,i)$ and $s(\cdot,\cdot,i)$ has the same smoothness as $F(\cdot|\cdot,i)$ as suggested by (8), then $\xi(\cdot,\cdot,i)$ has $R+1$ continuous bounded derivatives. Now, since $G(\cdot|\cdot,i)$ has also $R+1$ continuous bounded derivatives from (iii), then (iv) follows. As an important consequence, $g(\cdot|\cdot,\cdot)$ can be estimated uniformly at a faster rate, namely $(L/\log L)^{(R+1)/(2R+d+3)}$, than $f(\cdot|\cdot,\cdot)$ can be.

3.2. *Optimal Uniform Convergence Rate*

In this section, we study the optimal rate at which the latent density of private values can be estimated uniformly from observed bids. Indeed, uniform conver-

[13] If A2-(ii) is not tenable, our results still apply provided there exists a known transformation with $R+1$ continuous bounded derivatives such that the transformed density satisfies A1–A2. In this case, one should use the transformed X_l's and $\hat{V}_{pl}$'s in (13)–(14) and (20)–(21) defining our estimator.

gence results are crucial for recovering the shape of a density. As is well-known, however, nonparametric estimators typically converge at different rates depending on the choice of their smoothing parameters. An important issue in nonparametric statistics is to determine the best rate at which the functional of interest can be estimated uniformly. Though the optimal rate of uniform convergence is known for density estimation (see Stone (1982)), this rate does not apply in our case because private values are unobserved. To our knowledge, such a difficult problem has been seldom addressed in structural estimation. Hereafter, we focus upon the estimation of the conditional density $f(v|x)$.[14]

We adopt a minimax approach, as developed by Khas'minskii (1976). We consider joint densities $f(v,x,i)$ satisfying A2. Let $f_0(v,x,i)$ be one such density. In order to determine the optimal rate of uniform convergence r_L^* for estimating the corresponding conditional density $f_0(v|x)$, we study the quantity

$$\inf_{\hat{f}(\cdot|\cdot)} \sup_{f \in U_\epsilon(f_0)} \Pr_g\left(r_L \sup_{(v,x)\in \mathscr{C}(V)} |\hat{f}(v|x) - f(v|x)| > \kappa \right), \tag{10}$$

where κ is a positive constant, $\mathscr{C}(V)$ is an arbitrary inner compact subset with nonempty interior of the support $S(f_0(v,x)) = \bigcup_{i\in\mathscr{I}} S_i(F_0)$ of $f_0(v,x)$, and $\Pr_g(\cdot)$ denotes the probability distribution of (b,x,i) when the underlying density of (v,x,i) is $f(\cdot,\cdot,\cdot)$.

Because we consider the uniform convergence of estimators, the relevant discrepancy measure is the sup norm over $\mathscr{C}(V)$ of the difference between an arbitrary estimator $\hat{f}(\cdot|\cdot)$ and the conditional density $f(\cdot|\cdot)$ of interest. The latter is restricted to belong to the set of densities $U_\epsilon(f_0)$, which is a neighborhood of f_0 defined as

$$U_\epsilon(f_0) \equiv \left\{ f;\ \sup_{(v,x,i)\in S(F_0)} |f(v,x,i) - f_0(v,x,i)| < \epsilon,\quad \|f(\cdot,\cdot,\cdot)\|_R < M \right\},$$

where $M > 0$ and $\|\cdot\|_R < M$ requires densities to have all their derivatives *up to* the Rth order bounded by M uniformly on $S(F_0)$. As in the standard theory, considering the supremum over such a neighborhood avoids superefficient estimators.

The relevance of (10) in determining the optimal uniform convergence rate r_L^* can be explained as follows. First, consider an arbitrary estimator $\hat{f}(\cdot|\cdot)$. Intuitively, if r_L diverges to infinity sufficiently fast, the probability in (10) will tend to one. On the other hand, if r_L does not diverge sufficiently fast, this probability will tend to zero. Now, to determine the optimal uniform conver-

[14] Though our results also apply to the estimation of the conditional density $f(v|x,i)$, our interest in $f(v|x)$ is justified by the economic model, which assumes that the private values and number of bidders are independent conditionally on x so that $f(v|x,i) = f(v|x)$ for every i. If the latter does not hold, a more complex bidding model with a game of entry should be considered.

gence rate for our problem, we must consider all possible estimators of $f(\cdot|\cdot)$. Suppose that r_L diverges to infinity extremely fast; then the probability in (10) will converge to one for every possible estimator and hence (10) will be bounded away from zero. The optimal rate r_L^* is the infimum of such r_Ls.[15]

The next theorem gives an *upper bound* for the optimal uniform convergence rate for estimating $f(\cdot|\cdot)$ from observed bids. Its proof adapts the argument used in Khas'minskii (1976) to our problem, and crucially relies on Theorem 1 and Proposition 1.

THEOREM 2: *Assume that A1–A2 hold and* $\|f_0(\cdot,\cdot,\cdot)\|_R < M$. *Let* $\mathscr{C}(V)$ *be an inner compact subset of* $S(f_0(v,x))$ *with nonempty interior. There exists a constant* $\kappa > 0$ *such that*

$$\lim_{\epsilon \to 0} \varliminf_{L \to +\infty} \inf_{\hat{f}(\cdot|\cdot)} \sup_{f \in U_\epsilon(f_0)} \Pr_g\left(\left(\frac{L}{\log L}\right)^{R/(2R+d+3)} \times \sup_{(v,x)\in\mathscr{C}(V)} |\hat{f}(v|x) - f(v|x)| > \kappa\right) > 0.$$

We consider a lower limit (or limit inf) as $L \to \infty$ because the simple limit of (10) may not exist. Also, because (10) is nondecreasing in ϵ, by taking the limit as $\epsilon \to 0$, we establish in fact the desired result for all ϵ. Specifically, Theorem 2 implies that there exists a strictly positive constant κ such that, for any $\epsilon > 0$, any $L \geq L_0(\epsilon)$:

$$\inf_{\hat{f}(\cdot|\cdot)} \sup_{f \in U_\epsilon(f_0)} \Pr_g\left(\left(\frac{L}{\log L}\right)^{R/(2R+d+3)} \sup_{(v,x)\in\mathscr{C}(V)} |\hat{f}(v|x) - f(v|x)| > \kappa\right) \geq \delta > 0.$$

Thus the optimal uniform convergence rate r_L^* and hence the rate of uniform convergence of any estimator of $f(\cdot|\cdot)$ cannot be larger than $(L/\log L)^{R/(2R+d+3)}$ over $\mathscr{C}(V)$.

Note that this rate is slower than $(L/\log L)^{R/(2R+d+1)}$, which is the optimal rate if private values were observed. This slower rate of convergence is specific to our auction problem, where the variables associated with the density of interest $f(\cdot|\cdot)$ are not observed. Intuitively, this result can be understood as follows. Since $g_0(\cdot|\cdot,\cdot)$ has $R+1$ derivatives (see Proposition 1), $g_0(\cdot|\cdot,\cdot)$ and hence $\xi_0(\cdot,\cdot,\cdot)$ can be estimated at the rate $(L/\log L)^{(R+1)/(2R+d+3)}$ from

[15] Hence, in the statistical literature, determining optimal rates is referred as finding lower bounds.

Stone (1982). But

$$(11)\qquad f_0(v|x,i) = \frac{g_0(\xi_0^{-1}(v,x,i)|x,i)}{\xi_0'(\xi_0^{-1}(v,x,i),x,i)}.$$

Hence the derivative ξ_0' must be estimated.[16] As our results indicate, this is the hardest statistical estimation problem when estimating $f_0(\cdot|\cdot)$, because it requires to estimate g_0' in view of (9). Since the best rate for estimating g_0' and hence ξ_0' is $(L/\log L)^{R/(2R+d+3)}$, this actually gives the best rate at which $f_0(\cdot|\cdot)$ can be estimated from observed bids.

As noted above, Theorem 2 only provides an upper bound to the optimal uniform convergence rate r_L^*. As usual, such a bound would not be much useful if it cannot be attained, i.e., if there does not exist an estimator of $f(\cdot|\cdot)$ that converges at the rate $(L/\log L)^{R/(2R+d+3)}$. In the next subsections, we establish that our two-step nonparametric estimator converges at the rate $(L/\log L)^{R/(2R+d+3)}$ given appropriate choice of the smoothing parameters (see Theorem 3). As a consequence, the optimal uniform convergence rate r_L^* for estimating the latent density $f(\cdot|\cdot)$ from observed bids is $(L/\log L)^{R/(2R+d+3)}$. It also follows that our two-step nonparametric estimator converges at the best possible rate, i.e., is optimal.

3.3. *Definition of the Estimator*

The purpose of this section is to generalize the two-step procedure presented in Section 2.3 to heterogeneous auctions. At the same time, we make precise the assumptions on the kernels and bandwidths, which define our estimator.[17]

Note that (9) can be rewritten as

$$(12)\qquad V_{pl} \equiv \xi(B_{pl}, X_l, I_l) = B_{pl} + \frac{1}{I_l - 1}\frac{G(B_{pl}, X_l, I_l)}{g(B_{pl}, X_l, I_l)},$$

where $G(b,x,i) = G(b|x,i)f_m(x,i) = \int_{\underline{b}(x)}^{b} g(u,x,i)\,du$. Hence, using the observations $\{(B_{pl}, X_l, I_l);\ p = 1,\dots,I_l,\ l = 1,\dots,L\}$, our first step consists in estimating the ratio $\psi(\cdot,\cdot,\cdot) = G(\cdot,\cdot,\cdot)/g(\cdot,\cdot,\cdot)$ by $\tilde{\psi} = \tilde{G}/\tilde{g}$, where

$$(13)\qquad \tilde{G}(b,x,i) = \frac{1}{Lh_G^d}\sum_{l=1}^{L}\frac{1}{I_l}\sum_{p=1}^{I_l}\mathbb{1}(B_{pl} \le b)K_G\left(\frac{x - X_l}{h_G}, \frac{i - I_l}{h_{GI}}\right),$$

[16] Throughout, we use prime to denote a derivative with respect to the first argument of a function.

[17] An alternative estimator is to use directly (11). Specifically, estimation of $g(\cdot|\cdot,\cdot)$, and hence of $\xi(\cdot,\cdot,\cdot)$, followed by determination of $\xi'(\cdot,\cdot,\cdot)$ and $\xi^{-1}(\cdot,\cdot,\cdot)$ would give an estimate of $f(\cdot|\cdot)$. The major problem of this procedure is that $\hat{\xi}(\cdot,\cdot,\cdot)$ may not be invertible. In addition to converging at the optimal rate, our two-step estimator avoids this drawback.

(14) $$\tilde{g}(b,x,i)=\frac{1}{Lh_g^{d+1}}\sum_{l=1}^{L}\frac{1}{I_l}\sum_{p=1}^{I_l}K_g\left(\frac{b-B_{pl}}{h_g},\frac{x-X_l}{h_g},\frac{i-I_l}{h_{gI}}\right),$$

which clearly extend (4)–(5). Here h_G, h_{GI}, h_g, and h_{gI} are some bandwidths, and K_G and K_g are kernels with bounded supports. Following Bierens (1987), we apply kernel techniques to the discrete variable I, which allows for averaging over values of i. Note that we choose different bandwidths for continuous and discrete variables.[18]

In view of (12) it would be natural to estimate each private value V_{pl} by

(15) $$\tilde{V}_{pl}=B_{pl}+\frac{1}{I_l-1}\tilde{\psi}(B_{pl},X_l,I_l),$$

and to use these estimates in a second-step estimation of the conditional density $f(v|x)$. Unfortunately, it is well-known that $\tilde{\psi}$ is an asymptotically biased estimator of ψ at the boundaries of the support of (B,X,I). Because of this boundary effect, we modify slightly (15) by introducing a trimming near the boundaries.

In this aim we estimate the boundary of the support of the joint distribution of (B,X,I), which is unknown.[19] We focus on the estimation of the support $[\underline{b}(x,i),\bar{b}(x,i)]$ of the conditional distribution of B given (X,I) since the support of (X,I) can be assumed to be known (or can be readily estimated). We propose some nonparametric boundary estimators that generalize Geffroy's (1984) estimators to the multidimensional case. Let $h_\partial>0$. We consider the following partition of $\mathbb{R}^d$ with a generic hypercube of side h_∂:

$$\pi_{k_1,\dots,k_d}=[k_1h_\partial,(k_1+1)h_\partial)\times\cdots\times[k_dh_\partial,(k_d+1)h_\partial),$$

where $(k_1,\dots,k_d)$ runs over $\mathbb{Z}^d$. This induces a partition of $[\underline{x},\bar{x}]$. Given an integer i and a value x, the estimate of the upper boundary $\bar{b}(x,i)$ is the maximum of those bids for which $I_l=i$ and the corresponding value of X_l falls in the hypercube $\pi_{k_1,\dots,k_d}$ containing x. The estimate of the lower boundary is similarly defined. Formally, our boundary estimates of the support of the conditional density of B given $(X,I)=(x,i)$ are

(16) $$\hat{\bar{b}}(x,i)=\sup\{B_{pl},\,p=1,\dots,I_l,\,l=1,\dots,L;\ X_l\in\pi_{k_1,\dots,k_d},\,I_l=i\},$$

(17) $$\hat{\underline{b}}(x,i)=\hat{\underline{b}}(x)=\inf\{B_{pl},\,p=1,\dots,I_l,\,l=1,\dots,L;\ X_l\in\pi_{k_1,\dots,k_d}\}.$$

Because $\underline{b}(x,i)=\underline{v}(x)$ is independent of i, we need not restrict (17) to bids such

[18] Because ψ resembles the hazard rate $g/(1-G)$, various nonparametric estimators of the latter can be used. See Hassani, Sarda, and Vieu (1986) and Singpurwalla and Wong (1983) for recent surveys. Our estimator $\tilde{\psi}$ has the characteristic that it takes into account a repeated aspect of our data due to the number of bidders $I_l\geq 2$. In addition, to minimize boundary effects, we have chosen the sum averaging instead of the more common integral averaging for estimating G.

[19] Note that even if the boundary functions $\underline{v}(\cdot)$ and $\bar{v}(\cdot)$ were known the upper boundary $\bar{b}(x,i)$ would be unknown since $\bar{b}(x,i)=s(\bar{v}(x),x,i)$, which depends on the underlying density $f(\cdot|\cdot,\cdot)$.

that $I_l = i$. Our estimator for $S_i(G)$ is $\hat{S}_i(G) \equiv \{(b,x): b \in [\hat{\underline{b}}(x,i), \hat{\bar{b}}(x,i)],\ x \in [\underline{x}, \bar{x}]\}$.

We now turn to the trimming. Basically, for each i we will trim out observations (B_{pl}, X_l, I_l) with $I_l = i$ that are close to the boundaries of the estimated support $\hat{S}_i(G)$ by less than a distance, which is a function of the smoothing parameters. Specifically, let $S(h_G)$ and $S(h_g)$ be the supports of $K_G(\cdot/h_G, 0)$ and $K_g(\cdot/h_g, \cdot/h_g, 0)$, respectively. For instance, if the supports of $K_G(\cdot, 0)$ and $K_g(\cdot,\cdot,0)$ are hypercubes of sides equal to 1, then $S(h_G)$ and $S(h_g)$ are hypercubes of sides h_G and h_g in $\mathbb{R}^d$ and $\mathbb{R}^{d+1}$, respectively. Because we consider points (b,x) in $\mathbb{R}^{d+1}$, we consider hereafter that $S(h_G)$ is $\{0 \times S(h_G)\} \subset \mathbb{R}^{d+1}$. Instead of (15), we define the pseudo private value as

$$\hat{V}_{pl} = B_{pl} + \frac{1}{I_l - 1}\hat{\psi}(B_{pl}, X_l, I_l), \tag{18}$$

where

$$\hat{\psi}(b,x,i) \equiv \begin{cases} \tilde{\psi}(b,x,i) & \text{if } (b,x) + S(2h_G) \subset \hat{S}_i(G) \text{ and} \\ & (b,x) + S(2h_g) \subset \hat{S}_i(G), \\ +\infty & \text{otherwise}. \end{cases} \tag{19}$$

This extends (6) to the heterogeneous case.[20]

In the second step we use the *pseudo sample* $\{(\hat{V}_{pl}, X_l),\ p = 1, \dots, I_l,\ l = 1, \dots, L\}$, to estimate nonparametrically the density $f(v|x)$ by $\hat{f}(v|x) = \hat{f}(v,x)/\hat{f}(x)$, where

$$\hat{f}(v,x) = \frac{1}{Lh_f^{d+1}} \sum_{l=1}^{L} \frac{1}{I_l} \sum_{p=1}^{I_l} K_f\left(\frac{v - \hat{V}_{pl}}{h_f}, \frac{x - X_l}{h_f}\right), \tag{20}$$

$$\hat{f}(x) = \frac{1}{Lh_X^d} \sum_{l=1}^{L} K_X\left(\frac{x - X_l}{h_X}\right), \tag{21}$$

h_f and h_X are bandwidths, and K_f and K_X are kernels. Because the latter have compact supports, the contribution of infinite pseudo private values is nil in our estimate $\hat{f}(v,x)$. Hence, (19) can be interpreted as trimming the observations (B_{pl}, X_l, I_l) that are too close to the boundary of the estimated support $\hat{S}_{I_l}(G)$.

We turn to the choice of kernels and bandwidths defining our two-step estimator.

ASSUMPTION A3:

(i) *The kernels* $K_G(\cdot,\cdot), K_g(\cdot,\cdot,\cdot), K_f(\cdot,\cdot), K_X(\cdot)$ *are symmetric with bounded hypercube supports and twice continuous bounded* (*uniformly in* I) *derivatives with respect to their continuous arguments.*

[20] In fact, as the proof of Proposition 3-(i) shows, the trimming in (19) can be defined using $(1+\epsilon)h_G$ and $(1+\epsilon)h_g$ with $\epsilon > 0$ instead of $2h_G$ and $2h_g$, respectively. Moreover, in the simple case where there are no exogenous variables X, this proof shows that we can have $\epsilon = 0$ because $\underline{b} \le B_{min}$ and $B_{max} \le \bar{b}$. This gives the trimming (6).

(ii) $\int K_G(x, 0)\, dx = 1$, $\int K_g(b, x, 0)\, db\, dx = 1$, $\int K_f(v, x)\, dv\, dx = 1$, *and* $\int K_X(x)\, dx = 1$.

(iii) $K_G(\cdot, 0), K_g(\cdot, \cdot, 0), K_f(\cdot, \cdot), K_X(\cdot)$ *are of order* $R+1, R+1, R, R+1$. *Thus moments of order strictly smaller than the given order vanish.*

Assumption A3 is standard. The orders of the kernels have been chosen according to the smoothness of the estimated functions. In particular, K_g is of order $R+1$, since the density $g(\cdot, \cdot, \cdot)$ admits up to $R+1$ bounded continuous derivatives from Proposition 1.

ASSUMPTION A4:

(i) *As* $L \to \infty$, *the "discrete" bandwidths* h_{GI} *and* h_{gI} *vanish.*

(ii) *The "continuous" bandwidths* h_G, h_g, h_f, *and* h_X *are of the form*:

$$h_G = \lambda_G (\log L/L)^{1/(2R+d+2)}, \quad h_g = \lambda_g (\log L/L)^{1/(2R+d+3)},$$

$$h_f = \lambda_f (\log L/L)^{1/(2R+d+3)}, \quad h_X = \lambda_X (\log L/L)^{1/(2R+d+2)},$$

where the λ*'s are strictly positive constants.*

(iii) *The "boundary" bandwidth is of the form* $h_\partial = \lambda_\partial (\log L/L)^{1/(d+1)}$ *if* $d > 0$.

Part (i) combined with A3-(i) implies that averaging over observations (B_{pl}, X_l, I_l) such that $I_l \neq i$ will disappear from (13) and (14), as $L \to \infty$.

The log L arises because we deal with uniform consistency. From this point of view, h_G, h_g, and h_X are optimal bandwidths given Proposition 1 and A2-(iii) (see, e.g., Hardle (1991)). Hence our kernel estimators (13), (14), and (21) of $G(\cdot, \cdot, \cdot)$, $g(\cdot, \cdot, \cdot)$, and $f(\cdot)$ converge uniformly at the best possible rate. If the private values were observed, the optimal bandwidth for estimating $f(\cdot, \cdot)$ would be of order $(\log L/L)^{1/(2R+d+1)}$, which is asymptotically smaller than the rate for h_f given in (ii). Thus our choice of h_f implies oversmoothing and would be suboptimal in this sense. However, private values are unobserved. As we show below, given the optimal rates for h_G, h_g, and h_X, our choice for h_f is the only rate that achieves the optimal rate of Theorem 2.

