Semiparametric Estimation of First-Price Auctions with Risk-Averse Bidders

Author(s): SANDRA CAMPO, EMMANUEL GUERRE, ISABELLE PERRIGNE and QUANG VUONG

Source: *The Review of Economic Studies*, January 2011, Vol. 78, No. 1 (January 2011), pp. 112-147

Published by: Oxford University Press

Stable URL: https://www.jstor.org/stable/23015850

Review of Economic Studies (2011) **78**, 112–147 doi: 10.1093/restud/rdq001

Semiparametric Estimation of First-Price Auctions with Risk-Averse Bidders

SANDRA CAMPO
University of North Carolina at Chapel Hill
EMMANUEL GUERRE
Queen Mary University of London
ISABELLE PERRIGNE
Pennsylvania State University
and
QUANG VUONG
Pennsylvania State University

First version received March 2007; *final version accepted September* 2009 (*Eds.*)

In view of the non-identification of the first-price auction model with risk-averse bidders, this paper proposes some parametric identifying restrictions and a semiparametric estimator for the risk aversion parameter(s) and the latent distribution of private values. Specifically, we exploit heterogeneity across auctioned objects to establish semiparametric identification under a conditional quantile restriction of the bidders' private value distribution and a parameterization of the bidders' utility function. We develop a multistep semiparametric method and we show that our semiparametric estimator of the utility function parameter(s) converges at the optimal rate, which is slower than the parametric one but independent of the dimension of the exogenous variables thereby avoiding the curse of dimensionality. We then consider various extensions including a binding reserve price, affiliation among private values, and asymmetric bidders. The method is illustrated on U.S. Forest Service timber sales, and bidders' risk neutrality is rejected.

Keywords: Risk aversion, Private value, Semiparametric identification, Semiparametric estimation, Optimal rate, Timber auctions

JEL Codes: C14, D44, L70

1. INTRODUCTION

Since the seminal work by Kenneth Arrow and its formalization by Pratt (1964), risk aversion has become a fundamental concept whenever agents face uncertainties. This is the case in auctions especially when bidders' valuations are large relative to their assets. Auction theory has extensively studied risk aversion. Maskin and Riley (1984) and Matthews (1987) derive the optimal auction mechanism that involves complex transfers. Within a private value framework, the main implication of risk aversion is that it induces more aggressive bidding than under risk neutrality. As a corollary, the first-price auction mechanism dominates the ascending one in terms of revenue for the seller since bidding is not affected by risk aversion in the latter thereby providing a rationale for the use of the former. In experimental data where the underlying distribution of private values is known, overbidding over the risk-neutral Nash equilibrium has been exploited to show that bidders are risk averse. For instance, Goeree, Holt and Palfrey (2002) mainly explain deviations from the risk-neutral Nash equilibrium by bidders' risk aversion, while Bajari and Hortacsu (2005) show that a risk aversion model provides the best fit to experimental bid data among several competing models. Despite its importance, few studies have assessed the extent of bidders' risk aversion on field auction data. This is because auction theory does not provide testable implications of risk aversion that can be used within a reduced-form approach.[1] This calls for a structural approach that assumes that observed bids are the Bayesian Nash equilibrium outcomes and is thus well adapted to assess and test for bidders' risk aversion.[2] A major difficulty, however, lies in the non-identification of the general first-price auction model with risk-averse bidders as shown by Guerre, Perrigne and Vuong (2009).

Our paper proposes a semiparametric approach to identify the auction model with risk aversion based on minimal parametric restrictions.[3] Throughout, we consider first-price sealed-bid auctions with risk-averse bidders within the private value paradigm. The first part of our paper presents the benchmark model with independent private values. The structural elements are the bidders' von Neuman Morgenstern (vNM) utility function and the bidders' private value distribution. First, we show that any smoothed bid distribution can be rationalized by some constant relative risk aversion (CRRA) model or some constant absolute risk aversion (CARA) model. Such a striking result arises from the weak restrictions imposed on observed bids by the game theoretical model. Second, we show that the model is non-identified from observed bids even when the utility function is restricted to belong to well-known families of risk aversion such as CRRA or CARA. Third, since little is known on the utility function, an alternative identifying strategy is to parameterize the private value distribution, while leaving the utility function non-parametric. Again, we show that this model is not identified from observed bids in general.

The second part of our paper seeks weak and palatable restrictions that lead to identification. Specifically, we exploit heterogeneity across auctioned objects through a parametric quantile restriction of the private value distribution, while restricting the bidders' vNM utility function to be parametric. Under these parametric restrictions, we show that the utility function parameter(s) and the conditional private value distribution can be semiparametrically identified. We show that dropping either one of these two conditions looses identification. In this sense, our semiparametric modelling is minimal. The third part of the paper characterizes the best (optimal) rate that any estimator of the risk aversion parameter(s) can achieve relying on the minimax theory developed by, *e.g.*, Ibragimov and Has'minskii (1981). Because estimation of an upper boundary can be achieved at a faster rate than for any other quantile, we focus on parameterizing the upper boundary. When auctioned objects' heterogeneity is characterized by d continuous variables and the underlying density is R continuously differentiable, we show that the optimal rate for estimating the risk aversion parameter(s) is $N^{(R+1)/(2R+3)}$. Though slower than $\sqrt{N}$, this rate is independent of d thereby avoiding the curse of dimensionality associated with non-parametric estimators.

The fourth part of the paper develops a multistep semiparametric estimator. The first step consists in estimating non-parametrically the conditional bid density at its upper boundary. The

1. Notable examples are Baldwin (1995) and Athey and Levin (2001) who suggest that bidding diversification across species in U.S. Forest Service (USFS) auctions is consistent with bidders' risk aversion.

2. See Perrigne and Vuong (1999) and Athey and Haile (2007) for recent surveys.

3. Guerre, Perrigne and Vuong (2009) develop a different identification approach based on exclusion restrictions. In particular, they consider an exogenous bidders' participation leading to an underlying private value distribution independent of the number of bidders. Lu and Perrigne (2008) propose another identification approach by combining data from ascending and first-price auctions. Ackerberg, Hirano and Shahriar (2006) adopt a fully parametric approach to estimate risk aversion and impatience in eBay auctions by exploiting the buy-it-now option.

second step combines weighted non-linear least squares (NLLS) with the non-parametric estimates obtained in the first step to estimate the utility function parameter(s). The third step recovers the bidders' private values and their conditional density following Guerre, Perrigne and Vuong (2000). We show that our estimator of the utility function parameter(s) attains the optimal rate $N^{(R+1)/(2R+3)}$. This contrasts with most $\sqrt{N}$-consistent semiparametric estimators developed in the econometric literature as surveyed by Newey and McFadden (1994) and Powell (1994).[4] A notable feature of our estimation problem is that the variance of the error term in the non-linear regression diverges with N thereby leading to a non-standard convergence rate for our semiparametric estimator.

The fifth part of our paper studies extensions of the benchmark model. We show that our identification results extend to a binding reserve price, affiliated private values, and asymmetric auctions when asymmetry arises from different private value distributions. In contrast, asymmetry arising from different utility functions provides additional restrictions that help to identify the model. Our estimation method can be readily adapted to the first three cases, while estimation of the fourth case is briefly discussed. One advantage of our method is its computational simplicity as it circumvents both the numerical determination and the inversion of the equilibrium bidding strategy. This is especially convenient when there is no closed-form solution to the differential equations defining the equilibrium strategies such as for general risk aversion and asymmetric bidders. We then illustrate our procedure on the USFS timber auctions. In particular, bidders' risk neutrality is rejected.

The paper is organized as follows. Section 2 presents the benchmark model and some non-identification results in a semiparametric context. This leads to the identifying restrictions in Section 3. Section 4 provides an upper bound for the optimal convergence rate that can be attained by semiparametric estimators of the utility function parameter(s). Section 5 presents our semiparametric estimation procedure with its statistical properties. Section 6 considers extensions of the benchmark model and discusses their semiparametric identification and estimation. Section 7 illustrates our method to timber auction data. Section 8 concludes. Three appendices collect the proofs.

2. MODEL AND NON-IDENTIFICATION RESULTS

This section introduces the independent private value (IPV) first-price sealed-bid auction model with risk-averse bidders. Bidders' private values v_i are drawn independently from a distribution $F(\cdot \mid I)$, which is absolutely continuous with density $f(\cdot \mid I)$ on a compact support $[\underline{v}(I), \overline{v}(I)] \subset \boldsymbol{R}_+$. The distribution $F(\cdot \mid I)$ and the number of potential bidders $I \geq 2$ are common knowledge.[5] Let $U(\cdot)$ be the bidders' vNM utility function with $U(0) = 0$, $U'(\cdot) > 0$, and $U''(\cdot) \leq 0$ because of potential risk aversion. All bidders are thus identical *ex ante* and the game is symmetric. Bidder i maximizes his expected utility $\mathrm{E}\Pi_i = U(v_i - b_i)\Pr(b_i \geq b_j, j \neq i)$ with respect to his bid b_i, where b_j is the jth player's bid. Because the scale is irrelevant, we impose the normalization $U(1) = 1$. The risk-neutral case is obtained when $U(\cdot)$ is the identity function.[6]

4. Notable exceptions of semiparametric estimators converging at a slower rate than $\sqrt{N}$ are those proposed by Horowitz (1992), Kyriazidou (1997), and Honore and Kyriazidou (2000) though the latter suffers from the curse of dimensionality.

5. The dependence of the private value distribution on I captures the idea that private values and I can be dependent. For instance, objects of higher value may attract more bidders. The number of bidders may capture some unobserved heterogeneity. See Campo, Perrigne and Vuong (2003). It may also result from endogenous participation.

6. Bidders' wealth w can be readily introduced in the model. In this case, the expected profit becomes $[U(w + v_i - b_i) - U(w)]\Pr(b_i \geq b_j, j \neq i) + U(w)$. On the other hand, allowing different wealths w_i leads to an asymmetric game if the w_is are common knowledge and to a multisignal game if the w_is are private information. The first case is studied in Section 6, while the second case is beyond the scope of this paper. For multisignals, see Che and Gale (1998) for a model with budget constraints.

From Maskin and Riley (1984), if a symmetric Bayesian Nash equilibrium strategy $s(\cdot, U, F, I)$ exists, then it is strictly increasing, continuous, and differentiable. Thus, equation (1) becomes $E\Pi_i = U(v_i - b_i)F^{I-1}(s^{-1}(b_i) \mid I)$, where $s^{-1}(\cdot)$ denotes the inverse of $s(\cdot)$. Hence, imposing bidder i's optimal bid b_i to be $s(v_i)$ gives the following differential equation:

$$s'(v_i) = (I-1)\frac{f(v_i \mid I)}{F(v_i \mid I)}\lambda(v_i - b_i) \tag{1}$$

for all $v_i \in [\underline{v}(I), \overline{v}(I)]$, where $\lambda(\cdot) = U(\cdot)/U'(\cdot)$. As shown by Maskin and Riley (1984), the boundary condition is $U(\underline{v}(I) - s(\underline{v}(I))) = 0$, *i.e.*, $s(\underline{v}(I)) = \underline{v}(I)$ because $U(0) = 0$. We assume that $U(\cdot)$ and $F(\cdot \mid I)$ satisfy some smoothness conditions as given in Definitions 1 and 2 in Guerre, Perrigne and Vuong (2009), *i.e.*, $U(\cdot) \in \mathcal{U}_R$ and $F(\cdot \mid I) \in \mathcal{F}_R$.[7]

In a first-price sealed-bid auction with a non-binding reserve price, the number I of bidders is observed. The distribution $G(\cdot \mid I)$ of an equilibrium bid is known. Knowledge of $G(\cdot \mid I)$ from observed bids is an estimation problem, which is addressed in Section 5. Following Guerre, Perrigne and Vuong (2000), we express equation (1) using the bid distribution $G(\cdot \mid I)$. For every $b \in [\underline{b}(I), \overline{b}(I)] = [\underline{v}(I), s(\overline{v}(I))]$, we have $G(b \mid I) = F(s^{-1}(b) \mid I) = F(v \mid I)$ with density $g(b \mid I) = f(v \mid I)/s'(v)$, where $v = s^{-1}(b)$. Thus, equation (1) can be written as

$$1 = (I-1)\frac{g(b_i \mid I)}{G(b_i \mid I)}\lambda(v_i - b_i). \tag{2}$$

Since $\lambda(\cdot)$ is strictly increasing as $\lambda'(\cdot) = 1 - (U(\cdot)U''(\cdot)/U'(\cdot)^2) \geq 1$, equation (2) gives

$$v_i = b_i + \lambda^{-1}\left(\frac{1}{I-1}\frac{G(b_i \mid I)}{g(b_i \mid I)}\right) \equiv \xi(b_i, U, G, I), \tag{3}$$

where $\lambda^{-1}(\cdot)$ denotes the inverse of $\lambda(\cdot)$. This equation expresses each bidder's private value as a function of his corresponding bid, the bid distribution, the number of bidders, and the utility function. Note that $\xi(\cdot)$ is the inverse of the bidding strategy $s(\cdot)$. The bid distribution $G(\cdot \mid I)$ satisfies some smoothness properties embodied in $\mathcal{G}_R$ implied by the smoothness of $s(\cdot)$ and the smoothness of $[U, F]$. See Theorem 1, Definition 3, and Lemma 1(i) in Guerre, Perrigne and Vuong (2009).[8]

Guerre, Perrigne and Vuong (2009, Proposition 1) show that the auction model with general risk aversion imposes weak restrictions on observables since a bid distribution $G(\cdot \mid I)$ can be rationalized by a risk-averse structure $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$ if and only if $G(\cdot \mid I) \in \mathcal{G}_R$. The first question is whether such a result still holds if we restrict ourselves to some parametric families for $U(\cdot)$. Specifically, for $F(\cdot \mid I) \in \mathcal{F}_R$, we consider CRRA structures $[U, F]$ with $U(x) = x^{1-c}$ for $0 \leq c < 1$ and CARA structures with $U(x) = (1 - \exp(-ax))/(1 - \exp(-a))$ for $a > 0$.

Proposition 1. *Let $I \geq 2$ and $R \geq 1$. Any distribution $G(\cdot \mid I) \in \mathcal{G}_R$ can be rationalized by some CRRA structure as well as some CARA structure with $F(\cdot \mid I) \in \mathcal{F}_R$.*

7. Specifically, in addition to the aforementioned properties, $U(\cdot)$ is $R+2$ continuously differentiable and $\lim_{x \downarrow x} \lambda^{(r)(x)}$ is finite for $1 \leq r \leq R+1$, while $F(\cdot \mid I)$ is $R+1$ continuously differentiable with $f(\cdot \mid I)$ bounded away from zero.

8. Specifically, the bid distribution is $R+1$ continuously differentiable on its support $[\underline{b}(I), \overline{b}(I)]$ with $0 \leq \underline{b}(I) < \overline{b}(I) < +\infty$. Its density $g(\cdot \mid I)$ is bounded away from zero on its support and $R+1$ continuously differentiable on $(\underline{b}(I), \overline{b}(I)]$. Moreover, $\lim_{b \downarrow \underline{b}(I)} d^r[G(b \mid I)/g(b \mid I)]/db^r$ exists and is finite for $1 \leq r \leq R+1$. The fact that $g(\cdot \mid I)$ is smoother than $f(\cdot \mid I)$ is a property that is specific to auction models. A large class of bid distributions satisfy such properties.