3.4. *Uniform Consistency*

Our next main result establishes the uniform consistency of our two-step estimator with its rate of convergence. To do so, we need two results, which are of independent interests. The first proposition establishes the uniform consistency with rates of convergence of our nonparametric boundary estimators (16)–(17).[21]

[21] The bandwidth h_∂ in A4-(iii) does not depend upon the smoothness of the upper and lower bounds $\underline{b}(\cdot, i)$ and $\overline{b}(\cdot, i)$. Therefore, the rate of convergence of our boundary estimators can be improved. See, for instance, Korostelev and Tsybakov (1993).

PROPOSITION 2: *Let* $r_\partial = (L/\log L)^{1/(d+1)}$. *Under* $A1$–$A2$ *and* $A4$-(*iii*),

$$\sup_{(x,i)\in[\underline{x},\bar{x}]\times\mathcal{I}} \left|\hat{\bar{b}}(x,i) - \bar{b}(x,i)\right| = O(1/r_\partial) \quad a.s. \quad and$$

$$\sup_{x\in[\underline{x},\bar{x}]} \left|\hat{\underline{b}}(x) - \underline{b}(x)\right| = O(1/r_\partial) \quad a.s.$$

The second proposition studies the rate at which the pseudo private values $\hat{V}_{pl}$ converge uniformly to the true values.[22]

PROPOSITION 3: *Let* $r_g = (L/\log L)^{R/(2R+d+3)}$, $r_g^* = (L/\log L)^{(R+1)/(2R+d+3)}$. *Under* $A1$–$A4$:

(i) $\sup_{p,l} \mathbb{1}(\hat{V}_{pl} \neq +\infty)|\hat{V}_{pl} - V_{pl}| = O(1/r_g)$ *a.s.*;

(ii) *for any closed inner subset* $\mathscr{C}(V)$ *of* $S(f(v,x))$, *we have*

$$\sup_{p,l} \mathbb{1}_{\mathscr{C}(V)}(V_{pl}, X_l)|\hat{V}_{pl} - V_{pl}| = O(1/r_g^*) \quad a.s.$$

Proposition 3 shows that the pseudo private values $\hat{V}_{pl}$ converge uniformly to the true private values, with the exception of those corresponding to observations (B_{pl}, X_l, I_l) close to the boundaries (see (19)). However, r_g is smaller than r_g^*. Hence the rate of convergence in part (i) is slower than in (ii). The reason is that (i) considers observations that are arbitrary close to the support $S(G)$ as $L \to \infty$, while (ii) deals with observations bounded away from the boundary. Moreover, from (12) and (18) the rate of uniform convergence of $\hat{V}_{pl}$ to V_{pl} depends mainly on that of $\tilde{g}(\cdot,\cdot,i)$ to $g(\cdot,\cdot,i)$. And we know from Proposition 1 that $g(\cdot,\cdot,i)$ is smoother on a closed inner subset of its support.

We now state the main result of this section. Because kernel estimators suffer from boundary effects, we restrict uniform convergence to inner closed subsets of $S(f(v,x))$.

THEOREM 3: *Suppose that* $A1$–$A4$ *hold. We have*

$$\sup_{(v,x)\in\mathscr{C}(V)} |\hat{f}(v|x) - f(v|x)| = O\Big((\log L/L)^{R/(2R+d+3)}\Big) \quad a.s.$$

for any closed inner subset $\mathscr{C}(V)$ *of* $S(f(v,x))$.

In addition to establishing the uniform consistency of our two-step estimator, Theorem 3 is important for two reasons. First, it implies that the upper bound $(L/\log L)^{R/(2R+d+3)}$ of Theorem 2 is in fact the optimal uniform convergence rate r_L^* for estimators of the conditional density $f(\cdot|\cdot)$ from observed bids.

[22] Proposition 3 is interesting in itself because it can also be used for the estimation of the conditional mean, variance, or quantiles, etc... of private values.

Second, it shows that our two-step estimator attains the optimal rate r_L^*, and hence is asymptotically optimal.

We present the proof of Theorem 3 as it helps understand the role of Proposition 3 and our bandwidth choices.

PROOF OF THEOREM 3: We have $\hat{f}(v|x) = \hat{f}(v,x)/\hat{f}(x)$. Given the optimal bandwidth choice for h_X in A4, we know that $\hat{f}(x)$ converges uniformly to $f(x)$ at the rate $(L/\log L)^{(R+1)/(2R+d+2)}$ on any inner compact subset of its support (see Hardle (1991)). Because this rate is faster than that of the theorem and $f(x)$ is bounded away from 0 by A2-(ii), it suffices to show that $\hat{f}(v,x)$ converges at the rate $(L/\log L)^{R/(2R+d+3)}$.

Our proof relies upon the (infeasible) nonparametric estimator of the density of (V,X) using the unobserved true private values V_{pl}:

$$(22)\qquad \tilde{f}(v,x) = \frac{1}{Lh_f^{d+1}} \sum_{l=1}^{L} \frac{1}{I_l} \sum_{p=1}^{I_l} K_f\left(\frac{v-V_{pl}}{h_f}, \frac{x-X_l}{h_f}\right).$$

Lemma B2 in Appendix B implies that the suboptimal bandwidth h_f leads to a uniform convergence of $\tilde{f}(v,x)$ to $f(v,x)$ on $\mathscr{C}(V)$ at the rate $(L/\log L)^{R/(2R+d+3)}$. Since $\hat{f}(v,x) - f(v,x) = [\hat{f}(v,x) - \tilde{f}(v,x)] + [\tilde{f}(v,x) - f(v,x)]$, we are left with the first term.

Let $\mathscr{C}'(V)$ be an inner closed subset of $S(f(v,x))$ containing all hypercubes of size δ (small enough) centered at a point (v,x) in $\mathscr{C}(V)$. Define $\mathscr{C}''(V)$ similarly with respect to $\mathscr{C}'(V)$. Hence $\mathscr{C}(V) \subset \mathscr{C}'(V) \subset \mathscr{C}''(V) \subset S(f(v,x))$. Now, for $(v,x) \in \mathscr{C}(V)$ and L large enough, $\hat{f}(v,x)$ uses at most observations $(\hat{V}_{pl}, X_l)$ in $\mathscr{C}'(V)$ and hence for which (V_{pl}, X_l) is in $\mathscr{C}''(V)$ by Proposition 3-(i). Because $\tilde{f}(v,x)$ uses at most (V_{pl}, X_l) in $\mathscr{C}''(V)$ for any (v,x) in $\mathscr{C}(V)$, we obtain almost surely for L large enough,

$$\begin{aligned}
&\hat{f}(v,x) - \tilde{f}(v,x)\\
&\quad = \frac{1}{Lh_f^{d+1}} \sum_{l=1}^{L} \frac{1}{I_l} \sum_{p=1}^{I_l} \mathbb{1}_{\mathscr{C}''(V)}(V_{pl}, X_l)\\
&\qquad \times \left(K_f\left(\frac{v-\hat{V}_{pl}}{h_f}, \frac{x-X_l}{h_f}\right) - K_f\left(\frac{v-V_{pl}}{h_f}, \frac{x-X_l}{h_f}\right)\right).
\end{aligned}$$

Next, a second-order Taylor expansion gives

$$\begin{aligned}
(23)\qquad &|\hat{f}(v,x) - \tilde{f}(v,x)|\\
&\quad \leq \left| \frac{1}{Lh_f^{d+1}} \sum_{l=1}^{L} \frac{1}{I_l} \sum_{p=1}^{I_l} \mathbb{1}_{\mathscr{C}''(V)}(V_{pl}, X_l)\left(\hat{V}_{pl} - V_{pl}\right)\right.
\end{aligned}$$

$$\times \frac{1}{h_f}\frac{\partial K_f}{\partial v}\left(\frac{v-V_{pl}}{h_f},\frac{x-X_l}{h_f}\right)\Bigg|$$

$$+\frac{1}{2Lh_f^{d+1}}\sum_{l=1}^{L}\frac{1}{I_l}\sum_{p=1}^{I_l}\mathbb{1}_{\mathscr{C}''(V)}(V_{pl},X_l)\left(\hat{V}_{pl}-V_{pl}\right)^2$$

$$\times \frac{1}{h_f^2}\sup_v\left|\frac{\partial^2 K_f}{\partial v^2}\left(v,\frac{x-X_l}{h_f}\right)\right|$$

$$\leq \frac{\sup_{p,l}\mathbb{1}_{\mathscr{C}''(V)}(V_{pl},X_l)|\hat{V}_{pl}-V_{pl}|}{h_f}\frac{1}{Lh_f^{d+1}}$$

$$\times\sum_{l=1}^{L}\frac{1}{I_l}\sum_{p=1}^{I_l}\left|\frac{\partial K_f}{\partial v}\left(\frac{v-V_{pl}}{h_f},\frac{x-X_l}{h_f}\right)\right|$$

$$+\frac{\sup_{p,l}\mathbb{1}_{\mathscr{C}''(V)}(V_{pl},X_l)|\hat{V}_{pl}-V_{pl}|^2}{2h_f^3}\frac{1}{Lh_f^d}$$

$$\times\sum_{l=1}^{L}\sup_v\left|\frac{\partial^2 K_f}{\partial v^2}\left(v,\frac{x-X_l}{h_f}\right)\right|$$

$$\leq O\left(\left(\frac{\log L}{L}\right)^{R/(2R+d+3)}\right)\frac{1}{Lh_f^{d+1}}$$

$$\times\sum_{l=1}^{L}\frac{1}{I_l}\sum_{p=1}^{I_l}\left|\frac{\partial K_f}{\partial v}\left(\frac{v-V_{pl}}{h_f},\frac{x-X_l}{h_f}\right)\right|$$

$$+O\left(\left(\frac{\log L}{L}\right)^{(2R-1)/(2R+d+3)}\right)\frac{1}{Lh_f^d}$$

$$\times\sum_{l=1}^{L}\sup_v\left|\frac{\partial^2 K_f}{\partial v^2}\left(v,\frac{x-X_l}{h_f}\right)\right|,$$

by Proposition 3-(ii) and the definition of h_f in A4-(ii). The two sums appearing in (23) may be viewed as kernel estimators, and hence converge uniformly on $\mathscr{C}(V)$ to

$$f(v,x)\int\left|\frac{\partial K_f}{\partial v}(u,y)\right|du\,dy \quad\text{and}\quad f(x)\int\sup_v\left|\frac{\partial^2 K_f}{\partial^2 v}(v,y)\right|dy,$$

respectively. Thus they stay bounded almost surely. Since $R\geq 1$ implies $(2R-1)/(2R+d+3)\geq R/(2R+d+3)$, it follows that $\hat{f}(v,x)-\tilde{f}(v,x)=O(\log L/L)^{R/(2R+d+3)}$. *Q.E.D.*

The above proof shows that, under our bandwidth choice A4, $\tilde{f}(v,x)-f(v,x)$ and the first-order term in the Taylor expansion of $\hat{f}(v,x)-\tilde{f}(v,x)$ have the same order of magnitude, namely $O((\log L/L)^{R/(2R+d+3)})$, while the second-order term is smaller. In fact, given our choice for h_G and h_g, it can be shown that our choice for h_f, which implies oversmoothing, is the only choice that achieves the optimal uniform convergence rate $(L/\log L)^{R/(2R+d+3)}$ for estimating the density $f(v,x)$ for any combination (d,R) of dimension and smoothness. In particular, the standard "optimal" choice $(\log L/L)^{1/(2R+d+1)}$ for h_f would lead to a suboptimal uniform convergence rate for estimating $f(v,x)$. There may, however, exist alternative bandwidth choices for (h_G,h_g,h_f) leading to the optimal uniform convergence rate $(L/\log L)^{R/(2R+d+3)}$. Our choice for h_G and h_g, which corresponds to the optimal bandwidth rates for estimating $G(\cdot,\cdot,i)$ and $g(\cdot,\cdot,i)$ (see Hardle (1991)) have the advantage of providing an estimate of the inverse equilibrium strategy $\xi(\cdot,\cdot,i)$ for every i with the best uniform convergence rate.

Theorem 3 is useful in practice only if the support $S(f(v,x))$ is known so that one can choose appropriately the compact set $\mathscr{C}(V)$. This support can be estimated. Specifically, because $\underline{v}(x)=\underline{b}(x)$, then the lower bound $\underline{v}(\cdot)$ is estimated from observed bids using (17) directly. Regarding the upper bound $\bar{v}(\cdot)$, we can use an estimator of the form (17) with inf and B_{pl} replaced by sup and $\hat{V}_{pl}$, respectively. Thus from Proposition 3-(ii) and a proof analogous to that of Proposition 2, it can be readily shown that the support $S(f(v,x))$ can be estimated uniformly.

Lastly, asymptotic normality of our estimator is not covered by Theorem 3. The proof of Theorem 3 indicates that classical asymptotic normality results are likely to be imprecise because only the leading term in the expansion of $\hat{f}(v,x)-f(v,x)$ is used. In our case, the second order term of the Taylor expansion can be close to the first one, especially if the degree of smoothness R is small. Such drawbacks can be circumvented by establishing an exponential-type inequality for all L of the form

$$\Pr\left(\left(\frac{L}{\log L}\right)^{R/(2R+d+3)} \sup_{(v,x)\in\mathscr{C}(V)} |\hat{f}(v|x)-f(v|x)| > e_f^*(t)\right) \le P_f^*(t), \tag{24}$$

for $t>0$ and some positive functions $e_f^*(t)$ and $P_f^*(t)$, analogous to the inequalities obtained in Lemma C4. Though we shall not pursue such an issue, which is outside the scope of the paper, such an inequality is useful in practice for two reasons.

First, (24) can be used to obtain conservative (uniform) confidence intervals. Specifically, choose t such that $P_f^*(t)\le\alpha$. Hence the probability that $f(v|x)$ is in the interval

$$\left[\hat{f}(v|x)-e_f^*(t)(\log L/L)^{R/(2R+3)}, \hat{f}(v|x)+e_f^*(t)(\log L/L)^{R/(2R+3)}\right]$$

$$\text{for all } (v,x)\in\mathscr{C}(V)$$

is larger than $1-\alpha$. Second, (24) can be used to choose the constants λ appearing in the definitions of the various bandwidths. Indeed, Fubini's Theorem yields

$$E\left[\sup_{(v,x)\in\mathscr{C}(V)}|\hat{f}(v|x)-f(v|x)|\right]$$
$$=\int_0^{+\infty}\Pr\left(\sup_{(v,x)\in\mathscr{C}(V)}|\hat{f}(v|x)-f(v|x)|>t\right)dt.$$

Combined with (24), this gives an upper bound for $E[\sup_{(v,x)\in\mathscr{C}(V)}|\hat{f}(v|x)-f(v|x)|]$, which can be used to assess the choice of the λ's based on the expected supnorm loss function.

4. *Reservation Price and Number of Bidders*

Up to now, we have assumed that the reservation price is nonbinding. In practice, the seller may announce a reservation price sufficiently high prior to the bidding as a screening device for participating in the auction. Though the Bayesian Nash equilibrium strategy is still given by (1), a binding reservation price raises a new difficulty due to the unobserved number I of potential bidders. This number is typically different from the observed number I^* of actual bidders who have proposed a bid ($\geq p_0$). Hence there is a new structural element, namely I, in addition to the latent distribution of bidders' private values $F(\cdot)$. In this section we show how our results can be extended to this situation.

4.1. *Nonparametric Identification*

As in Section 2.2, we first consider one auction only and hence we omit the subscript l. Alternatively, our reasoning can be viewed as conditional upon (x, p_0, I). Following the derivation leading to (3), we rewrite the differential equation (2) in terms of observables. Unlike Section 2, however, a binding reservation price p_0 introduces a truncation because a potential bidder with a private value lower than p_0 does not bid. Let b_i^* denote the equilibrium bid of the ith actual bidder, $i=1,\ldots,I^*$, and $G^*(\cdot)$ be its distribution. Thus $G^*(b^*)=\Pr(s(\tilde{v})\leq b^*|\tilde{v}\geq p_0)=[F(v)-F(p_0)]/[1-F(p_0)]$ where $v=s^{-1}(b^*)$. Differentiating with respect to b^* gives the conditional density $g^*(b^*)=(1/s'(v))(f(v)/(1-F(p_0)))$. Hence, from (2) elementary algebra gives

$$\text{(25)}\qquad v_i=\xi(b_i^*,G^*,I,F(p_0))$$
$$\equiv b_i^*+\frac{1}{I-1}\left(\frac{G^*(b_i^*)}{g^*(b_i^*)}+\frac{F(p_0)}{1-F(p_0)}\frac{1}{g^*(b_i^*)}\right),$$

for $i = 1, \ldots, I^*$. Equation (25) is the analog of (3), but involves I and $F(p_0)$, which are unknown. This complicates the identification and estimation of the model.

The next result solves the identification problem.

THEOREM 4: *Let $G^*(\cdot) \in \mathscr{P}$ with support $[p_0, \bar{b}]$, and $\pi(\cdot)$ be a discrete distribution. There exist a distribution of buyers' private values $F(\cdot) \in \mathscr{P}$ and a number $I \geq 2$ of potential buyers such that (i) $G^*(\cdot)$ is the truncated distribution of the equilibrium bid in a first-price sealed-bid auction with reservation price $p_0 \in (\underline{v}, \bar{v})$ and (ii) $\pi(\cdot)$ is the distribution of the number of actual bidders I^* if and only if the following conditions hold*:

C1: $\pi(\cdot)$ is Binomial with parameters $(I, 1 - F(p_0))$, where $0 < F(p_0) < 1$.

C2: The observed bids are i.i.d. as $G^*(\cdot)$ conditionally upon I^* and $\lim_{b \downarrow p_0} g^*(b) = +\infty$.

C3: The function $\xi(\cdot, G^*, I, F(p_0))$ defined in (25) is strictly increasing on $[p_0, \bar{b}]$ and its inverse is differentiable on $[\underline{v}, \bar{v}] \equiv [\xi(p_0, G^*, I, F(p_0)), \xi(\bar{b}, G^*, I, F(p_0))]$.

Moreover, if Conditions C1–C3 hold, then I and $F(p_0)$ are unique while $F(\cdot)$ is uniquely defined on $[p_0, \bar{v}]$ as $F(\cdot) = F(p_0) + [1 - F(p_0)] G^*(\xi^{-1}(\cdot, G^*, I, F(p_0)))$. In addition, $\xi(\cdot, G^*, I, F(p_0))$ is the quasi inverse of the equilibrium strategy $s(\cdot, F, p_0, I)$ in the sense that $\xi(b, G^*, I, F(p_0)) = s^{-1}(b, F, p_0, I)$ for all $b \in [p_0, \bar{b}]$.

Theorem 4 parallels Theorem 1. In particular, it shows that the game theoretical model imposes some restrictions on the joint distribution of the observables $(b_1^*, \ldots, b_{I^*}^*, I^*)$. Moreover, the number of potential bidders I is identified, while the latent private values distribution $F(\cdot)$ is identified nonparametrically on $[p_0, \bar{v}]$.

4.2. *Nonparametric Estimation*

As in Section 3 we consider heterogeneous auctioned objects characterized by X_l with the difference that the reservation price is now binding. The next assumption clarifies the nature of the number of potential bidders and the reservation price.

ASSUMPTION A5:

(i) *The number of potential bidders $I \geq 2$ is constant.*

(ii) *The reservation price P_0 is a possibly unknown deterministic $(R+1)$ continuously differentiable function $h(\cdot)$ of the characteristics X.*

(iii) *For some $\epsilon > 0$, $\underline{v}(x) + \epsilon \leq h(x) \leq \bar{v}(x) - \epsilon$ for all $x \in [\underline{x}, \bar{x}]$.*

By A5-(i), the size of the market is assumed constant across auctions. Though known to potential bidders in the theoretic model, I is unknown to the analyst, as is typically the case.[23] Assumption A5-(ii) is suited when the seller determines the reservation price as a function of the object's characteristics, as it is the case in practice. In particular, this is satisfied when the seller's private value V_0 is a function of the object's characteristics, and the reservation price is determined optimally (see Riley and Samuelson (1981)).[24] Assumption A5-(iii) requires that the reservation price be bounded away from $\underline{v}(x)$ and $\bar{v}(x)$. This ensures that the reservation price is always binding and that there is always a positive probability of having at least one actual bidder. For instance, given A2, A5-(iii) is satisfied when the reservation price is chosen optimally and $\underline{v}(X) \leq V_0 \leq \bar{v}(X)$.