Proposition 1 is striking and strengthens Guerre, Perrigne and Vuong (2009, Proposition 1). Our result indicates that a single parameter family of utility functions such as CRRA or CARA is sufficient to rationalize any bid distribution $G(\cdot \mid I) \in \mathcal{G}_R$. Second, because a structure $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$ leads to a bid distribution $G(\cdot \mid I) \in \mathcal{G}_R$, Proposition 1 implies that there always exist some CARA and CRRA structures that are observationally equivalent to the general risk aversion structure $[U, F]$. In particular, one cannot discriminate between a CRRA or a CARA model.[9] Third, Proposition 1 implies that any risk-neutral model is observationally equivalent to a CRRA or a CARA model. The converse, *i.e.*, whether any risk-averse model is observationally equivalent to a risk-neutral model, is not true.[10] Thus, by allowing for risk aversion, one does enlarge the set of rationalizable bid distributions relative to risk neutrality.

We now turn to identification. A model is a set of structures $[U, F]$. Hereafter, a structure $[U, F]$ is "non-identified" if there exists another structure $[\tilde{U}, \tilde{F}]$ within the model that leads to the same equilibrium bid distribution. If no such $[\tilde{U}, \tilde{F}]$ exists for any $[U, F]$, the model is (globally) identified. In view of the non-identification of the general risk aversion model established in Guerre, Perrigne and Vuong (2009, Proposition 2), additional restrictions must be imposed to achieve identification. Two natural parametric identifying strategies are available. First, since little is known about $U(\cdot)$, while $F(\cdot \mid I)$ is approximately log-normal in empirical studies, we can think of parameterizing the private value distribution as $F(\cdot \mid I; \gamma)$ giving the semiparametric model $\mathcal{U}_R \times \mathcal{F}(\Gamma)$. The next proposition states that parameterizing $F(\cdot)$ is not sufficient.

Proposition 2. *Let $I \geq 2$ and $R \geq 1$. The semiparametric model $\mathcal{U}_R \times \mathcal{F}(\Gamma)$ is not necessarily identified.*

It is sufficient to exhibit a non-identified semiparametric model. Let $I = 2$ and $\mathcal{F}(\Gamma) = \{F(v; \gamma) = v/\gamma, v \in [0, \gamma], \gamma \in \Gamma = (1, 2]\}$. Any structure $[U, F]$ with $U(x) = x^{\gamma - 1}$ leads to a uniform bid distribution $G(\cdot)$ on $[0, 1]$ as solving equation (1) gives $s(v) = v/\gamma$. Thus, there exists an infinity of structures $[U, F] \in \mathcal{U}_R \times \mathcal{F}(\Gamma)$ leading to the same bid distribution. More generally, using the monotonicity of the equilibrium strategy, equation (3) evaluated at the α-quantile $b(\alpha; I)$ of $G(\cdot \mid I)$ gives $F^{-1}(\alpha \mid I; \gamma) - b(\alpha; I) = \lambda^{-1}[\alpha/g(b(\alpha; I) \mid I)]$ for $\alpha \in [0, 1]$. This equation does not contain enough information to identify both γ and $\lambda(\cdot)$. In particular, Proposition 2 implies that parameterizing one or more quantiles of $F(\cdot \mid I)$ is not sufficient to identify $U(\cdot)$ non-parametrically.[11]

9. Proposition 1 also implies that any auction model with some wealth $w > 0$, as defined in footnote 6, is observationally equivalent to some CRRA/CARA model with zero wealth. Note that $\lambda(\cdot)$ is independent of w under a CARA specification as $\lambda(\cdot) = (\exp(a\cdot) - 1)/a$.

10. The following is an example with $I = 2$ of a CRRA structure that is not observationally equivalent to any risk-neutral structure. Let $G(b) = kb$ for $b \in [0, 1/2]$ and $G(b) = \left[\frac{x_2 - 1}{1 - x_1} \frac{b - x_1}{x_2 - b}\right]^{3/[8(x_2 - x_1)]}$ for $b \in [1/2, 1]$, where $x_1 < x_2$ are roots of $-8x^2 + 11x - 2 = 0$ and k such that $G(\cdot)$ is continuous at $b = 1/2$. This distribution belongs to $\mathcal{G}_R$ with $R = 1$. Because $\lambda(x) = x/(1 - c)$, $G(\cdot)$ can be rationalized by a CRRA structure, where $\xi(b, c, G) = b + (1 - c)G(b)/g(b)$ as soon as $2/5 < c < 1$. On the other hand, from Guerre, Perrigne and Vuong (2000), $G(\cdot)$ is rationalized by a risk-neutral structure if and only if $\xi(b, G) = b + G(b)/g(b)$ is strictly increasing. This function is $\xi(b, G) = 2b$ for $0 \leq b \leq 1/2$ and $\xi(b, G) = -\frac{8}{3}(b - \frac{1}{2})(b - \frac{5}{4}) + 1$ for $1/2 \leq b \leq 1$, which is not strictly increasing. Hence, there does not exist a risk-neutral structure that is observationally equivalent to the preceding CRRA structure.

11. If $F(\cdot \mid I)$ is known, this equation shows that $\lambda(\cdot)$ and hence $U(\cdot)$ are non-parametrically identified on $[0, \max_v (v - s(v))]$. This property is exploited in Lu and Perrigne (2008), who rely on ascending auction data to estimate $F(\cdot \mid I)$ and hence to identify $U(\cdot)$ non-parametrically from first-price sealed-bid auction data. Alternatively, if $F(\cdot \mid \cdot)$ does not depend on I, which is the case with exogenous participation, $[U, F]$ is identified non-parametrically as shown in Guerre, Perrigne and Vuong (2009).

The second identifying strategy is to parameterize the utility function as $U(\cdot;\theta)$ leading to the semiparametric model $\mathcal{U}(\Theta) \times \mathcal{F}_R$. We define the CARA model (with smoothness R) as the set of structures $[U, F] \in \mathcal{U}^{\text{CARA}} \times \mathcal{F}_R$. The CRRA model is similarly defined as $\mathcal{U}^{\text{CRRA}} \times \mathcal{F}_R$. The next proposition shows that parameterizing $U(\cdot)$ is not sufficient.

Proposition 3. *Let $I \geq 2$ and $R \geq 1$. Any structure $[U, F]$ in $\mathcal{U}^{\text{CARA}} \times \mathcal{F}_R$ or $\mathcal{U}^{\text{CRRA}} \times \mathcal{F}_R$ is not identified.*

It is useful to understand the source of non-identification by considering the CRRA model. Let $[U, F]$ be a CRRA structure and $G(\cdot \mid I)$ the corresponding bid distribution. Consider the alternative CRRA structure $[\tilde{U}, \tilde{F}]$ with $c < \tilde{c} < 1$ and $\tilde{F}(\cdot \mid I)$ the distribution of

$$\tilde{v} = b + \frac{1-\tilde{c}}{I-1}\frac{G(b \mid I)}{g(b \mid I)} = \frac{\tilde{c}-c}{1-c}b + \frac{1-\tilde{c}}{1-c}\left(b + \frac{1-c}{I-1}\frac{G(b \mid I)}{g(b \mid I)}\right),$$

where $b \sim G(\cdot \mid I)$. Because the above function is strictly increasing in b when $c < \tilde{c} < 1$, then $G(\cdot \mid I)$ can also be rationalized by $[\tilde{U}, \tilde{F}]$ by Guerre, Perrigne and Vuong (2009, Lemma 1). Hence, $[\tilde{U}, \tilde{F}]$ is observationally equivalent to $[U, F]$. This result contrasts with Donald and Paarsch (1996, Theorem 1), who obtain parametric identification of the CRRA model by restricting $\tilde{F}(\cdot \mid I)$ and $F(\cdot \mid I)$ to have the same known support. At $b = \overline{b}(I)$, the above equation indicates that the support of $\tilde{F}(\cdot \mid I)$ must shrink, *i.e.*, $\overline{\tilde{v}}(I) < \overline{v}(I)$, to compensate for the increase in the CRRA parameter $\tilde{c} > c$. More generally, all the quantiles of $\tilde{F}(\cdot \mid I)$ are smaller than the corresponding ones for $F(\cdot \mid I)$ as $\tilde{v}(\alpha; I) = [(\tilde{c}-c)/(1-c)]b(\alpha; I) + [(1-\tilde{c})/(1-c)]v(\alpha; I)$ with $b(\alpha; I) < v(\alpha; I)$ for $\alpha \in (0, 1]$. In other words, an increase in risk aversion can be compensated by a shrinkage in the private value distribution.