Given A5-(ii), the conditional distribution of an observed bid B_{pl}^* given (X_l, P_{0l}) is the conditional distribution given X_l only, namely $G^*(\cdot|X_l)$ with support $[h(X_l), \bar{b}(X_l)]$. Thus, for $p = 1, \ldots, I_l^*$ and $l = 1, \ldots, L$, (25) becomes

$$(26) \qquad V_{pl} = \xi(B_{pl}^*, X_l) \equiv B_{pl}^* + \frac{1}{I-1}\left(\frac{G^*\left(B_{pl}^*|X_l\right)}{g^*\left(B_{pl}^*|X_l\right)} + \frac{\Phi(X_l)}{1-\Phi(X_l)}\frac{1}{g^*\left(B_{pl}^*|X_l\right)}\right),$$

where $\Phi(X_l) \equiv F(P_{0l}|X_l)$. Provided one can estimate I and $\Phi(X_l)$, this equation can be used to develop a two-step estimation procedure analogous to that of Section 3. A technical difficulty arises as the density $g^*(\cdot|X_l)$ is unbounded at $B^* = P_{0l}$ (see C2 of Theorem 4). This is because $s^*(P_{0l}, X_l) = 0$. Hence Theorems 2 and 3 no longer apply.

In fact, the density $g^*(\cdot|X_l)$ behaves as $1/\sqrt{b^* - P_{0l}}$ in the neighborhood of P_{0l} since the behavior of $s(\cdot, X_l) - P_{0l}$ is quadratic in the neighborhood of P_{0l}. This leads to considering the transformation $B_\dagger = (B^* - P_0)^{1/2}$. Hence $B_\dagger = s_\dagger(V, X)$, where $s_\dagger(V, X) = [s(V, X) - P_0]^{1/2}$ with $V \geq P_0$ and $P_0 = h(X)$. Thus the distribution of $B_\dagger$ is $G_\dagger(b_\dagger|X) = G^*(P_0 + b_\dagger^2|X)$ with bounded density $g_\dagger(b_\dagger|X) = 2b_\dagger g^*(P_0 + b_\dagger^2|X)$ on its support $[0, \bar{b}_\dagger(X)] = [0, (\bar{b}(X) - P_0)^{1/2}]$. Under A5-(ii), (26) becomes

$$(27) \qquad V_{pl} = \xi_\dagger(B_{\dagger pl}, X_l) \equiv P_{0l} + B_{\dagger pl}^2 + \frac{2B_{\dagger pl}}{I-1}\left(\frac{G_\dagger(B_{\dagger pl}, X_l)}{g_\dagger(B_{\dagger pl}, X_l)} + \frac{\Phi(X_l)}{1-\Phi(X_l)}\frac{f(X_l)}{g_\dagger(B_{\dagger pl}, X_l)}\right),$$

using $G_\dagger(\cdot|\cdot)/g_\dagger(\cdot|\cdot) = G_\dagger(\cdot,\cdot)/g_\dagger(\cdot,\cdot)$ and $1/g_\dagger(\cdot|\cdot) = f(\cdot)/g_\dagger(\cdot,\cdot)$.

[23] In fact, we only need that I be constant on some known subsets of auctions, in which case our estimation method applies to each subset. Tests of the constancy of I can be based on estimates of I (see below) for each subset. Because of A5-(i), hereafter I is dropped as an argument from any function.

[24] If there are many sellers across auctions, A5-(ii) may no longer hold as $h(\cdot)$ may vary across sellers. Thus P_0 given X can be viewed as stochastic, which leads to a random truncation given X (see Wang, Jewell, and Tsai (1986)). It can be shown that our two-step estimator is still uniformly consistent on compact subsets at the rate $(L/\log L)^{R/(2R+d+4)}$.

PROPOSITION 4: *Given A2 and A5, the conditional distribution* $G_\dagger(\cdot|\cdot)$ *satisfies properties (i)–(iv) of Proposition* 1, *where* $G(\cdot|\cdot,\cdot)$, $S_i(G)$, $\underline{b}(x,i)$, $\overline{b}(x,i)$, *and* $\mathscr{C}_i(B)$ *are replaced respectively by* $G_\dagger(\cdot|\cdot)$, $S(G_\dagger)=\{(b_\dagger,x): x\in[\underline{x},\bar{x}],\ b_\dagger\in[0,\overline{b}_\dagger(x)]\}$, 0, $\overline{b}_\dagger(x)$, *and* $\mathscr{C}(B_\dagger)$, *which is a closed inner subset of* $S(G_\dagger)$.

Proposition 4 is important as it is analogous to Proposition 1, which is the backbone of the optimal rate and uniform consistency of our estimator (Theorems 2 and 3).

Analogously to Section 3, (27) suggests a two-step estimation procedure. In the first step, $f(\cdot)$ is estimated by (21), while $G_\dagger(\cdot,\cdot)$ and $g_\dagger(\cdot,\cdot)$ are estimated by

$$\tilde{G}_\dagger(b_\dagger,x)=\frac{1}{Lh_G^d}\sum_{l=1}^{L}\frac{1}{I_l^*}\sum_{p=1}^{I_l^*}\mathbb{1}(B_{\dagger pl}\le b_\dagger)K_G\left(\frac{x-X_l}{h_G}\right),$$

$$\tilde{g}_\dagger(b_\dagger,x)=\frac{1}{Lh_g^{d+1}}\sum_{l=1}^{L}\frac{1}{I_l^*}\sum_{p=1}^{I_l^*}K_g\left(\frac{b_\dagger-B_{\dagger pl}}{h_g},\frac{x-X_l}{h_g}\right).$$

This follows because the transformed actual bids $(B^*_{\dagger 1l},\dots,B^*_{\dagger I_l^* l})$ are independent of I_l^* conditionally upon X_l in view of Condition C2 of Theorem 4.

In the second step, (27) is used to recover the pseudo private values provided one can estimate I and $\Phi(X_l)$. A natural estimator for I is $\hat{I}=\max_{l=1,\dots,L} I_l^*$. Given Condition C1 of Theorem 4, $\hat{I}=I$ almost surely. To estimate $\Phi(X_l)$, we note that $E[I_l^*|X_l]=I[1-\Phi(X_l)]$. Solving for $\Phi(X_l)$ suggests to estimate it by $\hat{\Phi}(X_l)$, where

$$\hat{\Phi}(x)=1-\frac{1}{\hat{I}Lh_x^d\hat{f}(x)}\sum_{l=1}^{L}I_l^*K_x\left(\frac{x-X_l}{h_x}\right),$$

using the usual Nadaraya (1964)–Watson (1964) nonparametric regression estimator with $\hat{f}(x)$ as in (21). Note that $\hat{\Phi}(\cdot)$ is always between 0 and 1.

Let $\tilde{\xi}_\dagger(\cdot,\cdot)$ be given by (27), where I, $\Phi(\cdot)$, $f(\cdot)$, $G_\dagger(\cdot,\cdot)$, and $g_\dagger(\cdot,\cdot)$ are replaced by their estimators. Following (16), let $\hat{\overline{b}}_\dagger(x)=\sup\{B_{\dagger pl},\ p=1,\dots,I_l^*,\ l=1,\dots,L,\ X_l\in\pi_{k_1,\dots,k_d}\}$. The estimated support of $G_\dagger(\cdot,\cdot)$ is $\hat{S}(G_\dagger)\equiv\{(b,x): b\in[0,\hat{\overline{b}}_\dagger(x)], x\in[\underline{x},\bar{x}]\}$. The pseudo private values are given by $\hat{V}_{pl}=\hat{\xi}_\dagger(B_{\dagger pl},X_l)$, where

$$\hat{\xi}_\dagger(b_\dagger,x)\equiv\begin{cases}\tilde{\xi}_\dagger(b_\dagger,x) & \text{if } (b_\dagger,x)+S(2h_G)\subset\hat{S}(G_\dagger),\\ & (b_\dagger,x)+S(2h_g)\subset\hat{S}(G_\dagger),\\ & x+S(2h_f)\in[\underline{x},\bar{x}],\\ +\infty & \text{otherwise},\end{cases}$$

for $p=1,\dots,I_l^*$, $l=1,\dots,L$. The nonparametric estimator of the conditional density $f(v|x)$ is $\hat{f}(v|x)\equiv[1-\hat{\Phi}(x)]\hat{f}^*(v|x)$, where $\hat{f}^*(v|x)$ is the estimator of the truncated conditional density $f^*(v|x)=f(v|x)/[1-\Phi(x)]$ and is obtained as in Section 3 with I_l replaced by I_l^* in (20). Let $S^* \equiv \{(v, x): x \in [\underline{x}, \bar{x}], v \in [h(x), \bar{v}(x)]\}$ be the support of $f^*(\cdot|\cdot)$. The next result gives the asymptotic properties of the estimator $\hat{f}(\cdot|\cdot)$.

THEOREM 5: *Suppose that $A1$–$A5$ hold. Then $\hat{f}(\cdot|\cdot)$ is uniformly consistent with optimal rate $(L/\log L)^{R/(2R+d+3)}$ on any closed inner subset of S^* with nonempty interior.*

In view of Theorem 4 and A5-(ii), S^* is the largest subset of the support of $f(\cdot|\cdot)$ where the latter is identified.

5. CONCLUSION

For first-price sealed-bid auctions with binding or nonbinding reservation prices, we have shown that the underlying distribution of bidders' private values within the independent private values paradigm is identified from observables, namely, observed bids and the number of actual bidders, without any parametric assumptions. Moreover, using the recently developed minimax theory, we have established the best rate of uniform convergence at which the latent density of private values can be estimated from available data. We then have proposed a computationally convenient two-step nonparametric estimator of this density that converges at this optimal rate.[25]

As a matter of fact, our results go well beyond the auction mechanism and paradigm studied here as they can be generalized to other auction models as considered by Milgrom and Weber (1982). For instance, as shown by Guerre, Perrigne, and Vuong (1995) for the independent private value paradigm, they apply to descending or Dutch auctions, where only the winning bid, if any, is observed. Alternatively, our results can be extended to the affiliated private value paradigm, which is the most general auction model that can be identified from observed bids (see Laffont and Vuong (1996)). Li, Perrigne, and Vuong (1996, 1999) show how our two-step nonparametric procedure can be modified accordingly despite the complexity of the Bayesian Nash equilibrium strategy.

More generally, because the determination of the equilibrium strategy is avoided, our indirect estimation procedure is especially convenient when the equilibrium strategy cannot be obtained explicitly. In such situations, direct (parametric) estimation methods become cumbersome, if not computationally

[25]Alternatively, our general indirect two-step principle can also be used to estimate parametrically the distribution of observed bids in the first step and the distribution of private values in the second step. Besides some technical difficulties arising from the relationship between $G(\cdot)$ and $F(\cdot)$, such a parametric two-step procedure would constitute a powerful alternative to existing parametric methods such as the simulation-based method proposed in Laffont, Ossard, and Vuong (1995). Specifically, this new parametric method will be computationally much easier when the equilibrium strategy cannot be easily simulated or cannot be solved explicitly as in asymmetric auctions.

infeasible, since the differential equation must be solved numerically for the equilibrium strategy for each trial value of the parameters. An example of such a situation arises with a random reservation price. This occurs when the seller adopts a random rejection rule of the highest bid, or when the reservation price is announced after bidding. Within the independent private values paradigm, Elyakime, Laffont, Loisel, and Vuong (1994, 1997) extend our results to the latter situation. Another important case arises when bidders are asymmetric ex ante due to joint bidding, informational advantages such as in Outer Continental Shelf auctions (see Hendricks and Porter (1988), Hendricks, Porter, and Wilson (1994)), the large size of a bidder, and more generally collusion among some bidders (see Graham and Marshall (1987)). In this case, the equilibrium strategies are solutions of a system of differential equations that cannot be solved explicitly (see Maskin and Riley (1983)). Using results similar to ours, Campo, Perrigne, and Vuong (1998) and Laffont, Li, and Vuong (1999) show how to circumvent this difficulty. Asymmetry also arises in procurement actions when firms bid in both price and quality. Applying our approach, Laffont, Oustry, Simioni, and Vuong (1996) identify and estimate the resulting model.

Lastly, an important feature of our identification results is that they provide necessary and sufficient conditions for the existence of a latent distribution that can "rationalize" the observed bid distribution. Because they are the restrictions imposed by the game theoretic model on observables, these conditions can constitute the basis of a test of the theory. They are of two types. The first type deals with independence of bids (see conditions C1 and C2 of Theorems 1 and 4, respectively), which relates to the paradigm and can be tested nonparametrically using, e.g., the Blum, Kiefer, and Rosenblatt (1961) test. The second type deals with the monotonicity of an estimable function (see conditions C2 and C3 of Theorem 1 and 4, respectively), which can be used to test whether bidders adopt the symmetric Bayesian Nash equilibrium strategy. A major difficulty in developing a nonparametric test, however, is to estimate the function $\xi(\cdot)$ under monotonicity constraint. Though progress has been made in the regression context (see, e.g., Wright and Wegman (1980) and Utreras (1985)), the different nature of $\xi(\cdot)$ calls for further work.

To conclude, by proposing a general and computationally convenient estimation principle, this paper contributes to the structural analysis of auction data that was plagued by the numerical complexities associated with the equilibrium strategies. The estimation of auction models can also be considered as a first step in the estimation of asymmetric information models used in the theory of regulation and contracts. Hence our paper contributes to the agenda of the new empirical industrial organization calling for the data evaluation of game theoretic models (see Laffont and Tirole (1993) and Salanié (1997)).

Université de Jussieu, LSTA-ISUP, Boite 158, 4, Place Jussieu, 75252 Paris Cedex 05, France, *Dept. of Economics, University of Southern California, Los Angeles, CA 90089-0253, U.S.A.*

and

Dept. of Economics, University of Southern California, Los Angeles, CA 90089-0253, U.S.A.

Manuscript received April, 1995; final revision received April, 1998.

APPENDIX A

Proofs of Mathematical Properties

This Appendix gives (i) the proofs of our identification results (Theorems 1 and 4) and (ii) establishes the model's regularity properties (Propositions 1 and 4), when the reservation price is nonbinding and binding respectively. Lemmas are proved in Appendix C.

A.1. *Proofs of Identification Results*

Proof of Theorem 1: First, we prove that Conditions C1 and C2 are necessary. Because $b_i = s(v_i, F, I)$ and the v_i's are i.i.d., it follows that the b_i's are i.i.d. so that Condition C1 must hold. To prove Condition C2, let $s(\cdot, F, I)$ be the strictly increasing differentiable and symmetric Bayesian Nash equilibrium strategy corresponding to $F(\cdot)$ with support $[\underline{v}, \bar{v}]$ (say). Let $G(\cdot)$ be the distribution defined by $G(b) = F(s^{-1}(b, F, I))$ for every $b \in [\underline{b}, \bar{b}] \equiv [s(\underline{v}, F, I), s(\bar{v}, F, I)]$. Note that $G(\cdot)$ must be the distribution of observed (equilibrium) bids. Now, $s(\cdot, F, I)$ must solve the first-order differential equation (2). But, because (3) follows from (2), then $s(\cdot, F, I)$ must satisfy $\xi(s(v, F, I), G, I) = v$ for all $v \in [\underline{v}, \bar{v}]$. Making the change of variable $b = s(v, F, I)$, we obtain $\xi(b, G, I) = s^{-1}(b, F, I)$ for all $b \in [\underline{b}, \bar{b}]$. Hence Condition C2 must hold because $s^{-1}(\cdot, F, I)$ is strictly increasing on $[\underline{b}, \bar{b}]$ and $s(\cdot, F, I)$ differentiable on $[\underline{v}, \bar{v}] = [\xi(\underline{b}, G, I), \xi(\bar{b}, G, I)]$.

To prove sufficiency, let $\mathbf{G}(\cdot)$ belong to $\mathscr{P}^I$ with support $[\underline{b}, \bar{b}]^I$. By Condition C1, the b_i's are i.i.d., each distributed as some $G(\cdot)$ with support $[\underline{b}, \bar{b}]$. Now note that

$$\text{(A.1)} \qquad \lim_{b \downarrow \underline{b}} \xi(b, G, I) = \underline{b}.$$

This follows from (3) and the fact that (i) $\underline{b}$ is finite (≥ 0), (ii) $\lim_{b \downarrow \underline{b}} \log G(b) = -\infty$, and (iii) $g(b)/G(b) = d \log G(b)/db$, so that $\lim_{b \downarrow \underline{b}} g(b)/G(b) = +\infty$.

Next define $F(\cdot) \equiv F(\xi^{-1}(\cdot, G, I))$ on $[\underline{v}, \bar{v}]$, where $\underline{v} \equiv \xi(\underline{b}, G, I) = \underline{b}$ by (A.1) and $\bar{v} \equiv \xi(\bar{b}, G, I)$. Thus $F(\cdot)$ is a valid distribution because $\xi(\cdot, G, I)$ is strictly increasing on $[\underline{b}, \bar{b}]$ by Condition C2. Moreover, because $G(\cdot)$ is strictly increasing on $[\underline{b}, \bar{b}]$ (see footnote 3), then $F(\cdot)$ is strictly increasing on $[\underline{v}, \bar{v}]$. Thus the support of $F(\cdot)$ is $[\underline{v}, \bar{v}]$, which is an interval of $\mathbb{R}_+$. Lastly, because $\xi^{-1}(\cdot, G, I)$ is differentiable and $G(\cdot)$ is absolutely continuous, then $F(\cdot)$ is absolutely continuous. Therefore $F(\cdot)$ belongs to $\mathscr{P}$ as required.

It remains to show that this distribution $F(\cdot)$ of buyers' private values can rationalize $G(\cdot)$ in a first-price sealed-bid auction with no reservation price, i.e., that $G(\cdot) = F(s^{-1}(\cdot, F, I))$ on $[\underline{b}, \bar{b}]$, where $s(\cdot, F, I)$ solves (2) with the boundary condition $s(\underline{v}, F, I) = \underline{v}$. By construction of $F(\cdot)$, we have $G(\cdot) = F(\xi(\cdot, G, I))$. Thus it suffices to show that $\xi^{-1}(\cdot, G, I)$ solves (1) with boundary condition $\xi^{-1}(\underline{v}, G, I) = \underline{v}$. It is easy to see that the boundary condition is satisfied. From the construction of $F(\cdot)$ note that $f(\cdot)/F(\cdot) = \xi^{-1'}(\cdot, G, I) g(\xi^{-1}(\cdot, G, I))/G(\xi^{-1}(\cdot, G, I))$ after differentiating and taking a ratio. Thus $\xi^{-1}(\cdot,G,I)$ solves (2) if

$$(A.2)\qquad 1=(v-\xi^{-1}(v,G,I))(I-1)\frac{g(\xi^{-1}(v,G,I))}{G(\xi^{-1}(v,G,I))}\quad \forall v\in[\underline{v},\bar{v}].$$

But (A.2) clearly holds by definition of $\xi(\cdot,G,I)$. This completes the proof of sufficiency.

It remains to establish the last part of Theorem 1. From the proof of necessity, we know that $\xi(\cdot,G,I)=s^{-1}(\cdot,F,I)$ holds when $F(\cdot)$ exists. Since $F(\cdot)=G(s(\cdot,F,I))$, then $F(\cdot)=G(\xi^{-1}(\cdot,G,I))$. Because $\xi(\cdot,G,I)$ is uniquely determined by $G(\cdot)$, it follows that $F(\cdot)$ is unique given $G(\cdot)$. Moreover, its support must be $[\xi(\underline{b},G,I),\xi(\bar{b},G,I)]=[\underline{b},\xi(\bar{b},G,I)]$ by (A.1). *Q.E.D.*

Proof of Theorem 4: Consider necessity. Condition C1 must hold since a potential bidder bids if and only if his private value is at least p_0. Thus $0<F(p_0)<1$ follows from $p_0\in(\underline{v},\bar{v})$.

Regarding Condition C2, let $B_j=0$ if $V_j<p_0$ for $j=1,\dots,I$. For any $(b_1,\dots,b_i)\in\mathbb{R}_+^i$,

$$\begin{aligned}
&\Pr(B_1^*\le b_1,\dots,B_i^*\le b_i,\ I^*=i|I,p_0)\\
&\quad=\sum_{1\le r_1\ne\cdots\ne r_i\le I}\Pr(p_0\le B_{r_1}\le b_1,\dots,p_0\le B_{r_i}\le b_i,\ B_j=0,\ j\notin\{r_1,\dots,r_i\}|I,p_0)\\
&\quad=\frac{I!}{i!(I-i)!}\Pr(p_0\le B_1\le b_1,\dots,p_0\le B_i\le b_i,\ B_{i+1}=\cdots=B_I=0|I,p_0)\\
&\quad=\frac{I!}{i!(I-i)!}\Pr{}^{I-i}(B_{i+1}=0|I,p_0)\prod_{r=1}^{i}\Pr(p_0\le B_r\le b_r|I,p_0),
\end{aligned}$$

because B_j's are iid given (I,p_0). Since I^* is binomial with parameters $(I,1-F(p_0))$, then

$$\begin{aligned}
\Pr(B_1^*\le b_1,\dots,B_i^*\le b_i|I^*=i,\ I,p_0)&=\prod_{r=1}^{i}\frac{\Pr(p_0\le B_r\le b_r|I,p_0)}{1-F(p_0)}\\
&=\prod_{r=1}^{i}\frac{F(s^{-1}(b_r))-F(p_0)}{1-F(p_0)},
\end{aligned}$$

as desired. Moreover, because $s(p_0)=p_0$ and $F(p_0)>0$ it follows from (2) that $\lim_{v\downarrow p_0}s'(v)/f(v)=0$. Since $g^*(b)=f(s^{-1}(b))/[s'(s^{-1}(b))(1-F(p_0)]$, then $\lim_{b\downarrow p_0}g^*(b)=+\infty$. Regarding the necessity of Condition C3, the proof is similar to that of Condition C2 in Theorem 1, noting that $\xi(\cdot,G^*,I,F(p_0))=s^{-1}(\cdot,F,I,p_0)$ on $[p_0,\bar{b}]=[p_0,s(\bar{v},F,I,p_0)]$.

Turning to sufficiency, choose $\underline{v}<p_0$ and $0<F(p_0)<1$. The proof then follows that of Theorem 1 by replacing $G(\cdot)$ and $F(\cdot)$ by the truncated distributions $G^*(\cdot)$ and $F^*(\cdot)\equiv[F(\cdot)-F(p_0)]/[1-F(p_0)]$ defined on $[p_0,\bar{b}]$ and $[p_0,\bar{v}]$, respectively. The condition $\lim_{b\downarrow p_0}\xi(b,G^*,I,F(p_0))=p_0$, which is analogous to (A.1), follows now from (25) and $\lim_{b\downarrow p_0}g^*(b)=+\infty$. In particular, if Conditions C1–C3 hold, we can establish that there exists a distribution $F(\cdot)$ on $[\underline{v},\bar{v}]$ with $\bar{v}=\xi(\bar{b},G^*,I,F(p_0))$ that (i) is absolutely continuous on $[p_0,\bar{v}]$ and (ii) can rationalize $G^*(\cdot)$ in a first-price sealed-bid auction with reservation price p_0. A distribution $F(\cdot)\in\mathscr{P}$ is obtained by extension in an absolute continuous fashion to the interval $[\underline{v},p_0]$.

To prove the last part of Theorem 4, note that I and $F(p_0)$ are identified from $\pi(\cdot)$. Moreover, similar to the end of the proof of Theorem 1, we have $\xi(\cdot,G^*,I,F(p_0))=s^{-1}(\cdot,F,I,p_0)$ on $[p_0,\bar{b}]$, and the truncated distribution $F^*(\cdot)\equiv[F(\cdot)-F(p_0)]/[1-F(p_0)]$ is unique and equal to $G^*(\xi^{-1}(\cdot,G^*,I,F(p_0)))$ with support $[p_0,\bar{v}]=[p_0,\xi(\bar{b},G^*,I,F(p_0))]$. Because I and $F(p_0)$ are identified, it follows that $F(\cdot)$ is uniquely determined on $[p_0,\bar{v}]$ and equal to $F(p_0)+[1-F(p_0)]G^*(\xi^{-1}(\cdot,G^*,I,F(p_0)))$ on this interval. *Q.E.D.*

A.2. *Proofs of Regularity Properties*

To prove Proposition 1 we need two lemmas. Lemma A1 studies the regularity of the boundaries $\underline{v}(\cdot)$ and $\bar{v}(\cdot)$ as well as that of the bid function $s(\cdot,\cdot,i)$. Lemma A2 translates these regularity properties into regularity of $\underline{b}(\cdot,i)$, $\bar{b}(\cdot,i)$, and $\xi(\cdot,\cdot,i)$.

LEMMA A1: *Under A2*:

(i) *The boundaries* $\bar{v}(\cdot)$ *and* $\underline{v}(\cdot)$ *admit up to* $R+1$ *continuous bounded derivatives on* $[\underline{x}, \bar{x}]$, *and* $\inf_{x \in [\underline{x}, \bar{x}]}(\bar{v}(x) - \underline{v}(x)) > 0$.

(ii) *For each* $i \in \mathscr{I}$, $s(\cdot, \cdot, i)$ *admits up to* $R+1$ *continuous bounded partial derivatives on* $S_i(F)$. *Moreover, for any* $(v, x) \in S_i(F)$, $s'(v, x, i) \geq c_s > 0$.

LEMMA A2: *Under A2, for each* $i \in \mathscr{I}$, *we have the following*:

(i) *The boundaries* $\underline{b}(x, i)$ *and* $\bar{b}(x, i)$ *admit up to* $R+1$ *continuous bounded derivatives on* $[\underline{x}, \bar{x}]$, *and* $\inf_{x \in [\underline{x}, \bar{x}]}(\bar{b}(x, i) - \underline{b}(x, i)) > 0$.

(ii) $\xi(\cdot, \cdot, i)$ *admits up to* $R+1$ *continuous bounded partial derivatives on* $S_i(G)$. *Moreover, for any* $(b, x) \in S_i(G)$, $\xi'(b, x, i) \geq c_\xi > 0$.

PROOF OF PROPOSITION 1: (i) is Lemma A2-(i) plus the boundary condition. Next,

$$g(b|x, i) = \frac{f(\xi(b, x, i)|x, i)}{s'(\xi(b, x, i), x, i)}. \tag{A.3}$$

Because f is bounded away from 0 by A2-(ii) and s' is bounded by Lemma A1-(ii), then (ii) follows. To prove (iii), it suffices to note that $G(b|x, i) = F(\xi(b, x, i)|x, i)$, where $F(\cdot|\cdot, i)$ and $\xi(\cdot, \cdot, i)$ have $R+1$ continuous bounded derivatives on $S_i(F)$ and $S_i(G)$ by A2-(iii) and Lemma A2-(ii), respectively. Lastly, to prove (iv) we note that (9) gives

$$g(b|x, i) = \frac{G(b|x, i)}{(i-1)(\xi(b, x, i) - b)} \tag{A.4}$$

with $\xi(b, x, i) - b > 0$ for $(b, x) \in \mathscr{C}_i(B)$. Because $G(\cdot|\cdot, i)$ and $\xi(\cdot, \cdot, i)$ admit up to $R+1$ bounded continuous derivatives by (iii) and Lemma A2-(ii), the desired result follows. *Q.E.D.*

The proof of Proposition 4 follows that of Proposition 1. In particular, the following lemmas similar to Lemmas A1 and A2 are needed. As I is constant by A5-(i), then $\mathscr{I} = \{I\}$ below.

LEMMA A3: *Under A2 and A5, properties (i) and (ii) of Lemma A1 hold with* $\underline{v}(\cdot)$, $S_i(F)$ *and* $s(v, x, i)$ *replaced by* $h(\cdot)$, $S^* \equiv \{(v, x) : x \in [\underline{x}, \bar{x}], v \in [h(x), \bar{v}(x)]\}$ *and* $s_\dagger(v, x)$, *respectively*.

LEMMA A4: *Under A2 and A5, properties (i) and (ii) of Lemma A2 hold with* $\underline{b}(x, i)$, $\bar{b}(x, i)$, $S_i(G)$, *and* $\xi(b, x, i)$ *replaced by* $h(x)$, $\bar{b}_\dagger(x)$, $S(G_\dagger)$, *and* $\xi_\dagger(b_\dagger, x)$, *respectively*.

PROOF OF PROPOSITION 4: The proof is similar to that of Proposition 1, using $F^*(v|x) = [F(v|x) - F(h(x)|x)]/[1 - F(h(x)|x)]$, $G_\dagger(b_\dagger|x)$, Lemmas A3 and A4 instead of $F(v|x, i), G(b|x, i)$, Lemmas A1 and A2, respectively. In particular, instead of (A.4) we have from (27)

$$g_\dagger(b_\dagger|x) = \frac{2b_\dagger}{\xi_\dagger(b_\dagger, x) - p_0 - b_\dagger^2} \frac{1}{I-1}\left(G_\dagger(b_\dagger|x) + \frac{F(p_0|x)}{1 - F(p_0|x)}\right) \tag{A.5}$$

with $p_0 = h(x)$. *Q.E.D.*

APPENDIX B

PROOFS OF STATISTICAL PROPERTIES

This Appendix gives the proofs of Theorem 2, Propositions 2 and 3, and Theorem 5. Lemmas are proved in Appendix C. Throughout $|\cdot|_{r,*}$ denotes the supnorm of the rth derivatives of $\cdot$ on the set $*$.

B.1. *Optimal Uniform Convergence Rate*

PROOF OF THEOREM 2: The proof relies on the argument of Khas'minskii (1976), though it is somewhat more complicated due to the indirect nature of our statistical problem. The proof is divided in 4 steps, where the fourth step uses Fano's lemma. Let $r_L^* \equiv (L/\log L)^{R/(2R+d+3)}$.

Step 1: We show that the rate r_L^* is given by the estimation of the joint density $f(v,x)$. Let $\hat{f}(\cdot)$ be (say) the kernel density estimator (21) of the density $f(\cdot)$ of X. From the triangular inequality, for any estimator $\hat{f}(\cdot|\cdot)$ of the conditional density, we have

$$r_L^*|\hat{f}(v|x)-f(v|x)| \geq \frac{r_L^*}{\hat{f}(x)}(|\hat{f}(v|x)\hat{f}(x)-f(v|x)f(x)|-f(v|x)|\hat{f}(x)-f(x)|)$$

$$\geq C_1 r_L^*|\hat{f}(v|x)\hat{f}(x)-f(v,x)|-C_0 r_L^*|\hat{f}(x)-f(x)|, \quad \text{a.s.}$$

for some constants C_1 and C_0 since $f(\cdot)$ and $f(\cdot|\cdot)$ are uniformly bounded above on $\mathscr{C}(V)$ by definition of $U_\epsilon(f_0)$, and since $f(\cdot)$ converges uniformly to $f(\cdot)$. In addition, it is well-known (see Hardle (1991)) that the uniform rate of convergence of $\hat{f}(\cdot)$ to $f(\cdot)$ is $(L/\log L)^{(R+1)/(2R+d+2)}$, which is faster than r_L^*. Hence the above second term converges uniformly to zero. Now, $\hat{f}(\cdot|\cdot)\hat{f}(\cdot)$ is an estimator of the joint density $f(\cdot,\cdot)$. It is easy to show that, if Theorem 2 holds for some positive constant κ' where $\hat{f}(v,x)$ and $f(v,x)$ replace $\hat{f}(v|x)$ and $f(v|x)$, respectively, then Theorem 2 would hold for some κ.

Step 2: The set $U_\epsilon(f_0)$ can be replaced by any subset $U \subset U_\epsilon(f_0)$ since

$$\sup_{f\in U_\epsilon(f_0)} \Pr_g(r_L^*|\hat{f}(\cdot,\cdot)-f(\cdot,\cdot)|_{0,\mathscr{C}(V)}>\kappa) \geq \sup_{f\in U} \Pr_g(r_L^*|\hat{f}(\cdot,\cdot)-f(\cdot,\cdot)|_{0,\mathscr{C}(V)}>\kappa).$$

The subset U will be discrete and of the form $\{f_{mk}(\cdot,\cdot,\cdot),\ k=1,\dots,m^{d+1}\}$, where m is increasing with the sample size. To construct the joint density $f_{mk}(\cdot,\cdot,\cdot)$ for (V,X,I), we consider a nonconstant and odd C_∞-function ϕ, with support $[-1,1]^{d+1}$, such that

$$\text{(B.1)} \qquad \int_{[-1,0]} \phi(b,x)\,db=0, \quad \phi(0,0)=0, \quad \phi'(0,0)\neq 0.$$

Let $\mathscr{C}_i(B) \equiv s(\mathscr{C}(V),F_0,i)$ be the image of $\mathscr{C}(V)$ by the bidding function associated to f_0 when $I=i$. Without loss of generality, assume hereafter that $\Pr_0(I=2)\neq 0$ so that $\mathscr{C}_2(B)$ is a nonempty inner compact subset of $S_2(G_0)$. Let (b_k,x_k), $k=1,\dots,m^{d+1}$, be distinct points in the interior $C_2^o(B)$ such that the distance between (b_k,x_k) and (b_j,x_j), $j\neq k$, is larger than λ_1/m. Thus, one can choose a constant λ_2 such that the m^{d+1} functions

$$\phi_{mk}(b,x) \equiv \frac{1}{m^{R+1}}\phi(m\lambda_2(b-b_k), m\lambda_2(x-x_k)) \qquad (k=1,\dots,m^{d+1})$$

have disjoint hypercube supports $S(\phi_{mk})$ centered at (b_k,x_k) with side equal to $2/(m\lambda_2)$. Hence $\int\phi_{mj}(b,x)\phi_{mk}(b,x)\,db\,dx=0$ for $j\neq k$.

Let C_3 be a positive constant (chosen below), and for each $k=1,\dots,m^{d+1}$ define

$$g_{mk}(b,x,i) \equiv \begin{cases} g_0(b,x,2)-C_3\phi_{mk}(b,x) & \text{if } i=2,\\ g_0(b,x,i) & \text{if } i\neq 2.\end{cases}$$

That is, $g_{mk}(\cdot,\cdot,i)$ is identical to $g_0(\cdot,\cdot,i)$ except when $i=2$, in which case it differs from $g_0(\cdot,\cdot,2)$ only on the set $S(\phi_{mk})$. For m large enough, $S(\phi_{mk})$ is in $S_2^o(G_0)$, while $g(\cdot,\cdot,2)$ is bounded away from zero on $S_2(G_0)$ and integrates to $\Pr_0(I=2)$ as $g_0(\cdot,\cdot,2)$ is bounded away from zero (Proposition 1-(ii) and A2(ii)) and $\int_{[-1,+1]}\phi(b,x)\,db=0$.

Now consider the function $\xi_{mk}(b,x,2)=b+G_{mk}(b,x,2)/g_{mk}(b,x,2)$. It is easy to see that $G_{mk}(\cdot,\cdot,2)$, $g_{mk}(\cdot,\cdot,2)$, and $g'_{mk}(\cdot,\cdot,2)$ converge uniformly on $S_2(G_0)$ to $G_0(\cdot,\cdot,2)$, $g_0(\cdot,\cdot,2)$, and $g'_0(\cdot,\cdot,2)$, respectively. Thus $\xi_{mk}(\cdot,\cdot,2)$ is strictly increasing in b with a differentiable inverse. From Theorem 1, it follows that $g_{mk}(\cdot,\cdot,2)/\Pr_0(I=2)$ can be interpreted as the bid density associated with a unique density $f_{mk}(\cdot,\cdot,2)/\Pr_0(I=2)$ of (V,X) when $I=2$, namely,

$$\text{(B.2)}\qquad f_{mk}(v,x,2)=g_{mk}(\xi_{mk}^{-1}(v,x,2),x,2)/(\xi'_{mk}(\xi_{mk}^{-1}(v,x,2),x,2))$$
$$=\frac{g_{mk}^3(\xi_{mk}^{-1}(v,x,2),x,2)}{2g_{mk}^2(\xi_{mk}^{-1}(v,x,2),x,2)-G_{mk}(\xi_{mk}^{-1}(v,x,2),x,2)g'_{mk}(\xi_{mk}^{-1}(v,x,2),x,2)}.$$

As $g_{mk}(\cdot,\cdot,i)=g_0(\cdot,\cdot,i)$ for $i\neq 2$, we have $f_{mk}(\cdot,\cdot,i)=f_0(\cdot,\cdot,i)$ for those i's. This completes the construction of the densities $f_{mk}(\cdot,\cdot,\cdot)$, $m=1,\dots,m^{d+1}$, which compose the set U.

It remains to show that U is a subset of $U_\epsilon(f_0)$, i.e., that all $f_{mk}(\cdot,\cdot,\cdot)$'s belong to $U_\epsilon(f_0)$. For this, we use the following lemma.

LEMMA B1: *Given A1–A2, the following properties hold for m large enough*:

(i) *For any $k=1,\dots,m^{d+1}$, the supports of G_{mk} and F_{mk} are $S(G_0)$ and $S(F_0)$.*

(ii) *There is a positive constant C_4 depending upon ϕ, G_0, and $\mathscr{C}(V)$ such that for $j\neq k$,*

$$|f_{mk}-f_{mj}|_{0,\mathscr{C}(V)}\geq C_4\frac{C_3\lambda_2}{m^R}.$$

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

$$|f_{mk}-f_0|_{r,S(F_0)}=C_3\lambda_2^{r+1}O(1/m^{R-r})\qquad (r=0,\dots,R-1),$$
$$|f_{mk}-f_0|_{R,S(F_0)}=C_3\lambda_2^{R+1}O(1)+o(1),$$

where the big $O(\cdot)$ only depends upon ϕ and g_0.

As $\|f_0\|_R<M$ by assumption, there exists a positive constant λ_3 such that $|f_0|_{r,S(F_0)}<M-\lambda_3$ for all $r=1,\dots,R$. As $|f_{mk}|_{r,S(F_0)}\leq|f_{mk}-f_0|_{r,S(F_0)}+|f_0|_{r,S(F_0)}<|f_{mk}-f_0|_{r,S(F_0)}+M-\lambda_3$, it follows from Lemma B1-(iii) that the first term can be made smaller than λ_3 as soon as $m\geq m_0(\epsilon)$ and C_3 is small enough. Hence, f_{mk} is in $U_\epsilon(f_0)$ so that U is a subset of $U_\epsilon(f_0)$, as desired. As m is increasing with L, we have for $L\geq L_0(\epsilon)$:

$$\sup_{f\in U_\epsilon(f_0)}\Pr_g(r_L^*|\hat f(\cdot,\cdot)-f(\cdot,\cdot)|_{0,\mathscr{C}(V)}>\kappa)$$
$$\geq\max_{k=1,\dots,m^{d+1}}\Pr_{g_{mk}}(r_L^*|\hat f(\cdot,\cdot)-f_{mk}(\cdot,\cdot)|_{0,\mathscr{C}(V)}>\kappa).$$

Step 3: We now reduce our problem to the model selection problem of deciding which of the hypothesis $H_{mk}:\{f=f_{mk}\}$ holds. Consider the decision rule that selects H_{mk} whenever f_{mk} is closest to $\hat f$ in the supnorm $|\cdot|_{0,\mathscr{C}(V)}$. Let $\Pr_{g_{mk}}(\overline{H}_{mk})$ be the probability of error of this decision when $H_{mk}:\{f=f_{mk}\}$ holds. Choose m and κ such that

$$\text{(B.3)}\qquad m=(r_L^*)^{1/R}\quad\text{and}\quad \kappa=C_4C_3\lambda_2/2.$$

From Lemma B1-(ii), we have $r_L^*|f_{mk}-f_{mj}|_{0,\mathscr{C}(V)}\geq 2\kappa$ uniformly in $j\neq k$. Thus, from the triangular inequality, if $r_L^*|\hat f-f_{mk}|_{0,\mathscr{C}(V)}\leq\kappa$, then $r_L^*|\hat f-f_{mj}|_{0,\mathscr{C}(V)}\geq\kappa$, for any $j\neq k$, i.e., the decision rule of the closest accepts H_{mk}. Hence, for any estimator $\hat f(\cdot,\cdot)$ we have

$$\max_{k=1,\dots,m^{d+1}}\Pr_{g_{mk}}(r_L^*|\hat f(v,x)-f_{mk}(v,x)|_{0,\mathscr{C}(V)}>\kappa)\geq\max_{k=1,\dots,m^{d+1}}\Pr_{g_{mk}}(\overline{H}_{mk})$$
$$\geq\frac{1}{m^{d+1}}\sum_{k=1}^{m^{d+1}}\Pr_{g_{mk}}(\overline{H}_{mk})\equiv\Pr_e.$$

Step 4: Assume that the parameter k in g_{mk} is uniformly distributed over $1,\dots,m^{d+1}$. Thus $\Pr_e$ is the Bayesian posterior probability of misclassification of the above decision rule based on the closest f_{mk} to $\hat{f}$. We can bound $\Pr_e$ from below using Fano's Lemma (see Khas'minskii (1976), Ibragimov and Has'minskii (1981, p. 323)). Indeed, this lemma gives a lower bound on the probability of misclassification for any decision rule that selects a value among a finite number of equally probable values. In our case, because of i.i.d. sampling, we obtain

$$\Pr_e \geq 1 - \frac{L\mathscr{H}((B,X,I),k) + \log 2}{\log(m^{d+1}-1)},$$

where $\mathscr{H}((B,X,I),k)$ is the amount of Shannon information in $(X,I,B_p,\ p=1,\dots,I)$ relative to the parameter k, i.e.,

$$\begin{aligned}\mathscr{H}((B,X,I),k) &= E(\log(g(k|B,X,I)/g(k))) = E(\log(g(B,X,I|k)/g(B,X,I)))\\ &= E\left(\log\frac{g_{mk}(B,X,I)}{(1/m^{d+1})\Sigma_{j=1}^{m^{d+1}} g_{mj}(B,X,I)}\right)\\ &= -\frac{1}{m^{d+1}}\sum_{k=1}^{m^{d+1}}\sum_{i\in\mathscr{I}}\int\log\left(\frac{1}{m^{d+1}}\sum_{j=1}^{m^{d+1}}\frac{g_{mj}(b,x,i)}{g_{mk}(b,x,i)}\right)g_{mk}(b,x,i)\,db\,dx.\end{aligned}$$

We now bound $\mathscr{H}((B,X,I),k)$ from above. Using the concavity of the logarithm function, a simple second-order expansion, the definition of the g_{mk}'s, (B.1), the fact that g_{mk} is bounded away from 0, the orthogonality of the ϕ_{mk}'s, and a change of variables, we get:

$$\begin{aligned}\mathscr{H}((B,X,I),k) &\leq -\frac{1}{m^{2(d+1)}}\sum_{k,j=1}^{m^{d+1}}\int\log\left(\frac{g_{mj}(b,x,2)}{g_{mk}(b,x,2)}\right)g_{mk}(b,x,2)\,db\,dx\\ &\leq \frac{C_5}{m^{2(d+1)}}\sum_{k,j=1}^{m^{d+1}}\int(\phi_{mj}(b,x)-\phi_{mk}(b,x))^2\,db\,dx\\ &= \frac{C_5}{m^{d+1}}\sum_{k=1}^{m^{d+1}}\int\phi_{mk}^2(b,x)\,db\,dx\\ &= \frac{C_5}{m^{d+1}}\sum_{k=1}^{m^{d+1}}\int\frac{C_3^2}{m^{2R+2}}\phi^2(\lambda_2 m(b-b_k),\lambda_2 m(x-x_k))\,db\,dx\\ &= \frac{C_5}{\lambda_2^{d+1}}\frac{C_3^2}{m^{2R+d+3}}\int\phi^2(b,x)\,db\,dx\\ &= C_6\frac{\kappa^2}{(r_L^*)^{(2R+d+3)/R}} = C_6\frac{\kappa^2\log L}{L}\end{aligned}$$

by (B.3) and the definition of r_L^*. Thus Fano's Lemma gives

$$\Pr_e \geq 1 - \frac{C_6\kappa^2\log L + \log 2}{\log(m^{d+1}-1)}.$$

As $m = (r_L^*)^{1/R} = (L/\log L)^{1/(2R+d+3)}$, taking κ (i.e., C_3) small enough gives $\Pr_e \geq 1/2$ for $L \geq L_0(\epsilon)$. As this can be done for any value of ϵ, this completes the proof of Theorem 2. *Q.E.D.*

B.2. *Uniform Consistency of Boundary Estimators*

PROOF OF PROPOSITION 2: Because the proof for $\hat{\underline{b}}(x)$ is similar, we consider only the upper boundary estimator $\hat{\bar{b}}(x,i)$. The proof is divided into two steps.