3. SEMIPARAMETRIC IDENTIFICATION

Auction data typically provide variations in observed characteristics $Z \in \mathcal{Z} \subset \boldsymbol{R}^d$ of auctioned objects and/or in the number of bidders $I \in \mathcal{I}$. The purpose of this section is to investigate whether such variations help to identify $[U, F]$ through parameterizations. Hereafter, the structure is now defined by $[U(\cdot), F(\cdot \mid Z, I)]$, which excludes (Z, I) from the bidders' vNM utility function. This is justified in the case studied here as bidders do not face uncertainty about the quality and equivalent monetary value of the auctioned object. Restricting $U(\cdot)$ to be independent of (Z, I) is, however, insufficient for identifying $[U, F]$ as Propositions 2 and 3 extend to this case.[12] Thus, parameterizing either $F(\cdot \mid Z, I)$ or $U(\cdot)$ does not identify necessarily the structure $[U, F]$. Additional identifying restrictions need to be imposed. In the first case, parameterizing $U(\cdot)$ would lead to a full parametric model $\mathcal{U}(\Theta) \times \mathcal{F}(\Gamma)$. In the second case, parameterizing one quantile may help toward identification, while still providing flexibility about $F(\cdot \mid Z, I)$. Hereafter, the support of $F(\cdot \mid z, I)$ is denoted $[\underline{v}(z, I), \overline{v}(z, I)]$, while $G(\cdot \mid z, I)$ is the corresponding equilibrium bid distribution defined on $[\underline{b}(z, I), \overline{b}(z, I)]$ with density $g(\cdot \mid z, I)$. The α-quantiles of $F(\cdot \mid z, I)$ and $G(\cdot \mid z, I)$ are denoted $v(\alpha; z, I)$ and $b(\alpha; z, I)$, respectively.

Our identifying assumption is as follows.

Assumption A1. *For $\mathcal{I}$ a finite subset of $\{2, 3, \ldots\}$, $R \geq 1$, and some $\alpha \in (0, 1]$,*

(i) $U(\cdot) = U(\cdot;\theta) \in \mathcal{U}_R$ *for every* $\theta \in \Theta \subset \boldsymbol{R}^p$,

12. Consider the following example. Let $F(v \mid z, I)$ be a Uniform distribution on $[0, (\gamma + I - 2)z]$ and $U(x) = x^{\gamma-1}$ with $\gamma \in (1, 2]$. This leads to a uniform bid distribution on $[0, (I-1)z]$ irrespective of γ.

(ii) $F(\cdot \mid \cdot,\cdot) \in \mathcal{F}_R(\mathcal{Z} \times \mathcal{I}) \equiv \{F(\cdot \mid \cdot,\cdot)\colon F(\cdot \mid z, I) \in \mathcal{F}_R, \forall (z, I) \in \mathcal{Z} \times \mathcal{I}\}$,

(iii) $v(\alpha; z, I) = v(\alpha; z, I, \gamma)$ *for all* $(z, I) \in \mathcal{Z} \times \mathcal{I}$ *and some* $\gamma \in \Gamma \subset \mathbb{R}^q$,

(iv) *the system of equations* $\lambda(v(\alpha; z, I, \gamma) - b(\alpha; z, I); \theta) = \alpha/[(I-1)g(b(\alpha; z, I) \mid z, I)]$ *for* $(z, I) \in \mathcal{Z} \times \mathcal{I}$ *has a unique solution in* $(\theta, \gamma) \in \Theta \times \Gamma$.

Note that $\alpha = 0$ is excluded as the lower bound $\underline{v}(z, I)$ is non-parametrically identified from the boundary condition $\underline{v}(z, I) = \underline{b}(z, I)$. Condition (i) requires that $U(\cdot)$ belongs to a parametric family of smooth utility functions such as CRRA and CARA.[13] Condition (ii) requires that $F(\cdot \mid z, I)$ satisfies some smoothness conditions for every $(z, I) \in \mathcal{Z} \times \mathcal{I}$.

Condition (iii) requires a parameterization of a single conditional quantile $v(\alpha; z, I)$. See Powell (1994). For instance, $v(\alpha; \cdot, \cdot, \gamma)$ can be chosen to be a constant or a polynomial, where γ is a vector of unknown coefficients. It is worth noting that parameterizing the utility function and the α-quantile of the distribution of private values are jointly needed. In particular, dropping either A1-(i) or A1-(iii) would lead to a non-identified model as noted above. An alternative identification strategy would be to parameterize several quantiles of $F(\cdot \mid \cdot,\cdot)$ while leaving $U(\cdot)$ unspecified. The following example shows that this strategy does not lead to non-parametric identification of $U(\cdot)$. Let $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R(\mathcal{Z} \times \mathcal{I})$, where A1-(iii) is satisfied for, *e.g.*, two quantiles α_1 and α_2. Define $[\tilde{U}, \tilde{F}]$ with

$$\tilde{U}(x) = \begin{cases} c_1[U(x/\delta)]^{\delta} & \text{for } 0 \le x < \delta^2, \\ c_2 U(x + \delta(1-\delta)) & \text{for } x \ge \delta^2, \end{cases}$$

where $0 < \delta < 1$, $c_1 = c_2[U(\delta)]^{1-\delta}$, and $c_2 = 1/U(1 + \delta(1-\delta))$ so that $\tilde{U}(\cdot)$ is continuously differentiable at $x = \delta^2$ and $\tilde{U}(1) = 1$.[14] Let $\tilde{F}(\cdot \mid z, I)$ be the distribution of

$$\tilde{\xi}(b; z, I) = b + \tilde{\lambda}^{-1}\left(\frac{1}{I-1}\frac{G(b \mid z, I)}{g(b \mid z, I)}\right),$$

where $b \sim G(\cdot \mid z, I)$. We first show that $G(\cdot \mid \cdot,\cdot)$ is also rationalized by $[\tilde{U}, \tilde{F}]$. From Guerre, Perrigne and Vuong (2009, Lemma 1), it suffices that $\tilde{\xi}'(\cdot; z, I) > 0$ for any $(z, I) \in \mathcal{Z} \times \mathcal{I}$. After some algebra, we obtain

$$\tilde{\xi}(b; z, I) = \begin{cases} (1-\delta)b + \delta\xi(b; z, I) & \text{if } G(b \mid z, I)/[(I-1)g(b \mid z, I)] \le \lambda(\delta), \\ \xi(b; z, I) - \delta(1-\delta) & \text{if } G(b \mid z, I)/[(I-1)g(b \mid z, I)] \ge \lambda(\delta). \end{cases}$$

Because $\xi'(\cdot; z, I)$ is strictly positive, $\tilde{\xi}'(\cdot; z, I)$ is strictly positive as required. Hence, $[\tilde{U}, \tilde{F}]$ rationalizes the bid distribution $G(\cdot \mid \cdot,\cdot)$. It remains to show that $\tilde{F}(\cdot \mid \cdot,\cdot)$ satisfies A1-(iii) for the two quantiles α_1 and α_2. For $j = 1, 2$, the α_j-quantile $\tilde{v}(\alpha_j; z, I)$ of $\tilde{F}(\cdot \mid z, I)$ satisfies $\tilde{v}(\alpha_j; z, I) = \tilde{\xi}(b(\alpha_j; z, I); z, I)$. Let $X(\alpha_j; z, I) = G(b(\alpha_j; z, I) \mid z, I)/[(I-1)g(b(\alpha_j; z, I) \mid z, I)] = \alpha_j/[(I-1)g(b(\alpha_j; z, I) \mid z, I)]$. Because $\lambda(\cdot)$ is strictly increasing with $\lambda(0) = 0$, there exists δ sufficiently small so that $0 < \lambda(\delta) < \inf_{(z,I)\in\mathcal{Z}\times\mathcal{I}} X(\alpha_j; z, I)$ for $j = 1, 2$. Thus, $\tilde{v}(\alpha_j; z, I) = \xi(b(\alpha_j; z, I); z, I) - \delta(1-\delta) = v(\alpha_j; z, I) - \delta(1-\delta) > 0$ for δ sufficiently small. Hence, for

13. If wealth w is unknown, then w is included in θ. See footnote 6.

14. Note that $\tilde{U}(0) = 0$ and $\tilde{U}(\cdot)$ has $R+2$ continuous derivatives on $(0, \delta) \cup (\delta, +\infty)$. Thus, $\tilde{U}(\cdot)$ would belong to $\mathcal{U}_R$ if $\tilde{U}(\cdot)$ has $R+2$ continuous derivatives at $x = \delta^2$. In fact, $\tilde{U}(\cdot)$ has only one continuous derivative at $x = \delta^2$. Hence, $\tilde{U}(\cdot)$ should be smoothed out in the neighbourhood of $x = \delta^2$ to be $R+2$ continuously differentiable on $(0, +\infty)$. This smoothing requirement can be omitted by choosing δ sufficiently small as done below.