Step 1: First, we establish an exponential type inequality. Assume that x lies in $\pi_{k_1,\dots,k_d} \equiv \pi$. Define $\bar{b}^*(x,i) = \sup_{y\in\pi} \bar{b}(y,i)$. Using (16), the triangular inequality, and Proposition 1-(i), we have $|\hat{\bar{b}}(x,i) - \bar{b}(x,i)| \le \bar{b}^*(x,i) - \hat{\bar{b}}(x,i) + |\bar{b}|_1 h_\partial$ with $|\bar{b}|_1 = |\bar{b}|_{1,[\underline{x},\bar{x}]\times\mathcal{I}}$. Thus

$$\text{(B.4)} \qquad \Pr(|\hat{\bar{b}}(x,i) - \bar{b}(x,i)| > t) \le \Pr(\hat{\bar{b}}(x,i) < \bar{b}^*(x,i) - t + |\bar{b}|_1 h_\partial).$$

Let RHS be the right-hand side of (B.4). Using A1 we have

$$RHS = \Pr{}^L\Big(\sup_{p=1,\dots,i} B_p \mathbb{1}(X\in\pi,\ I=i) < \bar{b}^*(x,i) - t + |\bar{b}|_1 h_\partial\Big).$$

Provided t and h_∂ are sufficiently small so that $\bar{b}^*(x,i) - t + |\bar{b}|_1 h_\partial > 0$, then $(RHS)^{1/L}$ equals

$$\text{(B.5)} \qquad [1 - \Pr(I=i \text{ and } X\in\pi)] \\ + \int \Pr{}^i(B < \bar{b}^*(x,i) - t + |\bar{b}|_1 h_\partial | X=y,\ I=i)\mathbb{1}_\pi(y) f(y,i)\,dy.$$

Using Proposition 1-(ii), an upper bound for the second term is

$$\int \Pr{}^i(B < \bar{b}(y,i) - t + 2|\bar{b}|_1 h_\partial | X=y,\ I=i)\mathbb{1}_\pi(y) f(y,i)\,dy$$

$$\le \Big[1 - \inf_{(b,x,i)\in S(G)} g(b|x,i)(t - 2|\bar{b}|_1 h_\partial)\Big]^i \Pr(I=i \text{ and } X\in\pi),$$

provided $t - 2|\bar{b}|_1 h_\partial \ge 0$ (which is satisfied by our choice of t in Step 2), and provided t and h_∂ are sufficiently small to ensure that $\bar{b}(y,i) - t + 2|\bar{b}|_1 h_\partial \ge \underline{b}(y,i)$ and that the term in brackets is strictly between 0 and 1. Note that we have used $\inf_{(x,i)\in[\underline{x},\bar{x}]\times\mathcal{I}}(\bar{b}(x,i) - \underline{b}(x,i)) > 0$, which is ensured by Proposition 1-(i). Moreover we have $h_\partial^d \inf_{x\in\pi} f(x,i) \le \Pr(I=i \text{ and } X\in\pi) \le h_\partial^d \sup_{x\in\pi} f(x,i)$. Hence, because $i>1$, then (B.5) is bounded above by

$$1 - h_\partial^d \inf_{x\in\pi} f(x,i) + \Big[1 - \inf_{(b,x,i)\in S(G)} g(b|x,i)(t - 2|\bar{b}|_1 h_\partial)\Big] h_\partial^d \sup_{x\in\pi} f(x,i)$$

$$= 1 + h_\partial^d\Big(\sup_{x\in\pi} f(x,i) - \inf_{x\in\pi} f(x,i)\Big)$$

$$- h_\partial^d \sup_{(x,i)\in S(X)} f(x,i) \inf_{(b,x,i)\in S(G)} g(b|x,i)(t - 2|\bar{b}|_1 h_\partial)$$

$$\le 1 + c_1 h_\partial^{d+1} - c_2 h_\partial^d (t - 2|\bar{b}|_1 h_\partial)$$

$$\le (1 + (c_1 + 2c_2|\bar{b}|_1) h_\partial^{d+1})\Big(1 - \frac{c_2}{1 + (c_1 + 2c_2|\bar{b}|_1) h_\partial^{d+1}} h_\partial^d t\Big)$$

$$\le (1 + c_3 h_\partial^{d+1})(1 - c_4 h_\partial^d t),$$

for h_∂ and t sufficiently small, where c_i's are strictly positive constants independent of (x,i), and where we have used $\sup_{x\in\pi} f(x,i) - \inf_{x\in\pi} f(x,i) \le \sup_{(x,y)\in\pi^2}|f(x,i) - f(y,i)| \le c_1 h_\partial$, for $(\pi,i) \subset S(f(x,i))$ by A2-(iii). Therefore (B.4) and A4-(iii) give

$$\Pr(|\hat{\bar{b}}(x,i) - \bar{b}(x,i)| > t) \le (1 + c_5 \log L/L)^L (1 - c_6 t(\log L/L)^{d/(d+1)})^L$$

$$\le L^{c_5}(1 - c_6 t(\log L/L)^{d/(d+1)})^L,$$

for t sufficiently small and L sufficiently large.

Step 2: Let $x_\pi \in \pi$. Note that $\hat{\bar{b}}(x,i) = \hat{\bar{b}}(x_\pi, i)$ whenever $x \in \pi$ by (16). Thus, using the triangular inequality and letting $t = \mu/(2r_\partial)$ in the above inequality, we have for each $i \in \mathscr{I}$

$$\begin{aligned}
\Pr\left(r_\partial \sup_{x \in [\underline{x}, \bar{x}]} |\hat{\bar{b}}(x,i) - \bar{b}(x,i)| > \mu \right)
&\le \Pr\left(\sup_{\pi} |\hat{\bar{b}}(x_\pi, i) - \bar{b}(x_\pi, i)| > \mu/(2r_\partial) \right) \\
&\quad + \mathbb{1}\left(\sup_{\pi} \sup_{x \in \pi} |\bar{b}(x,i) - \bar{b}(x_\pi, i)| > \mu/(2r_\partial) \right) \\
&\le \sum_{\pi} \Pr(|\hat{\bar{b}}(x_\pi, i) - \bar{b}(x_\pi, i)| > \mu/(2r_\partial)) \\
&\quad + \sum_{\pi} \mathbb{1}\left(\sup_{x \in \pi} |\bar{b}(x,i) - \bar{b}(x_\pi, i)| > \mu/(2r_\partial) \right) \\
&\le \operatorname{card}(\pi) L^{c_5}(1 - c_6\, \mu(\log L/2L))^L + \operatorname{card}(\pi)\, \mathbb{1}(|\bar{b}|_1 r_\partial h_\partial > \mu/2),
\end{aligned}$$

where $\operatorname{card}(\pi)$ is the number of elements in the partition of $[\underline{x}, \bar{x}]$.

Now, take μ large enough. The second term vanishes, as $r_\partial h_\partial = \lambda_\partial$ from A4-(iii). Moreover, the first term gives a converging series since

$$\log(1 - c_6 t(\log L/2L))^L = L \log(1 - c_6\, \mu(\log L/2L)) \sim -(c_6/2)\mu \log L,$$

while $\operatorname{card}(\pi) = O(h_\partial^{-d}) = O((L/\log L)^{d/(d+1)})$ by A4-(iii). Hence, by the Borel-Cantelli Lemma, we have shown that there exists μ (sufficiently large) such that

$$\Pr\left(\limsup \left\{ r_\partial \sup_{x \in [\underline{x}, \bar{x}]} |\hat{\bar{b}}(x,i) - \bar{b}(x,i)| > \mu \right\} \right) = 0.$$

This establishes the desired result for each i. Since $\mathscr{I}$ is finite, the desired result follows. *Q.E.D.*

B.3. *Uniform Consistency of Pseudo Private Values*

To prove Proposition 3 we need a lemma that gives uniform convergence rates of $\tilde{G}$, $\tilde{g}$, $\hat{f}(x)$, and $\tilde{f}(v,x)$, where the latter is the infeasible estimator defined in (22). We consider uniform convergence on a fixed subset as well as on subsets expanding to the supports. Specifically, let

$$\begin{aligned}
\mathscr{C}_L(B) &\equiv \{(b,x,i) \in S(G);\ \{(b,x) + S(h_G) \cup S(h_g)\} \subset S_i(G)\}, \\
\mathscr{C}_L(V) &\equiv \{(v,x) \in S(f(v,x));\ \{(v,x) + S(h_f)\} \subset S(f(v,x))\}, \\
\mathscr{C}_L(X) &\equiv \{x \in S(f(x));\ \{x + S(h_X)\} \subset S(f(x))\},
\end{aligned}$$

where $S(h_f)$ and $S(h_X)$ are defined similarly to $S(h_G)$ and $S(h_g)$ in the text, i.e., are the supports of the kernels $K_f(\cdot/h_f, \cdot/h_f)$ and $K_X(\cdot/h_X)$, respectively. Let

$$\text{(B.6)} \qquad r_G = r_X = \left(\frac{L}{\log L} \right)^{(R+1)/(2R+d+2)}, \quad r_g = r_f = \left(\frac{L}{\log L} \right)^{R/(2R+d+3)},$$
$$r_g^* = \left(\frac{L}{\log L} \right)^{(R+1)/(2R+d+3)},$$

which give the rates of uniform consistency of our estimators.

LEMMA B2: *Given A1–A4, we have almost surely*

(i) $$|\tilde{G}(b,x,i)-G(b,x,i)|_{0,\mathscr{C}_L(B)}=O(1/r_G),\quad |\tilde{g}(b,x,i)-g(b,x,i)|_{0,\mathscr{C}_L(B)}=O(1/r_g),$$

$$|\tilde{f}(v,x)-f(v,x)|_{0,\mathscr{C}_L(V)}=O(1/r_f),\quad |\hat{f}(x)-f(x)|_{0,\mathscr{C}_L(X)}=O(1/r_X),$$

(ii) $$|\tilde{g}(b,x,i)-g(b,x,i)|_{0,\mathscr{C}(B)}=O(1/r_g^*),$$

where $\mathscr{C}_L(B)$, $\mathscr{C}_L(V)$ and $\mathscr{C}_L(X)$ are expanding subsets defined above, while $\mathscr{C}(B)$ is a (fixed) arbitrary inner subset of $S(G)$.

PROOF OF PROPOSITION 3: Let $G_{pl}=G(B_{pl},X_l,I_l)$, $\tilde{G}_{pl}=\tilde{G}(B_{pl},X_l,I_l)$, $g_{pl}=g(B_{pl},X_l,I_l)$, and $\tilde{g}_{pl}=\tilde{g}(B_{pl},X_l,I_l)$. Define $\tilde{c}_g=\min\{|\tilde{g}_{pl}|;\ l=1,\dots,L,\ p=1,\dots,I_l,\ \hat{V}_{pl}\neq+\infty\}$.

(i) Because $I_l\geq 2$, it follows from (12) and (18) that

$$\mathbb{1}(\hat{V}_{pl}\neq+\infty)|\hat{V}_{pl}-V_{pl}|\leq \mathbb{1}(\hat{V}_{pl}\neq+\infty)|\tilde{\psi}(B_{pl},X_l,I_l)-\psi(B_{pl},X_l,I_l)|.$$

Hence, using the set $\mathscr{C}_L(B)$ introduced in Lemma B2, we obtain

(B.7) $$\begin{aligned}&\mathbb{1}(\hat{V}_{pl}\neq+\infty)\mathbb{1}_{\mathscr{C}_L(B)}(B_{pl},X_l,I_l)|\tilde{\psi}(B_{pl},X_l,I_l)-\psi(B_{pl},X_l,I_l)|\\&\quad=\frac{\mathbb{1}(\hat{V}_{pl}\neq+\infty)\mathbb{1}_{\mathscr{C}_L(B)}(B_{pl},X_l,I_l)}{g_{pl}|\tilde{g}_{pl}|}|(\tilde{G}_{pl}-G_{pl})g_{pl}+(g_{pl}-\tilde{g}_{pl})G_{pl}|\\&\quad\leq\frac{\mathbb{1}(\hat{V}_{pl}\neq+\infty)\mathbb{1}_{\mathscr{C}_L(B)}(B_{pl},X_l,I_l)}{c_fc_g\tilde{c}_g}\Big||g|_0|\tilde{G}_{pl}-G_{pl}|+|f_m|_0|\tilde{g}_{pl}-g_{pl}|\Big|,\end{aligned}$$

because $G(b,x,i)\leq G(\bar{b}(x,i),x,i)=g(x,i)=f_m(x,i)$ and $g_{pl}=g(B_{pl}|X_l,I_l)f_m(X_l,I_l)\geq c_fc_g$ by A2-(ii) and Proposition 1-(ii). Thus

(B.8) $$\begin{aligned}&\mathbb{1}(\hat{V}_{pl}\neq+\infty)|\hat{V}_{pl}-V_{pl}|\\&\quad=\mathbb{1}(\hat{V}_{pl}\neq+\infty)(\mathbb{1}_{\mathscr{C}_L(B)}(B_{pl},X_l,I_l)+1-\mathbb{1}_{\mathscr{C}_L(B)}(B_{pl},X_l,I_l))|\hat{V}_{pl}-V_{pl}|\\&\quad\leq\frac{\mathbb{1}(\hat{V}_{pl}\neq+\infty)\mathbb{1}_{\mathscr{C}_L(B)}(B_{pl},X_l,I_l)}{c_fc_g\tilde{c}_g}\Big||g|_0|\tilde{G}_{pl}-G_{pl}|+|f_m|_0|\tilde{g}_{pl}-g_{pl}|\Big|\\&\qquad+\mathbb{1}(\hat{V}_{pl}\neq+\infty)(1-\mathbb{1}_{\mathscr{C}_L(B)}(B_{pl},X_l,I_l))|\hat{V}_{pl}-V_{pl}|.\end{aligned}$$

Now, from A3-(i) $S(2h_g)$ ($S(2h_G)$) is a hypercube centered at 0 with edge $2\rho_gh_g(2\rho_Gh_G)$. Thus, if $\hat{V}_{pl}\neq+\infty$, Proposition 2 and (19) imply that the distance in supnorm of (B_{pl},X_l) to the frontier of $S_{I_l}(G)$ is at least $\rho_gh_g-t(\log L/L)^{1/(d+1)}$, for some t and L large enough. But, for L large enough, $\rho_gh_g-t(\log L/L)^{1/(d+1)}\geq\rho_gh_g/2$ because $(\log L/L)^{1/(d+1)}=o(h_g)$ from A4-(ii). Hence, if $\hat{V}_{pl}\neq+\infty$, the distance in supnorm of (B_{pl},X_l) to the frontier of $S_{I_l}(G)$ is at least $\rho_gh_g/2$. Therefore $(B_{pl},X_l,I_l)\in\mathscr{C}_L(B)$. Hence we have shown that for any p,l, $\mathbb{1}(\hat{V}_{pl}\neq+\infty)(1-\mathbb{1}_{\mathscr{C}_L(B)}(B_{pl},X_l,I_l))=0$ almost surely as $L\to\infty$. It follows that the second term in (B.8) vanishes. Moreover, from Lemma B2, we have $\tilde{c}_g\to c_g>0$ by Proposition 1-(ii) and the first term is $O(\max\{1/r_G,1/r_g\})=O(1/r_g)$. The desired result follows.

(ii) Define $\mathscr{C}(B)=\bigcup_{i\in\mathscr{I}}\mathscr{C}_i(B)$ where $\mathscr{C}_i(B)=\{(b,x,i)\in S_i(G):(\xi(b,x,i),x,i)\in\mathscr{C}(V)\}$. Because ξ is a strictly increasing continuous function and $\mathscr{C}(V)$ is an inner compact subset of $S(f(v,x))$, then $\mathscr{C}_i(B)$ is a (fixed) inner compact subset of $S_i(G)$. Hence, from (19) and Proposition

2, it is easy to see that $\hat{V}_{pl} \neq +\infty$ if $(B_{pl}, X_l, I_l) \in \mathscr{C}(B)$ for L large enough. Thus

$$\mathbb{1}_{\mathscr{C}(V)}(V_{pl}, X_l)|\hat{V}_{pl} - V_{pl}| = \mathbb{1}_{\mathscr{C}(B)}(B_{pl}, X_l, I_l)|\hat{V}_{pl} - V_{pl}|$$
$$= \mathbb{1}_{\mathscr{C}(B)}(B_{pl}, X_l, I_l)\mathbb{1}(\hat{V}_{pl} \neq +\infty)|\hat{V}_{pl} - V_{pl}|$$

almost surely for L large enough. With $\mathscr{C}(B)$ replacing $\mathscr{C}_L(B)$ in the RHS of (B.7), the desired result follows applying Lemma B2-(ii) instead, and noting that $1/r_G < 1/r_g^*$. *Q.E.D.*

B.4. *Optimality and Uniform Consistency With a Binding Reservation Price*

PROOF OF THEOREM 5: Given A2 and A5, it is easy to see that the truncated conditional distribution $F^*(\cdot|x) = [F(\cdot|x) - F(h(x)|x)]/[1 - F(h(x)|x)]$ satisfies the requirements on $F(\cdot|\cdot, i)$ in A2, where the support is now S^*. Hence the proof of Theorem 2 with $f(\cdot|\cdot,\cdot)$ replaced by the truncated conditional density $f^*(\cdot|\cdot)$ establishes that the best rate for estimating uniformly the latter on inner compact subsets of S^* is bounded above by $r_L^* = (L/\log L)^{R/(2R+d+3)}$. On the other hand, we have $E[I^*|X] = I[1 - \Phi(X)]$. Because $\hat{I} = I$ almost surely and $\Phi(\cdot) = F(h(\cdot)|\cdot)$ is $(R+1)$-continuously differentiable on $[\underline{x}, \bar{x}]$ by A2 and A5, it follows from Stone (1982) that the best rate for estimating uniformly $\Phi(\cdot)$ is $(L/\log L)^{(R+1)/(2R+d+2)}$, which is larger than r_L^*. Hence, because $f(v|x) = [1 - \Phi(x)]f^*(v|x)$, the best rate for estimating uniformly $f(\cdot|\cdot)$ on inner compact subsets of S^* is bounded above by r_L^*.