$j=1,2$, $\tilde{v}(\alpha_j;z,I)$ satisfies A1-(iii) whenever $v(\alpha;z,I;\gamma)$ contains a constant term. Thus, the parameterization of two conditional quantiles of $F(\cdot\mid z,I)$ is not sufficient for identification.[15]

Condition (iv) is a high-level parametric identifying condition. It holds for a large class of structures $[U,F]$ satisfying A1-(i–iii). First, consider the case where I is exogenous, *i.e.*, $F(\cdot\mid Z,I)=F(\cdot\mid Z)$.

The key is to exploit variations in I to identify $U(\cdot)$. In particular, consider $I,\tilde{I}\geq 2$ and $I\neq\tilde{I}$. Because the quantiles of $F(\cdot\mid\cdot)$ are invariant in I, we obtain from equation (3) the so-called compatibility conditions

$$\lambda^{-1}(X(\alpha;z,I))-\lambda^{-1}(X(\alpha;z,\tilde{I}))=b(\alpha;z,\tilde{I})-b(\alpha;z,I),$$

for any $\alpha\in[0,1]$. Guerre, Perrigne and Vuong (2009) exploit these conditions to show that $U(\cdot)$ and $F(\cdot\mid\cdot)$ are non-parametrically identified on $[0,\sup_{(z,I)}\sup_{v\in[\underline{v}(z),\overline{v}(z)]}v-s(v;z,I)]$ and $\{(v,z)\colon v\in[\underline{v}(z),\overline{v}(z)],z\in\mathcal{Z}\}$, respectively. Consequently, under A1-(i) and A1-(iii), the parameters (θ,γ) are identified under the usual parametric conditions that θ and γ are uniquely determined by the knowledge of the functions $\lambda(\cdot;\theta)$ and $v(\alpha;\cdot,\gamma)$ on $[0,\sup_{(z,I)}\sup_{v\in[\underline{v}(z),\overline{v}(z)]}v-s(v;z,I)]$ and $\mathcal{Z}$, respectively. Hence, A1-(iv) is automatically satisfied and the class of identifiable $[U,F]$ contains all the smooth specifications.

Second, consider a CRRA model $U(x)=x^{\theta}$ with a polynomial specification for the α-quantile, *e.g.*, $v(\alpha;z,I,\gamma)=\gamma_0+\gamma_1 z+\gamma_2 I$. Because $\lambda(x;\theta)=x/\theta$, equation (3) gives

$$\gamma_0+\gamma_1 z+\gamma_2 I-\theta X(\alpha;z,I)=b(\alpha;z,I),$$

for all $(z,I)\in\mathcal{Z}\times\mathcal{I}$. Thus, A1-(iv) is satisfied if and only if this system of linear equations has a unique solution, *i.e.*, its determinant is not equal to zero, which can be tested from observables $(z,I,X(\alpha;z,I))$. As $X(\alpha;z,I)$ is likely to be non-linear in (z,I), A1-(iv) is satisfied for a large class of structures $[U,F]$. More generally, considering another parametric specification for $U(\cdot)$ would lead to a non-linear system of equations in (θ,γ) for which local identification conditions can be obtained through usual "rank" conditions. In particular, condition (iv) implies the "order" condition $\text{Card}(\mathcal{Z}\times\mathcal{I})\geq p+q$. Altogether, A1 may lead to overidentification. This issue will be further discussed in Section 5.[16]

The next proposition establishes the semiparametric identification of the first-price auction model with risk-averse bidders. It relies upon the key equation (3) evaluated at the α-quantile $b(\alpha;z,I)$, namely

$$g(b(\alpha;z,I)\mid z,I)=\frac{1}{I-1}\frac{\alpha}{\lambda(v(\alpha;z,I,\gamma)-b(\alpha;z,I);\theta)}, \tag{4}$$

for any $(z,I)\in\mathcal{Z}\times\mathcal{I}$, where $\lambda(\cdot;\theta)=U(\cdot;\theta)/U'(\cdot;\theta)$.

Proposition 4. *The semiparametric model defined as the set of structures $[U,F]$ satisfying A1 is identified.*

Proposition 4 provides a semiparametric identification result since $U(\cdot)$ is parametrically identified through θ while $F(\cdot\mid\cdot,\cdot)$ is non-parametrically identified subject to a parametric conditional quantile restriction.

15. The same argument applies for an infinity of α bounded away from zero so that $\inf_{\alpha}\inf_{(z,I)\in\mathcal{Z}\times\mathcal{I}}X(\alpha;z,I)>0$.

16. Alternatively, condition (iv) that delivers point identification could be dropped. In this case, we would introduce the identified set of parameters (θ,γ) leading to the same function $\lambda(v(\alpha;\cdot,\cdot,\gamma)-b(\alpha;\cdot,\cdot);\theta)$ in the spirit of Chernozhukov, Hong and Tamer (2007).

Next, we provide two results that are useful for the rest of the paper. The first establishes the existence, uniqueness, and smoothness of the equilibrium strategy $s(\cdot; z, I)$. The second provides the properties of the equilibrium bid distribution $G(\cdot \mid \cdot, \cdot)$.[17]

Theorem 1. *Let $\mathcal{I}$ be a finite subset of $\{2, 3, \ldots\}$ and $\mathcal{Z}$ be a rectangular compact of $\mathbb{R}^d$ with non-empty interior. If $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$, then there exists a unique (symmetric) equilibrium strategy $s(\cdot; \cdot, \cdot)$. Moreover, this strategy satisfies*

(i) *for every $(v, z, I) \in (\underline{v}(z, I), \overline{v}(z, I)] \times \mathcal{Z} \times \mathcal{I}$, $s(v; z, I) < v$ with $s(\underline{v}(z, I); z, I) = \underline{v}(z, I)$,*
(ii) *for every $(v, z, I) \in [\underline{v}(z, I), \overline{v}(z, I)] \times \mathcal{Z} \times \mathcal{I}$, $s'(v; z, I) > 0$ with $s'(\underline{v}(z, I); z, I) = (I - 1)\lambda'(0)/[(I - 1)\lambda'(0) + 1] < 1$,*
(iii) *for every $I \in \mathcal{I}$, $s(\cdot; \cdot, I)$ admits $R + 1$ continuous derivatives on $\{(v, z)\colon v \in [\underline{v}(z, I), \overline{v}(z, I)], z \in \mathcal{Z}\}$.*

Lemma 1. *Let $\mathcal{I}$ be a finite subset of $\{2, 3, \ldots\}$, $R \geq 1$, and $\mathcal{Z}$ be a rectangular compact of $\mathbb{R}^d$ with non-empty interior. For every $I \in \mathcal{I}$, the conditional distribution $G(\cdot \mid \cdot, I)$ corresponding to a structure $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$ satisfies the following:*