Turning to uniform consistency of our estimator $\hat{f}(v|x) = [1 - \hat{\Phi}(x)]\hat{f}^*(v|x)$, consider $\hat{f}^*(\cdot|\cdot)$. From Theorem 4-(ii), B^* and I^* are conditionally independent given X. Using a result analogous to Lemma B2 with $B_\dagger$ instead B, it can be shown that Proposition 3 still holds using a similar proof, where the pseudo private values are now given by (27) and the new trimming is needed because of the presence of $\hat{f}(X_l)$ in $\hat{\xi}_\dagger(\cdot,\cdot)$. It follows that Theorem 3 applies with $f(\cdot|\cdot)$ replaced by $f^*(\cdot|\cdot)$. That is, $\hat{f}^*(\cdot|\cdot)$ converges uniformly to $f^*(\cdot|\cdot)$ on inner compact subsets of S^* at the rate r_L^*. On the other hand, because $\hat{I} = I$ almost surely, we know from A4-(ii) and Hardle (1991) that $\hat{\Phi}(\cdot)$ converges uniformly on inner compact subsets of $[\underline{x}, \bar{x}]$ at the rate $(L/\log L)^{(R+1)/(2R+d+2)}$. The desired result follows. *Q.E.D.*

APPENDIX C

PROOFS OF LEMMAS

This Appendix gives the proofs of all lemmas in Appendices A and B.

C.1. *Proofs of Lemmas in Appendix A*

To prove Lemmas A1 and A2 we need a version of the Implicit Function Theorem.

LEMMA C1: *For every $i \in \mathscr{I}$, let $\alpha(\cdot,\cdot,i)$ be a function on $S_i(\alpha)$. Let $\beta(\cdot,\cdot,i)$ be a function on $S_i(\beta)$ such that $\beta(\cdot,\cdot,i) \in S_i(\alpha)$. For some nonnegative real number γ, assume that*

(C.1) $\alpha(\beta(b, x, i), x, i) = \gamma b \quad \forall (b, x) \in S_i(\beta)$ *and* $\forall i \in \mathscr{I}$.

Moreover, for each $i \in \mathscr{I}$, assume that the following conditions hold:

(i) *The mapping $\alpha(\cdot,\cdot,i)$ admits up to $(R+1)$ continuous bounded partial derivatives on $S_i(\alpha)$, with $\alpha'(v, x, i) \geq c_\alpha > 0$ for all $(v, x) \in S_i(\alpha)$.*

(ii) $S_i(\beta) = \{(b, x): x \in [\underline{x}, \bar{x}],\ b \in [\underline{b}(x, i), \bar{b}(x, i)]\}$, *where the boundaries $\underline{b}(\cdot, i)$ and $\bar{b}(\cdot, i)$ are continuous on $[\underline{x}, \bar{x}]$ and* $\inf_{x \in [\underline{x}, \bar{x}]}(\bar{b}(x, i) - \underline{b}(x, i)) > 0$.

Then $\beta(\cdot,\cdot,i)$ admits up to $R+1$ continuous bounded partial derivatives on $S_i(\beta)$.

PROOF OF LEMMA C1: For (b, x) in $S_i^o(\beta)$, (C.1) and the Implicit Function Theorem yield:

$$\beta'(b, x, i) = \frac{\gamma}{\alpha'(\beta(b, x, i), x, i)}, \tag{C.2}$$

$$\frac{\partial \beta(b, x, i)}{\partial x_k} = -\frac{1}{\alpha'(\beta(b, x, i), x, i)} \frac{\partial \alpha(\beta(b, x, i), x, i)}{\partial x_k}. \tag{C.3}$$

Hence, in view of (i), $\beta(\cdot, \cdot, i)$ is a Lipschitz function on $S_i^o(\beta)$. Thus, by continuity $\beta(\cdot, \cdot, i)$ can be extended to $\overline{S_i^o(\beta)}$, which is equal to $S_i(\beta)$ by (ii). Let $\tilde{\beta}(\cdot, \cdot, i)$ denote this extension, and note that $\tilde{\beta}(\cdot, \cdot, i)$ solves (C.1). But, because $\beta(b, x, i)$ is the unique solution of (C.1), then $\tilde{\beta}(\cdot, \cdot, i) = \beta(\cdot, \cdot, i)$. Therefore $\beta(\cdot, \cdot, i)$ is continuous and hence bounded on $S_i(\beta)$.

Moreover, since α' is bounded away from 0, (C.2)–(C.3) show that the first partial derivatives of $\beta(\cdot, \cdot, i)$ are continuous and bounded on $S_i(\beta)$. Proceeding by induction yields that $\beta(\cdot, \cdot, i)$ admits up to $R + 1$ continuous partial derivatives over $S_i(\beta)$. *Q.E.D.*

PROOF OF LEMMA A1: (i) Because $F(\cdot|x, i)$ is strictly increasing by A2-(ii), $\bar{v}(x)$ and $\underline{v}(x)$ are the unique solutions in $S_i(F)$ of $F(\underline{v}(x)|x, i) = 0$ and $F(\bar{v}(x)|x, i) - 1 = 0$, respectively. Hence, given A2, Lemma C1 applies with $\alpha(\cdot, \cdot, \cdot) = F(\cdot|\cdot, \cdot)$ or $F(\cdot|\cdot, \cdot) - 1$, $\beta(b, x, i) = \underline{v}(x)$ or $\bar{v}(x)$, $\gamma = 0$, and $S_i(\beta) = [-1, 1] \times [\underline{x}, \bar{x}]$ (say). Moreover, $1 = F(\bar{v}(x)|x, i) - F(\underline{v}(x)|x, i) \le (\bar{v}(x) - \underline{v}(x)) \sup_{(v, x) \in S_i(F)} |f(v|x, i)|$. Because $f(\cdot|\cdot, i)$ is bounded above, then $\inf_{x \in [\underline{x}, \bar{x}]} (\bar{v}(x) - \underline{v}(x)) > 0$.

(ii) For the sake of simplicity, we assume that $x \in \mathbb{R}$. From (8) and the Lebesgue Dominated Convergence Theorem it is immediate that $s(\cdot, \cdot, i)$ is $(R + 1)$-times continuously differentiable on $\{(v, x) : v \in (\underline{v}(x), \bar{v}(x)],\ x \in [\underline{x}, \bar{x}]\}$. Thus it remains to show that the derivatives of $s(\cdot, \cdot, i)$ up to order $R + 1$ are bounded near the lower boundary $\underline{v}(\cdot)$.

For each $x \in [\underline{x}, \bar{x}]$, an $(R + 1)$th order Taylor series expansion of $F(v|x)$ at $(\underline{v}(x), x)$ gives

$$F(v|x) = \frac{f(\underline{v}(x)|x)}{1!} t + \cdots + \frac{f^{(R)}(\underline{v}(x)|x)}{(R + 1)!} t^{R+1} + o(t^{R+1}),$$

uniformly in x, where $t = v - \underline{v}(x)$. Hence

$$F^{i-1}(v|x) = t^{i-1}(a_0(x) + \cdots + a_R(x) t^R + o(t^R)),$$

uniformly in x, for each $k = 0, \ldots, R$, $a_k(x)$ is a homogeneous polynomial of degree $(i - 1)$ in $f(\underline{v}(x)|x)$ and its derivatives $f^{(1)}(\underline{v}(x)|x), \ldots, f^{(k)}(\underline{v}(x)|x)$ with respect to v. In particular, $a_0(x) = f(\underline{v}(x)|x)^{i-1}$, which is nonzero by A2-(ii). It also follows that

$$\int_{\underline{v}(x)}^{v} F^{i-1}(u|x)\, du = t^i \left(\frac{a_0(x)}{i} + \cdots + \frac{a_R(x)}{R + i} t^R + o(t^R) \right),$$

uniformly in x. Therefore

$$\frac{1}{F^{i-1}(v|x)} \int_{\underline{v}(x)}^{v} F^{i-1}(u|x)\, du = t(b_0(x) + \cdots + b_R(x) t^R + o(t^R)),$$

where $b_k(x)$, $k = 0, \ldots, R$, is a polynomial in $a_1(x)/a_0(x), \ldots, a_k(x)/a_0(x)$ and hence in $f^{(0)}(\underline{v}(x)|x)/f(\underline{v}(x)|x), \ldots, f^{(k)}(\underline{v}(x)|x)/f(\underline{v}(x)|x)$. In particular, it is easy to see that $b_0(x) = 1/i$. Thus we obtain an $(R + 1)$th order Taylor series expansion of $s(v, x, i)$ around $(\underline{v}(x), x)$ as

$$s(v, x, i) = \underline{v}(x) + (1 - b_0(x))t - b_1(x)t^2 - \cdots - b_R(x) t^{R+1} - o(t^{R+1}),$$

uniformly in x. Hence $s(v, x, i)$ is continuous in v at $\underline{v}(x)$ with $R + 1$ bounded derivatives with respect to v in the neighborhood of $\underline{v}(x)$. In particular, $\partial s(\underline{v}(x), x, i)/\partial v = (i - 1)/i \neq 0$.

For each $r = 1, \ldots, R + 1$, it remains to prove that the derivatives $\partial^r s(\cdot, x, i)/\partial x^p\, \partial v^{r-p}$, $p = 1, \ldots, r$ are bounded in the neighborhood of $\underline{v}(x)$ for $x \in [\underline{x}, \bar{x}]$. The proof is by induction on r. For

$r=1$, differentiating the boundary condition $s(\underline{v}(x),x,i)=\underline{v}(x)$ gives

$$\frac{\partial s(\underline{v}(x),x,i)}{\partial v}\frac{d\underline{v}(x)}{dx}+\frac{\partial s(\underline{v}(x),x,i)}{\partial x}=\frac{d\underline{v}(x)}{dx}.$$

But $\partial s(\underline{v}(x),x,i)/\partial v$ is bounded, as shown above, while $d\underline{v}(x)/dx$ is bounded from Part (i). Hence $\partial s(\underline{v}(x),x,i)/\partial x$ is bounded.

Suppose next that all the partial derivatives up to order r of $s(\cdot,\cdot,i)$ are bounded, $1\le r\le R$. From the above Taylor series expansion of $s(\cdot,x,i)$ we obtain the system

$$s(\underline{v}(x),x,i)=\underline{v}(x),\quad \frac{\partial s(\underline{v}(x),x,i)}{\partial v}=1-b_0(x),\ldots,\frac{\partial^r s(\underline{v}(x),x,i)}{\partial v^r}=-b_{r-1}(x).$$

Now, differentiate the first equation $r+1$ times, the second equation r times,...,the last equation one time. This is possible because $b_k(x)$ is $R-k$ times differentiable (with bounded derivatives). Because $\partial^{r+1}s(\underline{v}(x),x,i)/\partial v^{r+1}$ and $d\underline{v}(x)/dx$ are bounded, differentiation of the last equation shows that $\partial^{r+1}s(\underline{v}(x),x,i)/\partial v^r\,\partial x$ is bounded. Combining this with the twice-differentiation of the next to the last equation shows that $\partial^{r+1}s(\underline{v}(x),x,i)/\partial v^{r-1}\,\partial x^2$ is bounded. Continuing up to the first equation shows that $\partial^{r+1}s(\underline{v}(x),x,i)/\partial x^{r+1}$ is bounded. This completes the proof that $s(\cdot,\cdot,i)$ admits $R+1$ bounded continuous derivatives on its support.

Lastly, $s'(\cdot,x,i)$ can only vanish at $\underline{v}(x)$ because

$$s'(v,x,i)=\frac{(i-1)f(v|x)}{F^i(v|x)}\int_{\underline{v}(x)}F^{i-1}(u|x)\,du.$$

But, as noted above, $s'(\underline{v}(x),x,i)=(i-1)/i$. Thus $s'(\cdot,\cdot,i)$ is bounded away from zero. *Q.E.D.*

PROOF OF LEMMA A2: Because $\underline{b}(x,i)=s(\underline{v}(x),x,i)$ and $\bar{b}(x,i)=s(\bar{v}(x,i),x,i)$, then (i) follows immediately from Lemma A1. To prove (ii), we note that the function ξ solves

$$s(\xi(b,x,i),x,i)=b\quad \forall(b,x)\in S_i(G).$$

The desired result follows from Part (i), Lemma A1-(ii), and Lemma C1. Moreover, $\xi'(b,x,i)=1/(s'(\xi(b,x,i),b,x,i))$ is bounded away from 0 by Lemma A1-(ii). *Q.E.D.*

PROOF OF LEMMA A3: (i) directly follows from Lemma A1-(i), A5-(ii), and A5-(iii). Regarding (ii), because $F(v|x)\ge F(p_0|x)$, which is bounded away from zero by A5-(iii), it is easy to see that $s_\dagger^2(\cdot,\cdot,\cdot)$ and hence $s_\dagger(\cdot,\cdot,\cdot)$ have $R+1$ continuous partial derivatives on $S(F^*)$, except possibly near $v=p_0$ as $s_\dagger(p_0,\cdot,\cdot)=0$. However, letting $v_\dagger=v-p_0$, we have

$$F(v_\dagger+p_0|x)=F(p_0|x)\left(\frac{f(p_0|x)}{F(p_0|x)}\frac{v_\dagger}{1!}+\cdots+\frac{f^{(R)}(p_0|x)}{F(p_0|x)}\frac{v_\dagger^{(R+1)}}{(R+1)!}+o(v_\dagger^{(R+1)})\right)$$

uniformly in (x,p_0). Similarly to the proof of Lemma A1-(ii), it can be shown that

$$s_\dagger^2(v,x)=a_2(x,p_0)v_\dagger^2+\cdots+a_{R+2}(x,p_0)v_\dagger^{R+2}+o(v_\dagger^{R+2})$$

uniformly in (x,p_0), where $a_k(x,p_0)$ is a polynomial in $f(p_0|x)/F(p_0|x),\ldots,f^{(k-2)}(p_0|x)/F(p_0|x)$ for $k=2,\ldots,R+2$. In particular, $a_2(x,p_0)=(I-1)f(p_0|x)/(2F(p_0|x))$. Hence

$$s_\dagger(v,x)=b_1(x,p_0)v_\dagger+\cdots+b_{R+1}(x,p_0)v_\dagger^{R+1}+o(v_\dagger^{R+1})$$

uniformly in (x,p_0), where $b_1(x,p_0)b_k(x,p_0)$ is a polynomial in $f(p_0|x)/F(p_0|x),\ldots,f^{(k-2)}(p_0|x)/F(p_0|x)$ for $k=1,\ldots,R+1$. In particular, $b_1(x,p_0)=[(I-1)f(p_0|x)/(2F(p_0|x))]^{1/2}$. Because $f(p_0|x)$ is bounded away from zero and infinity on $S(F^*)$ by A2, and because $1\ge F(p_0|x)$, which is bounded away from zero by A5-(iii), then $s_\dagger(\cdot,x)$ admits $R+1$ bounded derivatives at $v=p_0$ and hence on $S(F^*)$ with $s_\dagger'(\cdot,\cdot)$ bounded away from zero. It remains to show that the cross partial derivatives up to order $R+1$ are also bounded in the neighborhood of $v=p_0$. This is proved as in the end of the proof of Lemma A1-(ii). *Q.E.D.*

PROOF OF LEMMA A4: The proof is similar to that of Lemma A.2 *Q.E.D.*

C.2. Proofs of Lemmae in Appendix B

PROOF OF LEMMA B1: (i) The (b_k, x_k)'s are bounded away from the boundaries of $S_2(G_0)$. Thus for m large enough, $g_{mk}(\cdot, x, 2) = g_0(\cdot, x, 2)$, $G_{mk}(\cdot, x, 2) = G_0(\cdot, x, 2)$, and $\xi_{mk}(\cdot, x, 2) = \xi_0(\cdot, x, 2)$ in the neighborhood of $\underline{b}_0(x,2)$ or $\overline{b}_0(x,2)$. Because $g_0(\cdot,\cdot,2)$ is bounded away from zero on $S_2(G_0)$, then the support of G_{mk} is $S(G_0)$ for m large enough. Moreover, (B.2) implies that the support of $f_{mk}(\cdot,\cdot,2)$ is $S_2(F_0)$. The desired result follows as $f_{mk}(\cdot,\cdot,i) = f_0(\cdot,\cdot,i)$ for $i \neq 2$.

To prove (ii), note that $f_{mk}(\cdot,\cdot) - f_{mj}(\cdot,\cdot) = \sum_{i \in \mathscr{I}}[f_{mk}(\cdot,\cdot,i) - f_{mj}(\cdot,\cdot,i)] = f_{mk}(\cdot,\cdot,2) - f_{mj}(\cdot,\cdot,2)$. Note also that $(b_k, x_k) \in \mathscr{C}_2(B)$ implies $(v_k, x_k) \in \mathscr{C}(V)$, where $v_k \equiv \xi_0(b_k, x_k, 2)$. Thus it suffices to prove that $|f_{mk}(v_k, x_k, 2) - f_{mj}(v_k, x_k, 2)| = C_4 C_3 \lambda_2 / m^R$. Consider $f_{mk}(v_k, x_k, 2)$. Note that (B.1) implies $G_{mk}(b_k, x_k, 2) = G_0(b_k, x_k, 2)$ and $g_{mk}(b_k, x_k, 2) = g_0(b_k, x_k, 2)$. Hence $v_k \equiv \xi_0(b_k, x_k, 2) = \xi_{mk}(b_k, x_k, 2)$, which implies $\xi_{mk}^{-1}(v_k, x_k, 2) = b_k$. Therefore, computing $g'_{mk}(b_k, x_k, 2)$ and using (B.2) give

$$f_{mk}(v_k, x_k, 2) = \frac{g_0^3(b_k, x_k, 2)}{2g_0^2(b_k, x_k, 2) - G_0(b_k, x_k, 2)(g'_0(b_k, x_k, 2) + \lambda_2 C_3 \phi'(0,0)/m^R)}. \tag{C.4}$$

Consider next $f_{mj}(v_k, x_k, 2)$. From (B.1), we have $g_{mj}(\cdot,\cdot,2) = g_0(\cdot,\cdot,2)$, $G_{mj}(\cdot,\cdot,2) = G_0(\cdot,\cdot,2)$, and hence $\xi_{mj}(\cdot,\cdot,2) = \xi_0(\cdot,\cdot,2)$ except on $S(\phi_{mj})$, which is a hypercube centered at (b_j, x_j). Moreover, for m sufficiently large $S(\phi_{mk})$ and $S(\phi_{mj})$ are disjoint so that $(b_k, x_k) \notin S(\phi_{mj})$. Hence $\xi_{mj}(b_k, x_k, 2) = \xi_0(b_k, x_k, 2) = v_k$ so that $\xi_{mj}^{-1}(v_k, x_k, 2) = b_k$. Thus (B.2) gives

$$f_{mj}(v_k, x_k, 2) = \frac{g_0^3(b_k, x_k, 2)}{2g_0^2(b_k, x_k, 2) - G_0(b_k, x_k, 2) g'_0(b_k, x_k, 2)}. \tag{C.5}$$

Now compare (C.4) and (C.5). As $\phi'(0,0) \neq 0$ and $G_0(b_k, x_k, 2) > c \geq 0$ (because the (b_k, x_k)'s are far enough from the boundaries), the desired result follows.

The proof of (iii) is more involved and is divided in three steps of which the first two establish properties similar to (iii) for ξ_{mk} and s_{mk}, respectively.