(i) *The upper boundary $\overline{b}(z, I)$ admits $R + 1$ continuous derivatives with respect to $z \in \mathcal{Z}$ and $\inf_{z \in \mathcal{Z}}(\overline{b}(z, I) - \underline{b}(z, I)) > 0$, where $\underline{b}(z, I) = \underline{v}(z, I)$,*
(ii) *$G(\cdot \mid \cdot, I)$ admits $R+1$ continuous partial derivatives on $S_I(G) \equiv \{(b, z); b \in [\underline{b}(z, I), \overline{b}(z, I)], z \in \mathcal{Z}\}$,*
(iii) *$g(b \mid z, I) > c_g > 0$ for all $(b, z) \in S_I(G)$,*
(iv) *$g(\cdot \mid \cdot, I)$ admits $R+1$ continuous partial derivatives on $S_I^u(G) \equiv \{(b, z); b \in (\underline{b}(z, I), \overline{b}(z, I)], z \in \mathcal{Z}\}$,*
(v) *$\lim_{b \downarrow \underline{b}(z,I)} \partial^r [G(b \mid z, I)/g(b \mid z, I)]/\partial b^r$ exists and is finite for $r = 1, \ldots, R + 1$ and $z \in \mathcal{Z}$.*

Theorem 1 and Lemma 1 are extensions of Guerre, Perrigne and Vuong (2009, Theorem 1, Lemma 1-(i)) to the case of exogenous variables Z. Their proofs can be found in the Supplementary Material available on the web site of the review.

4. OPTIMAL CONVERGENCE RATE

The purpose of this section is to derive the optimal (best) convergence rate that can be attained by any semiparametric estimator of the risk aversion parameter(s) θ. The optimal convergence rate for estimating the conditional density $f(\cdot \mid \cdot, \cdot)$ will follow from Guerre, Perrigne and Vuong (2000). To derive the optimal convergence rate for θ, we consider the semiparametric model where the upper boundary is parameterized. Because estimating the upper boundary ($\alpha = 1$) is faster than estimating any other quantile with $\alpha \in (0, 1)$, the optimal convergence rate for estimating θ cannot be faster when considering another α-quantile restriction. Thus, we focus on $\alpha = 1$, and without loss of generality, we consider a constant but unknown upper boundary, *i.e.*, $\overline{v}(z, I) = \overline{v}$. Such an assumption is frequently made in parametric estimation. Formally, we consider distributions $F(\cdot \mid \cdot, \cdot)$ satisfying the following definition.

Definition 1. For $R \geq 1$ and some unknown $\overline{v}$, $0 < \overline{v} < +\infty$, let $\mathcal{F}_R^* \equiv \mathcal{F}_R^*(\mathcal{Z} \times \mathcal{I})$ be the set of conditional distributions $F(\cdot \mid \cdot, \cdot)$ satisfying

17. To simplify, discrete exogenous variables are excluded. If not, our results continue to hold with trivial modifications.

(i) $\forall (z,I) \in \mathcal{Z} \times \mathcal{I}, \overline{v}(z,I) = \overline{v}$,
(ii) $\forall I \in \mathcal{I}$, $F(\cdot \mid \cdot, I)$ admits $R+1$ continuous derivatives on $\{(v,z): v \in [\underline{v}(z,I), \overline{v}], z \in \mathcal{Z}\}$,
(iii) for every $(v,z,I) \in [\underline{v}(z,I), \overline{v}] \times \mathcal{Z} \times \mathcal{I}$, $f(v \mid z, I) > 0$.

Conditions (ii) and (iii) are straightforward extensions of Definition 2-(ii, iii) in Guerre, Perrigne and Vuong (2009). We then consider the semiparametric model composed of structures $[U, F]$ satisfying the following assumption.

Assumption A2. *Let $\mathcal{I}$ be a finite subset of $\{2,3,\ldots\}$, $R \geq 1$, and $\mathcal{Z}$ be a rectangular compact of $\mathbb{R}^d$ with non-empty interior.*

(i) *In addition to A1-(i), $U(\cdot;\cdot)$ is $R+2$ continuously differentiable on $(0,+\infty) \times \Theta$,*
(ii) *$F(\cdot \mid \cdot, \cdot) \in \mathcal{F}_R^*$,*
(iii) *the system of equations $\lambda(\overline{v} - \overline{b}(z,I); \theta) = \alpha/[(I-1) g(\overline{b}(z,I) \mid z, I)]$ for $(z,I) \in \mathcal{Z} \times \mathcal{I}$ has a unique solution in $(\theta, \overline{v}) \in \Theta \times (0, +\infty)$.*

Conditions (i) and (ii) strengthen A1-(i–iii). Condition (iii) simply expresses A1-(iv) at the upper boundary under a constant restriction. Thus, equation (4) becomes

$$g(\overline{b}(z,I) \mid z, I) = \frac{1}{I-1} \frac{1}{\lambda(\overline{v} - \overline{b}(z,I); \theta)}, \tag{5}$$

for all $(z,I) \in \mathcal{Z} \times \mathcal{I}$. Let $\beta = (\theta, \overline{v})$.

It remains to specify the data-generating process. For the ℓth auction, one observes all the bids $B_{i\ell}, i = 1, \ldots, I_\ell$, the number of bidders $I_\ell \geq 2$ as well as the d-dimensional vector Z_ℓ characterizing object heterogeneity. This gives a total number $N = \sum_{\ell=1}^{L} I_\ell$ of bids, where L is the number of auctions. Following the game theoretical model of Section 2, we make the following assumption.[18]

Assumption A3.

(i) *The variables (Z_ℓ, I_ℓ), $\ell = 1, 2 \ldots$, are i.i.d. with support $\mathcal{Z} \times \mathcal{I}$ and with density $f_{ZI}(\cdot, \cdot)$ satisfying $0 < \inf_{(z,I) \in \mathcal{Z} \times \mathcal{I}} f_{ZI}(z,I) \leq \sup_{(z,I) \in \mathcal{Z} \times \mathcal{I}} f_{ZI}(z,I) < +\infty$,*
(ii) *for every ℓ, the private values $V_{i\ell}, i = 1, \ldots, I_\ell$, are i.i.d. as $F_0(\cdot \mid Z_\ell, I_\ell)$ conditionally upon (Z_ℓ, I_ℓ),*
(iii) *the semiparametric model is correctly specified, i.e., the true utility function $U_0(\cdot)$ and conditional distribution $F_0(\cdot \mid \cdot, \cdot)$ satisfy A2 for some $\theta_0 \in \Theta \subset \mathbb{R}^p$ and $0 < \overline{v}_0 < +\infty$.*

We invoke the minimax theory developed by, *e.g.*, Ibragimov and Has'minskii (1981) to establish the optimal rate at which $\beta = (\theta, \overline{v})$ can be estimated. We consider the norms $\|\beta\|_\infty = \max(|\theta_1|, \ldots, |\theta_p|, |\overline{v}|)$ and $\|f(\cdot \mid \cdot, \cdot)\|_\infty = \sup_{(v,z,I) \in \{(v,z,I):\, v \in [\underline{v}(z,I), \overline{v}], (z,I) \in \mathcal{Z} \times \mathcal{I}\}} |f(v \mid z, I)|$ and define the set of conditional densities

$$\mathcal{F}_R^*(M) = \left\{ f(\cdot \mid \cdot, \cdot) \in \mathcal{F}_R^*; \left\| \frac{\partial^R f(\cdot \mid \cdot, \cdot)}{\partial v^R} \right\|_\infty < M \right\},$$

18. Assumption A3-(i) can be weakened allowing the (Z_ℓ, I_ℓ)s not to be i.i.d. distributed as Theorem 3 is derived conditionally upon $(Z_1, I_1, \ldots, Z_\ell, I_\ell)$.

for $M > 0$. As usual in studies of convergence rates, one considers a neighbourhood of the true parameters (β_0, f_0) in order to exclude superefficiency, *i.e.*,

$$\mathcal{V}_\varepsilon(\beta_0, f_0) = \{(\beta, f) \in \Theta \times (0, +\infty) \times \mathcal{F}_R^*(M);\ \|\beta - \beta_0\|_\infty < \varepsilon,$$
$$\|(f(\cdot \mid \cdot, \cdot) - f_0(\cdot \mid \cdot, \cdot))\boldsymbol{I}(f(\cdot \mid \cdot, \cdot) f_0(\cdot \mid \cdot, \cdot) > 0)\|_\infty < \varepsilon\}$$

where the indicator function restricts comparison of conditional densities on the intersection of their supports. Let $\Pr_{\beta,f}$ be the joint distribution of the $V_{i\ell}$s and the (Z_ℓ, I_ℓ)s under (θ, f, f_{ZI}). The next theorem gives an upper bound for the optimal rate when estimating β_0. Let Θ^o denote the interior of Θ.

Theorem 2. *Under A2–A3, for any $\beta_0 \in \Theta^o \times (0, +\infty)$, any $f_0 \in \mathcal{F}_R^*(M)$, and any deterministic sequence ρ_N such that $\rho_N N^{-(R+1)/(2R+3)} \to +\infty$, there exists a diverging deterministic sequence $t_N \to +\infty$ such that*

$$\lim_{\varepsilon \to 0} \lim_{N \to +\infty} \inf_{\tilde{\beta}_N} \sup_{(\beta, f) \in \mathcal{V}_\varepsilon(\beta_0, f_0)} \Pr_{\beta,f}(\|\rho_N(\tilde{\beta}_N - \beta)\|_\infty \geq t_N) \geq 1/2,$$

where the infimum is taken over all possible estimators $\tilde{\beta}_N$ of β based upon $(B_{i\ell}, Z_\ell, I_\ell)$, $i = 1, \ldots, I_\ell$, $\ell = 1, \ldots, L$.

Theorem 2 reveals the non-parametric nature of the parameter β, which cannot be estimated at a faster rate than $N^{(R+1)/(2R+3)}$. More precisely, for any estimator $\tilde{\beta}_N$, Theorem 2 shows that $\rho_N(\tilde{\beta}_N - \beta)$ diverges with probability at least 1/2. Thus, ρ_N diverges too fast and β cannot be estimated at a rate faster than $N^{(R+1)/(2R+3)}$, which is smaller than the parametric rate $\sqrt{N}$. On the other hand, Theorem 3 in Section 5 will show that there exists an estimator $\hat{\beta}_N$ converging at the rate $N^{(R+1)/(2R+3)}$. Therefore, the optimal rate of convergence for estimating β_0 in the minimax sense is $N^{(R+1)/(2R+3)}$, *i.e.*, $N^{2/5}$ when $R = 1$, which is "independent" of the dimension d of the exogenous variables Z.

The optimal convergence rate $N^{(R+1)/(2R+3)}$ corresponds to the optimal rate for estimating a univariate density with $R + 1$ bounded derivatives. This seems surprising in view of equation (5), which suggests that β is as hard to estimate as the conditional density $g(\cdot \mid \cdot, \cdot)$, while the latter cannot be estimated faster than $N^{(R+1)/(2R+3+d)}$ from Stone (1982) given the $(R + 1)$ bounded derivatives of $g(\cdot \mid \cdot, I)$. The faster rate $N^{(R+1)/(2R+3)}$ can be explained by noting that equation (5) leads to the moment conditions

$$\mathrm{E}\left[\left\{g(\overline{b}(Z, I) \mid Z, I) - \frac{1}{I-1} \frac{1}{\lambda(\overline{v} - \overline{b}(z, I); \theta)}\right\} W(Z, I)\right] = 0, \tag{6}$$

for any vector function $W(\cdot)$. Integrating with respect to Z intuitively improves the rate of convergence by eliminating the Z dimension.

The main idea of the proof is to consider suitable perturbations of the true parameters (β_0, f_0). For instance, when $R = 1$, we consider the bid density $g_N(b \mid z, I) = g_0(b \mid z, I) + [m(z, I; \beta_N) - m(z, I; \beta_0)]\psi(\kappa\sqrt{\rho_N}(b - \overline{b}_0(z, I)))$, where ψ: $\boldsymbol{R}^- \to \boldsymbol{R}$ is compactly supported with $\psi(0) = 1$ and $\int \psi(x)dx = 0$, while

$$m(z, I; \beta) = \frac{1}{I-1} \frac{1}{\lambda(\overline{v} - \overline{b}_0(z, I); \theta)}, \tag{7}$$

$\kappa > 0$, and $\|\beta_N - \beta_0\|_\infty = O(1/\rho_N)$. Using Lemma 1 in Guerre, Perrigne and Vuong (2009) and Lemma 1 in Section 3, we first establish that each such density can be rationalized by an auction model with $(\beta_N, f_N(\cdot \mid \cdot, \cdot)) \in \mathcal{V}_\varepsilon(\beta_0, f_0)$ for ρ_N sufficiently large. We then show that the probability distributions of the $B_{i\ell}$s under $g_N(\cdot \mid \cdot, \cdot)$ and $g_0(\cdot \mid \cdot, \cdot)$ cannot be distinguished statistically from each other.

5. SEMIPARAMETRIC ESTIMATION

This section proposes a semiparametric procedure for estimating the parameter(s) θ in the utility function $U(\cdot;\theta)$ and the conditional latent private value density $f(\cdot \mid \cdot, \cdot)$. Because $f(\cdot \mid \cdot, \cdot)$ is not parameterized, the estimation problem is semiparametric. The first subsection presents the different steps of our semiparametric procedure, while the second subsection establishes its asymptotic properties. In the continuation of Section 4, we consider the semiparametric model with a constant upper boundary $\overline{v}$. Relaxing this specification and the issue of overidentification will also be discussed.

5.1. *A semiparametric procedure*

Our semiparametric procedure relies on the identifying relation obtained from equations (5) and (7)

$$g_0(\overline{b}_0(z, I) \mid z, I) = \frac{1}{I-1} \frac{1}{\lambda(\overline{v}_0 - \overline{b}_0(z, I); \theta_0)} = m(z, I; \beta_0), \quad \forall (z, I) \in \mathcal{Z} \times \mathcal{I}, \tag{8}$$

where the subscript 0 indicates the truth. If one knew the upper boundary $\overline{b}_0(\cdot, \cdot)$ and the density $g_0(\cdot \mid \cdot, \cdot)$, one could recover $\beta_0 = (\theta_0, \overline{v}_0)$ from equation (8) given the parametric form for $\lambda(\cdot; \cdot)$. From the knowledge of $G_0(\cdot \mid \cdot, \cdot)$ and θ_0, one could then recover bidders' private values v_i from equation (3) to estimate $f_0(\cdot \mid \cdot, \cdot)$. This suggests the following three-step procedure:

- Step 1: From observed bids, estimate non-parametrically $\overline{b}_0(\cdot, \cdot)$ and $g_0(\overline{b}_0(\cdot, \cdot) \mid \cdot, \cdot)$ at the observed values (Z_ℓ, I_ℓ),
- Step 2: Using equation (8), where $g_0(\overline{b}_0(Z_\ell, I_\ell) \mid Z_\ell, I_\ell)$ and $\overline{b}_0(Z_\ell, I_\ell)$ are replaced by their estimates from the first step, estimate $\beta_0 \equiv (\theta_0, \overline{v}_0)$ using NLLS by $\hat{\beta}_N \equiv (\hat{\theta}_N, \hat{v}_N)$,
- Step 3: Using equation (3), where $G_0(\cdot \mid \cdot, \cdot)$, $g_0(\cdot \mid \cdot, \cdot)$, and $\lambda(\cdot; \theta_0)$ are replaced by their non-parametric estimators and $\lambda(\cdot; \hat{\theta}_N)$, respectively, recover the pseudo private values $\hat{v}_i$ to estimate non-parametrically $f_0(\cdot \mid \cdot, \cdot)$.