Step 1: As $g_{mk}(\cdot,\cdot,i) = g_0(\cdot,\cdot,i)$ except possibly when $i = 2$ and $(b, x) \in \mathscr{C}_2(B)$, then $|g_{mk}(\cdot,\cdot,\cdot) - g_0(\cdot,\cdot,\cdot)|_{r, S_2(G_0)} = |g_{mk}(\cdot,\cdot,2) - g_0(\cdot,\cdot,2)|_{r, \mathscr{C}_2(B)}$ and $|G_{mk}(\cdot,\cdot,\cdot) - G_0(\cdot,\cdot,\cdot)|_{r, S_2(G_0)} = |G_{mk}(\cdot,\cdot,2) - G_0(\cdot,\cdot,2)|_{r, \mathscr{C}_2(B)}$ for all $r \geq 0$. Now, using a change of variable, we have

$$\begin{aligned} G_{mk}(b, x, 2) &= G_0(b, x, 2) + \frac{C_3}{\lambda_2 m^{R+2}} \int_{m\lambda_2(\underline{b}(x) - b_k)}^{m\lambda_2(b - b_k)} \phi(u, m\lambda_2(x - x_k))\, du \\ &= G_0(b, x, 2) + \frac{C_3}{\lambda_2 m^{R+2}} \Phi(m\lambda_2(b - b_k), m\lambda_2(x - x_k)), \end{aligned}$$

where $\Phi(b, x) \equiv \int_{-1}^{b} \phi(u, x)\, du$. Since g_0 admits up to $R + 1$ bounded continuous derivatives on $\mathscr{C}_2(B)$ by Proposition 1-(iv), we have for any $n_0 + n_1 + \cdots + n_d = r$, $0 \leq r \leq R + 1$,

$$\begin{aligned} &\frac{\partial^r G_{mk}(b, x, 2)}{\partial^{n_0} b\, \partial^{n_1} x_1 \cdots \partial^{n_d} x_d} - \frac{\partial^r G_0(b, x, 2)}{\partial^{n_0} b\, \partial^{n_1} x_1 \cdots \partial^{n_d} x_d} \\ &\quad = \frac{C_3 (\lambda_2 m)^r}{\lambda_2 m^{R+2}} \frac{\partial^r \Phi(\lambda_2 m(b - b_k), \lambda_2 m(x - x_d))}{\partial^{n_0} b\, \partial^{n_1} x_1 \cdots \partial^{n_d} x_d}, \end{aligned} \tag{C.6}$$

$$\begin{aligned} &\frac{\partial^r g_{mk}(b, x, 2)}{\partial^{n_0} b\, \partial^{n_1} x_1 \cdots \partial^{n_2} x_d} - \frac{\partial^r g_0(b, x, 2)}{\partial^{n_0} b\, \partial^{n_1} x_1 \cdots \partial^{n_2} x_d} \\ &\quad = \frac{C_3 (\lambda_2 m)^r}{m^{R+1}} \frac{\partial^r \phi(\lambda_2 m(b - b_k), \lambda_2 m(x - x_d))}{\partial^{n_0} b\, \partial^{n_1} x_1 \cdots \partial^{n_d} x_d}, \end{aligned}$$

for all $(b,x) \in \mathscr{C}_2(B)$. Thus (C.6) immediately gives $|G_{mk} - G_0|_{r,S(G_0)} = C_3\lambda_2^{r-1}O(1/m^{R+2-r})$ and $|g_{mk} - g_0|_{r,S(G_0)} = C_3\lambda_2^r O(1/m^{R+1-r})$. Since $\xi_{mk}(b,x,i) = 1 + G_{mk}(b,x,i)/(i-1)(g_{mk}(b,x,i))$, the rth partial derivatives of $\xi_{mk}(\cdot,\cdot,i)$ are some polynomial functions in the partial derivatives of order less than or equal to r of G_{mk} and g_{mk}, divided by g_{mk}^{r+1}. Since g_{mk} is bounded away from 0, we have uniformly in k

$$\text{(C.7)} \qquad |\xi_{mk} - \xi_0|_{r,S(G_0)} = C_3\lambda_2^r O(1/m^{R+1-r}) \qquad (r = 0,\ldots,R+1).$$

Step 2: We have $\xi_{mk}(s_{mk}(v,x,i),x,i) = v$ for $(v,x,i) \in S(F_0)$. Lemma C1 and (C.7) show that s_{mk} admits up to $R+1$ bounded continuous derivatives on $S(F_0)$. First we consider $|s_{mk} - s_0|_{0,S(F_0)}$ and show that

$$\text{(C.8)} \qquad |s_{mk} - s_0|_{0,S(F_0)} = C_3\lambda_2 O(1/m^{R+1}).$$

Since $\xi_{mk} = \xi_0 + C_3\lambda_2 O(1/m^{R+1})$ uniformly by (C.7) and $\xi_0' \geq c_\xi > 0$ by Lemma A2, we get

$$\begin{aligned}
&|s_{mk}(v,x,i) - s_0(v,x,i)| \\
&\quad \leq \frac{1}{c_\xi}|\xi_0(s_{mk}(v,x,i),x,i) - \xi_0(x_0(v,x,i),x,i)| \\
&\quad \leq \frac{1}{c_\xi}|\xi_{mk}(s_{mk}(v,x,i),x,i) - \xi_0(s_0(v,x,i),x,i)| + C_3\lambda_2 O(1/m^{R+1}) \\
&\quad = C_3\lambda_2 O(1/m^{R+1}).
\end{aligned}$$

Next, letting $s_{mk} = s_{mk}(v,x,i)$, (C.2)–(C.3) give

$$s_{mk}' = \frac{1}{\xi_{mk}'(s_{mk},x,i)} \quad \text{and} \quad \frac{\partial s_{mk}}{\partial x_p} = -\frac{\partial \xi_{mk}(s_{mk},x,i)/\partial x_p}{\xi_{mk}'(s_{mk},x,i)}.$$

Differentiating one more time gives

$$\begin{aligned}
s_{mk}'' &= -\frac{\xi_{mk}''(s_{mk},x,i)s_{mk}'}{(\xi_{mk}'(s_{mk},x,i))^2}, \\
\frac{\partial s_{mk}'}{\partial x_q} &= -\frac{1}{(\xi_{mk}'(s_{mk},x,i))^2}\left(\xi_{mk}''(s_{mk},x,i)\frac{\partial s_{mk}}{\partial x_q} + \frac{\partial \xi_{mk}'(s_{mk},x,i)}{\partial x_q}\right), \\
\frac{\partial^2 s_{mk}}{\partial x_p\,\partial x_q} &= -\frac{1}{(\xi_{mk}'(s_{mk},x,i))^2} \\
&\quad \times \left(\xi_{mk}'(s_{mk},x,i)\left\{\frac{\partial \xi_{mk}'(s_{mk},x,i)}{\partial x_p}\frac{\partial s_{mk}}{\partial x_q} + \frac{\partial^2 \xi_{mk}(s_{mk},x,i)}{\partial x_p\,\partial x_q}\right\}\right. \\
&\quad \left. - \frac{\partial \xi_{mk}(s_{mk},x,i)}{\partial x_p}\left\{\xi_{mk}''(s_{mk},x,i)\frac{\partial s_{mk}}{\partial x_q} + \frac{\partial \xi_{mk}'(s_{mk},x,i)}{\partial x_q}\right\}\right).
\end{aligned}$$

An induction argument shows that any rth partial derivative of s_{mk} times $(\xi_{mk}'(s_{mk},x,i))^r$ is a polynomial function of the $(1,\ldots,r)$th's partial derivatives of ξ_{mk} taken at (s_{mk},x,i), and of the $(1,\ldots,r-1)$th's partial derivatives of s_{mk}. Thus, a bound for $|\xi_{mk}^{(r)}(s_{mk},x,i) - \xi_0^{(r)}(s_0,x,i)|_{0,S(F_0)}$ (where $\xi_{mk}^{(r)}$ is any rth partial derivatives of ξ_{mk}) shall give a similar bound for $|s_{mk} - s_0|_{r,S(F_0)}$,

$r=1,\ldots,R+1$. Thus, using equations (C.7) and (C.8) gives

$$
\begin{aligned}
&|\xi_{mk}^{(r)}(s_{mk},x,i)-\xi_0^{(r)}(s_0,x,i)|_{0,S(F_0)}\\
&\quad\le|\xi_{mk}^{(r)}(s_{mk},x,i)-\xi_0^{(r)}(s_{mk},x,i)|_{0,S(F_0)}+|\xi_0^{(r)}(s_{mk},x,i)-\xi_0^{(r)}(s_0,x,i)|_{0,S(F_0)}\\
&\quad\le C_3\lambda_2^r O(1/m^{R+1-r})\\
&\qquad+\begin{cases}|\xi_0|_{r+1,S(G_0)}C_3\lambda_2 O(1/m^{R+1}) & (r=0,\ldots,R),\\ \sup_{(v,x,i)\in S(F_0)}|\xi_0^{(r)}(s_{mk},x,i)-\xi_0^{(r)}(s_0,x,i)| & (r=R+1).\end{cases}
\end{aligned}
$$

Because the sup is $o(1)$, we obtain

$$
\text{(C.9)}\qquad\begin{cases}|s_{mk}-s_0|_{r,S(F_0)}=C_3\lambda_2^r O(1/m^{R+1-r}) & (r=0,\ldots,R),\\ |s_{mk}-s_0|_{R+1,S(F_0)}=C_3\lambda_2^{R+1}O(1)+o(1). & \end{cases}
$$

Step 3: By construction, we have $f_{mk}(v,x,i)=s'_{mk}(v,x,i)g_{mk}(s_{mk}(v,x,i),x,i)$. Thus the rth partial derivative of f_{mk} is a polynomial function of the $(0,\ldots,r)$th derivatives of s'_{mk}, g_{mk}, and s_{mk}. By the same argument as in Step 2, (iii) follows from (C.6) and (C.9). *Q.E.D.*

Further Lemmas: To prove Lemma B2 we need three lemmas. Lemma C2 studies the uniform bias of $\tilde{G}$, $\tilde{g}$, $\hat{f}(x)$, and $\tilde{f}(v,x)$, Lemma C3 studies their variances, and Lemma C4 establishes exponential-type inequalities. Hereafter, the condition "for L sufficiently large" means:

(i) for any (b,x), $K_G(x,(i-j)/h_{gI})=0$ and $K_g(b,x,(i-j)/h_{gI})=0$ if $j\neq i$, and

(ii) for inner closed possibly expanding subsets $\mathscr{C}(B)=\bigcup_{i\in\mathscr{I}}\mathscr{C}_i(B)$, $\mathscr{C}(V)$ and $\mathscr{C}(X)$ of $S(G)$, $S(f(v,x))$ and $S(f(x))$, respectively, we have

$$
\bigcup_{(v,x)\in\mathscr{C}(V)}\{(v,x)+S(h_f)\}\subset S(f(v,x)),\qquad \bigcup_{x\in\mathscr{C}(X)}\{x+S(h_X)\}\subset S(f(x)),
$$
$$
\bigcup_{(b,x)\in\mathscr{C}_i(B)}\{(b,x)+S(h_G)\cup S(h_G)\}\subset\mathscr{C}_i'(B),
$$

where $\mathscr{C}'(B)=\bigcup_{i\in\mathscr{I}}\mathscr{C}_i'(B)$ is an inner closed subset of $S(G)$.

Let $\|\cdot\|_a$ be the absolute norm for vectors, and

$$
M_g^r=\frac{1}{r!}\int\|(u,x)\|_a^r|K_g(u,x,0)|\,du\,dx,\quad M_G^r=\frac{1}{r!}\int\|x\|_a^r|K_G(x,0)|\,dx,
$$
$$
M_f^r=\frac{1}{r!}\int\|(u,x)\|_a^r|K_f(u,x)|\,du\,dx,\qquad M_X^r=\frac{1}{r!}\int\|x\|_a^r|K_X(x)|\,du\,dx,
$$

where $x=(x_1,\ldots,x_d)$. The next lemma on uniform bias considers two cases whether the subset $\mathscr{C}(B)$ expands to $S(G)$ (part (i)) or is fixed (part (ii)). In particular, given A4, $|E\tilde{g}-g|_{0,\mathscr{C}(B)}$ is of order $(\log L/L)^{R/(2R+d+3)}$ in (i), and has the optimal magnitude $(\log L/L)^{(R+1)/(2R+d+3)}$ in (ii). The reason is that $\mathscr{C}(B)$ can expand to $S(G)$ in (i), while $\mathscr{C}(B)$ is fixed in (ii) so that g admits up to $R+1$ continuous bounded derivatives on $\mathscr{C}(B)$ from Proposition 1.

LEMMA C2: *Given A1–A4, we have for L sufficiently large*

$$
\begin{aligned}
\text{(i)}\qquad &|E\tilde{G}-G|_{0,\mathscr{C}(B)}\le\lambda_G^{R+1}M_G^{R+1}|G|_{R+1,S(G)}/r_G,\\
&|E\tilde{g}-g|_{0,\mathscr{C}(B)}\le\lambda_g^R M_g^R|g|_{R,S(g)}/r_g,\\
&|E\tilde{f}(v,x)-f(v,x)|_{0,\mathscr{C}(V)}\le\lambda_f^R M_f^r|f(v,x)|_{R,S(f)}/r_f,\\
&|E\hat{f}(x)-f(x)|_{0,\mathscr{C}(X)}\le\lambda_X^{R+1}M_X^{R+1}|f(x)|_{R+1,S(f)}/r_X,\\
\text{(ii)}\qquad &|E\tilde{g}-g|_{0,\mathscr{C}(B)}\le\lambda_g^{R+1}M_g^{R+1}|g|_{R+1,\mathscr{C}'(B)}/r_g^*.
\end{aligned}
$$

PROOF OF LEMMA C2: Because the proofs are similar, we prove (ii) only. Using A1, we have

$$E\tilde{g}(b,x,i)=E\left[\frac{1}{h_g^{d+1}}K_g\left(\frac{b-B_p}{h_g},\frac{x-X}{h_g},0\right)\mathbb{1}(I=i)\right]$$

$$=\iint K_g(u,y,0)g(b-h_g u,x-h_g y,i)\,du\,dy.$$

Define $\gamma(t)=g(b-th_g u,x-th_g y,i)-g(b,x,i)$ for $t\in[0,1]$. Since the supports of kernels are intervals, $(b-th_g u,x-th_g y)\in(b,x)+S(h_g)\subset\mathscr{C}_i'(B)$ for $(b,x,i)\in\mathscr{C}(B)$ and $t\in[0,1]$. Moreover, using Proposition 1, $\gamma(t)$ admits up to $R+1$ continuous bounded derivatives with

$$\sup_{t\in[0,1]}|\gamma^{(R+1)}(t)|\leq h_g^{R+1}\|(u,y)\|_a^{(R+1)}|g|_{R+1,\mathscr{C}'(B)}.$$

Thus a Taylor expansion with integral remainder gives

$$\gamma(1)-\gamma(0)=\gamma^{(1)}(0)+\cdots+\frac{1}{R!}\gamma^{(R)}(0)+\int_0^1\frac{(1-t)^R}{R!}\gamma^{(R+1)}(t)\,dt,$$

where $\gamma^{(r)}(0)$ is a polynomial of order r in (u,y). Using A3 it follows that

$$|E\tilde{g}(b,x,i)-g(b,x,i)|_{0,\mathscr{C}(B)}$$

$$=\left|\int K_g(u,y,0)(\gamma(1)-\gamma(0))\,du\,dy\right|$$

$$\leq h_g^{R+1}|g|_{R+1,\mathscr{C}'(B)}\left(\int_0^1\frac{(1-t)^R}{R!}dt\right)\times\left(\int\|(u,y)\|_a^{R+1}K_g(u,y,0)\,du\,dy\right)$$

$$=h_g^{R+1}M_g^{R+1}|g|_{R+1,\mathscr{C}'(B)}.$$

The desired bound follows from A4-(ii) and (B.6). *Q.E.D.*

The next lemma obtains uniform variance bounds. Let

$$Q_G=\int K_G^2(x,0)\,dx,\qquad Q_g=\int K_g^2(u,x,0)\,du\,dx,$$

$$Q_f=\int K_f^2(u,x)\,du\,dx,\quad Q_X=\int K_X^2(x)\,dx.$$

LEMMA C3: *Given A1–A4, we have for L sufficiently large*

$$|\mathrm{Var}(\tilde{G})|_{0,\mathscr{C}(B)}\leq\frac{Q_G|G|_{0,S(G)}}{\lambda_G^d r_G^2\log L},\qquad |\mathrm{Var}(\tilde{g})|_{0,\mathscr{C}(B)}\leq\frac{Q_g|g|_{0,S(g)}}{\lambda_g^{d+1}(r_g^*)^2\log L},$$

$$|\mathrm{Var}(\tilde{f}(v,x))|_{0,\mathscr{C}(V)}\leq\frac{h_f^2Q_f|f(v,x)|_{0,S(f)}}{\lambda_f^{d+3}r_f^2\log L},\qquad |\mathrm{Var}(\tilde{f}(x))|_{0,\mathscr{C}(X)}\leq\frac{Q_f|f(x)|_{0,S(f)}}{\lambda_X^d r_X^2\log L}.$$

PROOF OF LEMMA C3: We consider $\tilde{G}$ only, since the other bounds are proved similarly. For $(b,x,i)\in\mathscr{C}(B)$, we have by A1 and the convexity of the square function,

$$\begin{aligned}\operatorname{Var}(\tilde{G}(b,x,i)) &\le \frac{1}{L}E\left(\frac{\mathbb{1}(I=i)}{i}\sum_{p=1}^{i}\mathbb{1}(B_p\le b)\frac{1}{h_G^d}K_G\left(\frac{x-X}{h_G},0\right)\right)^2\\ &\le \frac{1}{Lih_G^{2d}}\sum_{p=1}^{i}E\left(\mathbb{1}(I=i)\mathbb{1}(B_p\le b)K_G\left(\frac{x-X}{h_G},0\right)\right)^2\\ &= \frac{1}{Lh_G^d}\int K_G^2(y,0)G(b,x-h_Gy,i)\,dy.\end{aligned}$$

The desired bound follows from A4-(ii) and (B.6). Note the somewhat different bound for $\operatorname{Var}(\tilde{f}(v,x))$ because of the suboptimality of h_f. *Q.E.D.*

The next lemma derives some exponential-type inequalities for the probabilities of deviations of $\tilde{G}(v,x,i)-G(b,x,i)$, $\tilde{g}(b,x,i)-g(b,x,i)$, $\tilde{f}(v,x)-f(v,x)$, and $\hat{f}(x)-f(x)$ in supnorm over $\mathscr{C}_i(B)$, $\mathscr{C}_i(B)$, $\mathscr{C}(V)$, and $\mathscr{C}(X)$, respectively. To this end, we need to introduce coverings of these sets. For instance, $\mathscr{C}_i(B)$ is covered by N inner "balls" of the form

$$\mathscr{B}_{in}\equiv\mathscr{B}_i((b_n,x_n);\Delta)=\{(b,x)\in S_i(G):b\in[b_n-\Delta,b_n+\Delta],\ x\in[x_n-\Delta,x_n+\Delta]\},$$

where $\Delta>0$, and $(b_n,x_n)\in\mathscr{C}_i(B)$ for $n=1,\dots,N$. Moreover, we consider minimal coverings, i.e., coverings for which N is the smallest number denoted $N(\mathscr{C}_i(B),\Delta)$. Similar notions apply to the sets $\mathscr{C}(V)$ and $\mathscr{C}(X)$.

Hereafter, when there is no possible confusion, we simplify the notation by omitting the * in $|\cdot|_{r,*}$ when the supnorm is taken over the whole support of the function. Let

$$\begin{aligned}e_G(t,\tau) &= t+\frac{r_G|K_G|_0}{h_G^d}t\sqrt{\frac{\log L}{L}}+\lambda_G^{R+1}M_G^{R+1}|G|_{R+1}\\ &\quad+\tau(2d|K_G|_1+4h_G|K_G|_0|g(b,i)|_0),\\ e_g(t,\tau) &= t+2d|K_g|_1\tau+\lambda_g^RM_g^R|g|_R,\\ e_f(t,\tau) &= t+2d|K_f|_1\tau+\lambda_f^RM_f^R|f(v,x)|_R,\\ e_X(t,\tau) &= t+2d|K_X|_1\tau+\lambda_X^{R+1}M_X^{R+1}|f(x)|_{R+1},\\ e_g^*(t,\tau) &= t+2d|K_g|_1\tau+\lambda_g^{R+1}M_g^{R+1}|g|_{R+1,\mathscr{C}'(B)},\end{aligned}$$

where t and τ are arbitrary strictly positive numbers. Define

$$\begin{aligned}P_G(t,\tau) &= 2N(\mathscr{C}(B),\tau h_G^{d+1}/r_G)\exp\left(-\frac{\lambda_G^dt^2\log L}{2Q_G|G|_0+4t|K_G|_0/(3r_G)}\right),\\ P_G'(t,\tau) &= 2N(\mathscr{C}(B),\tau h_G^{d+1}/r_G)\exp\left(-\frac{t^2\log L}{4|g(b,i)|_0\tau h_G^{d+1}/r_G+2t\sqrt{\log L/L}/3}\right),\\ P_g(t,\tau) &= 2N(\mathscr{C}(B),\tau h_g^{d+2}/r_g)\exp\left(-\frac{(\lambda_g^{d+3}t^2\log L)/h_g^2}{2Q_g|g|_0+4t|K_g|_0/(3r_g)}\right),\end{aligned}$$

$$P_f(t,\tau)=2N(\mathscr{C}(V),\tau h_f^{d+2}/r_f)\exp\left(-\frac{(\lambda_f^{d+3}t^2\log L)/h_f^2}{2Q_f|f(v,x)|_0+4t|K_f|_0/(3r_f)}\right),$$

$$P_X(t,\tau)=2N(\mathscr{C}(X),\tau h_X^{d+1}/r_X)\exp\left(-\frac{\lambda_X^d t^2\log L}{2Q_X|f(x)|_0+4t|K_X|_0/(3r_X)}\right),$$

$$P_g^*(t,\tau)=2N(\mathscr{C}(B),\tau h_g^{d+2}/r_g^*)\exp\left(-\frac{\lambda_g^{d+1}t^2\log L}{2Q_g|g|_0+4t|K_g|_0/(3r_g^*)}\right).$$

LEMMA C4: *Given A1–A4, for any $t>0$, $\tau>0$ and $i\in\mathscr{I}$, we have for L sufficiently large*

(i) $\Pr(r_G|\tilde{G}-G|_{0,\mathscr{C}_i(B)}>e_G(t,\tau))\le P_G(t,\tau)+P'_G(t,\tau)$,

$\Pr(r_g|\tilde{g}-g|_{0,\mathscr{C}_i(B)}>e_g(t,\tau))\le P_g(t,\tau)$,

$\Pr(r_f|\tilde{f}-f|_{0,\mathscr{C}(V)}>e_f(t,\tau))\le P_f(t,\tau)$,

$\Pr(r_X|\hat{f}-f|_{0,\mathscr{C}(X)}>e_f(t,\tau))\le P_X(t,\tau)$,

(ii) $\Pr(r_g^*|\tilde{g}-g|_{0,\mathscr{C}_i(B)}>e_g^*(t,\tau))\le P_g^*(t,\tau)$.

PROOF OF LEMMA C4: We detail the proof for $\tilde{G}$, as it is the most involved.

Step 1: From Lemma C2 and the triangular inequality, we obtain

$$\begin{aligned}(\text{C.10})\quad &\Pr(r_G|\tilde{G}-G|_{0,\mathscr{C}_i(B)}>e_G(t,\tau))\\ &\le\Pr\Big(r_G|\tilde{G}-E\tilde{G}|_{0,\mathscr{C}_i(B)}+r_G|E\tilde{G}-G|_{0,\mathscr{C}_i(B)}>e_G(t,\tau)\Big)\\ &\le\Pr(r_G|\tilde{G}-E\tilde{G}|_{0,\mathscr{C}_i(B)}>e_G(t,\tau)-\lambda_G^{R+1}M_G^{R+1}|G|_{R+1}).\end{aligned}$$

Note that, for L sufficiently large, $\tilde{G}(b,x,i)-E\tilde{G}(b,x,i)=(1/L)\Sigma_{m=1}^L\zeta_{mL}(b,x,i)$, where

$$\zeta_{mL}(b,x,i)=\frac{1}{ih_G^d}\sum_{p=1}^{i}\left(\mathbb{1}(B_{pm}<b)K_G\left(\frac{x-X_m}{h_G},0\right)\mathbb{1}(I_m=i)\right.$$
$$\left.-E\left(\mathbb{1}(B<b)K_G\left(\frac{x-X}{h_G},0\right)\mathbb{1}(I=i)\right)\right).$$

The ζ_{mL}'s, $m=1,\dots,L$, are i.i.d. centered variables. Moreover, using Lemma C3, we have

$$|r_G\zeta_{mL}(b,x,i)|\le\frac{2|K_G|_0r_G}{h_G^d}=\frac{2L|K_G|_0}{\lambda_G^dr_G\log L},$$

$$\operatorname{Var}(r_G\zeta_{mL}(b,x,i))=Lr_G^2\operatorname{Var}(\tilde{G}(b,x,i))\le\frac{LQ_G|G|_0}{\lambda_G^d\log L}$$

for any $(b,x,i)\in\mathscr{C}(B)$. Hence, Bernstein inequality (see Serfling (1980, p. 95)) gives

$$\begin{aligned}(\text{C.11})\quad \Pr(r_G|\tilde{G}(b,x,i)-E\tilde{G}(b,x,i)|>t)&\le 2\exp\left(-\frac{t^2\log L}{2Q_G|G|_0/\lambda_G^d+(4t|K_G|_0)/(3\lambda_G^dr_G)}\right)\\ &=\frac{P_G(t,\tau)}{N(\mathscr{C}(B),\tau h_G^{d+1}/r_G)}\end{aligned}$$

for any $(b,x,i)\in\mathscr{C}(B)$, t, and L, where the equality follows from the definition of $P_G(t,\tau)$.

Step 2: Consider a minimal covering of $\mathscr{C}_i(B)$ for some $\Delta > 0$. For any $(b, x) \in \mathscr{B}_{in}$, we have

$$
\begin{aligned}
&r_G|\tilde{G}(b,x,i) - E\tilde{G}(b,x,i)| \\
&\quad \le \sup_{1\le n\le N}\left|\frac{r_G}{L}\sum_{m=1}^{L}\zeta_{mL}(b_n,x_n,i)\right| \\
&\qquad + \sup_{1\le n\le N}\ \sup_{(b,x)\in\mathscr{B}_{in}}\left|\frac{r_G}{L}\sum_{m=1}^{L}(\zeta_{mL}(b_n,x_n,i)-\zeta_{mL}(b,x,i))\right|.
\end{aligned}
$$

This gives

$$
\begin{aligned}
\text{(C.12)}\quad &\Pr(r_G|\tilde{G} - E\tilde{G}|_{0,\mathscr{C}_i(B)} > e_G(t,\tau) - \lambda_G^{R+1}M_G^{R+1}|G|_{R+1}) \\
&\le \Pr\left(\sup_{1\le n\le N} r_G|\tilde{G}(b_n,x_n,i) - E\tilde{G}(b_n,x_n,i)| > t\right) \\
&\quad + \Pr\left(\sup_{1\le n\le N}\ \sup_{(b,x)\in\mathscr{B}_{in}}\left|\frac{r_G}{L}\sum_{m=1}^{L}(\zeta_{mL}(b_n,x_n,i)-\zeta_{mL}(b,x,i))\right|\right. \\
&\qquad\qquad \left. > e_G(t,\tau) - t - \lambda_G^{R+1}M_G^{R+1}|G|_{R+1}\right).
\end{aligned}
$$

We now give some bounds for the increments at (b_n, x_n, i). For any $(b, x) \in \mathscr{B}_{in}$, we have

$$
\begin{aligned}
\text{(C.13)}\quad &\left|\frac{\mathbb{1}(B\le b_n)}{h_G^d}K_G\left(\frac{x_n - X}{h_G},0\right) - \frac{\mathbb{1}(B\le b)}{h_G^d}K_G\left(\frac{x-X}{h_G},0\right)\right| \\
&\le \frac{d\Delta|K_G|_1}{h_G^{d+1}} + \frac{|K_G|_0}{h_G^d}\mathbb{1}(b_n - \Delta \le B \le b_n + \Delta).
\end{aligned}
$$

Using (C.13) twice in the definition of $\zeta_{mL}(b, x, i)$ and the triangular inequality, we obtain

$$
\begin{aligned}
&|\zeta_{mL}(b_n,x_i,i) - \zeta_{mL}(b,x,i)| \\
&\quad \le \frac{1}{i}\sum_{p=1}^{i}\left(\frac{d\Delta|K_G|_1}{h_G^{d+1}} + \frac{|K_G|_0}{h_G^d}\mathbb{1}(b_n - \Delta\le B_{pm}\le b_n+\Delta)\right)\mathbb{1}(I_m = i) \\
&\qquad + E\left[\left(\frac{d\Delta|K_G|_1}{h_G^{d+1}} + \frac{|K_G|_0}{h_G^d}\mathbb{1}(b_n-\Delta\le B\le b_n+\Delta)\right)\mathbb{1}(I=i)\right] \\
&\quad \le \frac{2d\Delta|K_G|_1}{h_G^{d+1}} + \frac{|K_G|_0}{h_G^d}(2E[\mathbb{1}(b_n-\Delta\le B\le b_n+\Delta)\mathbb{1}(I=i)] + Z_{mL}(b_n))
\end{aligned}
$$

where we have used $(1/i)\sum_{p=1}^{i}\mathbb{1}(I_m = i) \le 1$, $E[\mathbb{1}(I = i)] \le 1$, and

$$
\begin{aligned}
Z_{mL}(b) \equiv \frac{1}{i}\sum_{p=1}^{i}(&\mathbb{1}(b-\Delta\le B_{pm}\le b+\Delta)\mathbb{1}(I_m = i) \\
&- E[\mathbb{1}(b-\Delta\le B\le b+\Delta)\mathbb{1}(I=i)]).
\end{aligned}
$$

Because $E[\mathbb{1}(b_n - \Delta \leq B \leq b_n + \Delta)\mathbb{1}(I = i)] \leq 2\Delta |g(b,i)|_0$ it follows that

$$\text{(C.14)} \qquad \sup_{1 \leq n \leq N} \sup_{(b,x) \in \mathscr{B}_{in}} \left| \frac{r_G}{L} \sum_{m=1}^{L} (\zeta_{mL}(b_n, x_n, i) - \zeta_{mL}(b, x, i)) \right|$$

$$\leq \frac{2dr_G \Delta |K_G|_1}{h_G^{d+1}} + \frac{4r_G \Delta |K_G|_0 |g(b,i)|_0}{h_G^d} + \sup_{1 \leq n \leq N} \frac{r_G |K_G|_0}{L h_G^d} \sum_{m=1}^{L} Z_{mL}(b_n).$$

Step 3: We study $\sup_{1 \leq n \leq N}(1/L)\Sigma_{m=1}^{L} Z_{mL}(b_n)$. We have $|Z_{mL}(b_n)| \leq 1$ while $\text{Var}(Z_{mL}(b_n)) \leq \text{Var}[\mathbb{1}(b_n - \Delta \leq B \leq b_n + \Delta)\mathbb{1}(I = i)] \leq 2\Delta |g(b,i)|_0$. Thus Bernstein inequality gives

$$\text{(C.15)} \qquad \Pr\left(\sup_{1 \leq n \leq N} \left| \frac{1}{L} \sum_{m=1}^{L} Z_{mL}(b_n) \right| > t\sqrt{\log L/L} \right)$$

$$\leq \sum_{n=1}^{N} \Pr\left(\left| \frac{1}{L} \sum_{m=1}^{L} Z_{mL}(b_n) \right| > t\sqrt{\log L/L} \right)$$

$$\leq 2N(\mathscr{C}(B), \Delta) \exp\left(- \frac{t^2 \log L}{4|g(b,i)|_0 \Delta + 2t\sqrt{\log L/L}/3} \right).$$

Step 4: Let $\Delta = \tau h_G^{d+1}/r_G$. Hence, (C.14) and the definition of $e_G(t,\tau)$ imply

$$\Pr\left(\sup_{1 \leq n \leq N} \sup_{(b,x) \in \mathscr{B}_{in}} \left| \frac{r_G}{L} \sum_{m=1}^{L} (\zeta_{ml}(b_n, x_n, i) - \zeta_{ml}(b, x, i)) \right| \right.$$

$$\left. > e_G(t,\tau) - t - \lambda_G^{R+1} M_G^{R+1} |G|_{R+1} \right)$$

$$\leq \Pr\left(\sup_{1 \leq n \leq N} \left| \frac{r_G |K_G|_0}{L h_G^d} \sum_{m=1}^{L} Z_{mL}(b_n) \right| > \frac{r_G |K_G|_0}{h_G^d} t\sqrt{\log L/L} \right)$$

$$\leq 2N(\mathscr{C}(B), \tau h_G^{d+1}/r_G) \exp\left(- \frac{t^2 \log L}{4|g(b,i)|_0 \tau h_G^{d+1}/r_G + 2t\sqrt{\log L/L}/3} \right)$$

$$= P_G'(t,\tau),$$

using (C.15). The desired result now follows from (C.10), (C.12), and (C.11) as

$$\Pr(r_G |\tilde{G} - G|_{0, \mathscr{C}_i(B)} > e_G(t,\tau))$$

$$\leq \sum_{n=1}^{N} \Pr(r_G |\tilde{G}(b_n, x_n, i) - E\tilde{G}(b_n, x_n, i)| > t) + P_G'(t,\tau)$$

$$= P_G(t,\tau) + P_G'(t,\tau).$$

Proofs of the other inequalities of the lemma are simpler because Step 3 can be dropped. Indeed, the corresponding bounds (C.13) only involve the derivative of the kernel. Thus the new bounds in (C.14) only depend upon the first term. For instance, we choose $\Delta = \tau h_g^{d+2}/r_g$ to prove the second bound, etc. *Q.E.D.*

PROOF OF LEMMA B2: The covering number N is of order $\Delta^{-\delta}$, where δ is the dimension of the covered set. Hence, in view of A4-(ii) and (B.6), the various covering numbers in the upper bounds of

Lemma C4 are of the order $(L/\log L)^{\eta}$ for some $\eta > 0$. For instance, $N(\mathscr{C}(B), \tau h_G^{d+1}/r_G) = O(L/\log L)^{(d+1)(R+d+2)/(2R+d+2)}$. Thus, by taking t sufficiently large, it is easy to see that the series $P_G(t,\tau)$ (say) can be made convergent as $L \to \infty$. The desired result follows from the Borel-Cantelli Lemma and the fact that $e_G(t,\tau) = O(1)$. *Q.E.D.*

REFERENCES

AHN, H. (1995): "Nonparametric Two-Stage Estimation of Conditional Probabilities in a Binary Choice Model under Uncertainty," *Journal of Econometrics*, 67, 337–378.

BIERENS, H. J. (1987): "Kernel Estimators of Regression Functions," in *Advances in Econometrics*, ed. by T. Bewley. Cambridge: Cambridge University Press.

BLUM, J. R., J. KIEFER, AND M. ROSENBLATT (1961): "Distribution Free Tests of Independence Based on the Sample Distribution Functions," *Annals of Mathematical Statistics*, 32, 485–498.

CAMPO, S., I. PERRIGNE, AND Q. VUONG (1998): "Asymmetry and Joint Bidding in OCS Wildcat Auctions," mimeo, University of Southern California.

DONALD, S., AND H. PAARSCH (1993): "Piecewise Pseudo-Maximum Likelihood Estimation in Empirical Models of Auctions," *International Economic Review*, 34, 121–148.

——— (1996): "Identification, Estimation, and Testing in Parametric Empirical Models of Auctions within the Independent Private Values Paradigm," *Econometric Theory*, 12, 517–567.

ELYAKIME, B., J. J. LAFFONT, P. LOISEL, AND Q. VUONG (1994): "First-Price Sealed-Bid Auctions with Secret Reservation Prices," *Annales d'Economie et de Statistiques*, 34, 115–141.

——— (1997): "Auctioning and Bargaining: An Econometric Study of Timber Auctions with Secret Reservation Prices," *Journal of Business Economics and Statistics*, 15, 209–220.

FRIEDMAN, L. (1956): "A Competitive Bidding Strategy," *Operations Research*, 4, 104–112.

GEFFROY, J. (1984): "Sur un Problème d'Estimation Géométrique," *Publication de l'Institut Statistique Universitaire de Paris*, 13, 191–220.

GRAHAM, D., AND R. MARSHALL (1987): "Collusive Bidder Behavior at Single Object Second-Price and English Auctions," *Journal of Political Economy*, 95, 1217–1239.

GROENEBOOM, P. (1996): "Lectures on Inverse Problems," in *Lectures on Probability and Statistics*, Lecture Notes in Mathematics, No. 1648, ed. by P. Bernard. New York: Springer Verlag.

GUERRE, E., I. PERRIGNE, AND Q. VUONG (1995): "Nonparametric Estimation of First-Price Auctions," mimeo, University of Southern California.

HARDLE, W. (1991): *Smoothing Techniques with Implementation in S*. New York: Springer Verlag.

HARSANYI, J. (1967): "Games with Incomplete Information Played by Bayesian Players," *Management Science*, 14, 159–182, 320–334, 486–502.

HASSANI, S., P. SARDA, AND P. VIEU (1986): "Approche Non Paramétrique en Théorie de la Fiabilité: Revue Bibliographique," *Revue de Statistiques Appliquées*, 35, 27–41.

HENDRICKS, K., AND R. H. PORTER (1988): "An Empirical Study of an Auction with Asymmetric Information," *American Economic Review*, 78, 865–883.

HENDRICKS, K., R. H. PORTER, AND C. WILSON (1994): "Auctions for Oil and Gas Leases with an Informed Bidder and a Random Reserve Price," *Econometrica*, 62, 1415–1444.

IBRAGIMOV, I. A., AND R. Z. HAS'MINSKII (1981): *Statistical Estimation. Asymptotic Theory*. New York: Springer Verlag.

KHAS'MINSKII, R. Z. (1976): "A Lower Bound on the Risks of Nonparametric Estimates of Densities in the Uniform Metric," *Theory of Probability and its Applications*, 23, 794–798.

KOROSTELEV, A. P., AND A. B. TSYBAKOV (1993): *Minimax Theory of Image Reconstruction, Lecture Notes in Statistics* 82. New York: Springer Verlag.

LAFFONT, J. J. (1997): "Game Theory and Empirical Economics: The Case of Auction Data," *European Economic Review*, 41, 1–35.

LAFFONT, J. J., T. LI, AND Q. VUONG (1999): "Identification of Asymmetric Independent Private Value Auction Models," mimeo, University of Southern California.

LAFFONT, J. J., H. OSSARD, AND Q. VUONG (1995): "Econometrics of First-Price Auctions," *Econometrica*, 63, 953–980.

LAFFONT, J. J., A. OUSTRY, M. SIMIONI, AND Q. VUONG (1996): "Econometrics of Optimal Procurement Auctions," mimeo, University of Toulouse.

LAFFONT, J. J., AND J. TIROLE (1993): *A Theory of Incentives in Procurement and Regulation*. Cambridge, MA: MIT Press.

LAFFONT, J. J., AND Q. VUONG (1993): "Structural Analysis of Descending Auctions," *European Economic Review*, 37, 329–341.

——— (1996): "Structural Analysis of Auction Data," *American Economic Review, Papers and Proceedings*, 86, 414–420.

LI, T., I. PERRIGNE, AND Q. VUONG (1996): "Affiliated Private Values in OCS Wildcat Auctions," mimeo, University of Southern California.

——— (1999): "Conditionally Independent Private Information in OCS Wildcat Auctions," *Journal of Econometrics*, forthcoming.

LI, T., AND Q. VUONG (1997): "Using All Bids in Parametric Estimation of First-Price Auctions," *Economic Letters*, 55, 321–325.

MASKIN, E., AND J. RILEY (1983): "Auctions with Asymmetric Beliefs," mimeo, University of California, Los Angeles.

MCAFEE, R. P., AND D. VINCENT (1992): "Updating the Reserve Price in Common Value Auctions," *American Economic Review, Papers and Proceedings*, 82, 512–518.

MILGROM, P., AND R. WEBER (1982): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50, 1089–1122.

NADAYARA, E. A. (1964): "On Estimating Regression," *Theory of Probability and its Applications*, 9, 141–142.

PAARSCH, H. J. (1992): "Deciding Between the Common and Private Value Paradigms in Empirical Models of Auctions," *Journal of Econometrics*, 51, 191–215.

PORTER, R. (1995): "The Role of Public Information in U.S. Offshore Oil and Gas Lease Auctions," *Econometrica*, 63, 1–28.

RILEY, J. G., AND W. F. SAMUELSON (1981): "Optimal Auctions," *American Economic Review*, 71, 381–392.

ROEHRIG, C. S. (1988): "Conditions for Identification in Nonparametric and Parametric Models," *Econometrica*, 56, 433–447.

SALANIÉ, B. (1997): *The Economics of Contracts*. Cambridge, MA: MIT Press.

SERFLING, R. J. (1980): *Approximation Theorems of Mathematical Statistics*. New York: Wiley and Sons.

SINGPURWALLA, N. D., AND M. Y. WONG (1983): "Estimation of the Failure Rate: A Survey of Non-parametric Methods. Part I: Non Bayesian Methods," *Communications in Statistics Theory and Methods*, 5, 559–588.

STONE, C. J. (1982): "Optimal Rates of Convergence for Nonparametric Regressions," *Annals of Statistics*, 10, 1040–1053.

SUTTON, J. (1993): "Much Ado about Auctions: Some Introductory Remarks," *European Economic Review*, 37, 319–320.

UTRERAS, F. (1985): "Smoothing Noisy Data under Monotonicity Constraints: Existence, Characterization and Convergence Rates," *Numerische Mathematik*, 47, 611–625.

VICKREY, W. (1961): "Counterspeculation, Auctions, and Sealed Tenders," *Journal of Finance*, 16, 8–37.

WANG, M. C., N. P. JEWELL, AND W. Y. TSAY (1986): "Asymptotic Properties of the Product Limit Estimate Under Random Truncation," *Annals of Statistics*, 14, 1597–1605.

WATSON, G. S. (1964): "Smooth Regression Analysis," *Sankhya*, A 26, 359–372.

WRIGHT, I. W., AND E. J. WEGMAN (1980): "Isotonic, Convex and Related Splines," *Annals of Statistics*, 8, 1023–1035.