5.1.1. Non-parametric boundary estimation. This step consists in estimating the upper boundary $\overline{b}_0(\cdot, \cdot)$ and the conditional density $g_0(\cdot \mid \cdot, \cdot)$ at the upper boundary. Fix $I \in \mathcal{I}$. By Lemma 1-(i), $\overline{b}_0(\cdot, I)$ is $R+1$ continuously differentiable on $\mathcal{Z}$. Following Korostelev and Tsybakov (1993), we introduce a partition of $\mathcal{Z}$ into bins increasing with N. The boundary estimator of $\overline{b}_0(z, I)$ for z in an arbitrary bin is obtained by minimizing the volume of the cylinder whose base is the bin and whose upper surface is defined by a polynomial of degree R in $z \in \boldsymbol{R}^d$ subject to the constraint that the observations are contained in such a cylinder. The optimal polynomial evaluated at z gives the boundary estimate $\hat{\overline{b}}_N(z, I)$. Under appropriate vanishing size Δ_N of the bins, namely $\Delta_N \propto (\log N/N)^{1/(R+1+d)}$, the resulting piecewise polynomial estimator converges to $\overline{b}_0(\cdot, I)$ uniformly on $\mathcal{Z}$ at the rate $(N/\log N)^{(R+1)/(R+1+d)}$, which is strictly faster that $\sqrt{N}$ whenever $R \geq d$. For instance, for $R = 1$ and $d = 1$, we partition $\mathcal{Z} = [\underline{z}, \overline{z}]$ into k_N bins $\{\mathcal{Z}_k; k = 1, \ldots, k_N\}$ of equal length $\Delta_N \propto (\log N/N)^{1/3}$. On each $\mathcal{Z}_k = [\underline{z}_k, \overline{z}_k)$, the estimate of the upper boundary is the straight line $\hat{a}_k + \hat{b}_k(z - \underline{z}_k)$, where $(\hat{a}_k, \hat{b}_k)$ is obtained by solving

$$\min_{\{(a_k,b_k):B_{i\ell}\le a_k+b_k(Z_\ell-\underline{z}_k),i=1,\dots,I_\ell=I,Z_\ell\in\mathcal{Z}_k\}}\int_{\underline{z}_k}^{\overline{z}_k} a_k+b_k(z-\underline{z}_k)dz=a_k\Delta_N+b_k\Delta_N^2/2.$$

This estimator converges at the uniform rate $(N/\log N)^{2/3}$.

Turning to the estimation of $g(\cdot\mid\cdot,\cdot)$, it is well known that standard kernel density estimators suffer from bias at the boundary. To minimize such boundary effects, we consider a one-sided kernel density estimator. Let $\Phi(\cdot)$ be a one-sided kernel with support $[-1,0]$ satisfying A4-(iii) given below. For every $\ell=1,\dots,L$ and $i=1,\dots,I_\ell$, define

$$Y_{i\ell}\equiv\frac{1}{h_N}\Phi\left(\frac{B_{i\ell}-\overline{b}_0(Z_\ell,I_\ell)}{h_N}\right),\quad \hat{Y}_{i\ell}\equiv\frac{1}{h_N}\Phi\left(\frac{B_{i\ell}-\hat{\overline{b}}_N(Z_\ell,I_\ell)}{h_N}\right)\tag{9}$$

where h_N is a bandwidth. Lemma B3 shows that $Y_{i\ell}$ is an asymptotically unbiased estimator of $g_0(\overline{b}_0(Z_\ell,I_\ell)\mid Z_\ell,I_\ell)$ given (Z_ℓ,I_ℓ) as h_N vanishes.[19] Because $\overline{b}_0(\cdot,\cdot)$ is unknown, we define $\hat{Y}_{i\ell}$ similarly to $Y_{i\ell}$, where $\overline{b}_0(\cdot,\cdot)$ is replaced by its estimator $\hat{\overline{b}}_N(\cdot,\cdot)$.

5.1.2. Semiparametric estimation of β_0. Let $\mathcal{F}_L$ be the σ-field generated by (Z_ℓ,I_ℓ), $\ell=1,\dots,L$. In view of equations (8) and (9), we consider

$$Y_{i\ell}=m(Z_\ell,I_\ell;\beta_0)+e_{i\ell}+\varepsilon_{i\ell},\tag{10}$$

where $e_{i\ell}\equiv\mathrm{E}[Y_{i\ell}\mid\mathcal{F}_L]-m(Z_\ell,I_\ell,\beta_0)$ and $\varepsilon_{i\ell}=Y_{i\ell}-\mathrm{E}[Y_{i\ell}\mid\mathcal{F}_L]$. Lemma B3 shows that the bias term $e_{i\ell}=O(h_N^{R+1})$, while the variance of the error term $\varepsilon_{i\ell}$ is an $O(1/h_N)$, namely

$$\mathrm{Var}[\varepsilon_{i\ell}\mid\mathcal{F}_L]=\frac{m(Z_\ell,I_\ell;\beta_0)+o(1)}{h_N}\int\Phi^2(x)dx.\tag{11}$$

Hence, the $Y_{i\ell}$s obey a regression model with a vanishing bias and an error variance diverging to infinity as h_N vanishes.

Equation (10) suggests to estimate β_0 by possibly weighted NLLS, *i.e.*, by minimizing

$$Q_N(\beta)=\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)[Y_{i\ell}-m(Z_\ell,I_\ell;\beta)]^2\tag{12}$$

with respect to $\beta=(\theta,\overline{v})\in\mathcal{B}_\delta$, where the $\omega(Z_\ell,I_\ell)$s are strictly positive weights and $\mathcal{B}_\delta=\{(\theta,\overline{v});\theta\in\Theta,\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\overline{b}_0(z,I)+\delta\le\overline{v}\le\overline{v}_{\sup}\}$ for some $\delta>0$ and $\overline{v}_{\sup}>0$.[20] The set $\mathcal{B}_\delta$ is introduced to bound $\lambda(\overline{v}-\overline{b}_0(z,I);\theta)$ away from 0. Because $\overline{b}_0(\cdot,\cdot)$ is unknown, it is replaced by its estimator. Thus, our estimator of β is $\hat{\beta}_N=\mathrm{Argmin}_{\beta\in\mathcal{B}_N}\hat{Q}_N(\beta)$, where

$$\hat{Q}_N(\beta)=\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)[\hat{Y}_{i\ell}-\hat{m}(Z_\ell,I_\ell;\beta)]^2,\tag{13}$$

$$\hat{m}(z,I;\beta)=\frac{1}{I-1}\frac{1}{\lambda(\overline{v}-\hat{\overline{b}}_N(z,I);\theta)},\quad \mathcal{B}_N=\left\{(\theta,\overline{v});\theta\in\Theta,\ \max_{1\le\ell\le L}\hat{\overline{b}}_N(Z_\ell,I_\ell)+\delta/2\le\overline{v}\le\overline{v}_{\sup}\right\}.$$

19. Note that $\overline{Y}_\ell=(1/I_\ell)\sum_{\ell=1}^{I_\ell}Y_{i\ell}$ has a kernel form with a one-sided kernel, though I_ℓ remains bounded and hence does not increase with N in our case.

20. Weighted NLLS can be viewed as generalized method of moments with instruments $W=\Omega[\partial m(\cdot,\cdot;\beta_0)/\partial\beta']$ and a weight matrix $([\partial m(\cdot,\cdot;\beta_0)'/\partial\beta]\Omega[\partial m(\cdot,\cdot;\beta_0)/\partial\beta'])^{-1}$, where Ω is the diagonal matrix of the weights $\omega(\cdot,\cdot)$.

5.1.3. Non-parametric estimation of $f(\cdot \mid \cdot, \cdot)$. This step is similar to the second step in Guerre, Perrigne and Vuong (2000) with the difference that $\lambda(\cdot;\theta_0)$ in equation (3) is estimated by $\lambda(\cdot;\hat{\theta}_N)$, while $\lambda(\cdot)$ was the identity in that paper. We first need an estimate of the ratio $G_0(\cdot \mid \cdot, \cdot)/g_0(\cdot \mid \cdot, \cdot)$ evaluated at $(B_{i\ell}, Z_\ell, I_\ell)$. For an arbitrary (b, z, I), the ratio $G_0(b \mid z, I)/g_0(b \mid z, I)$ is estimated by

$$\hat{\Lambda}(b, z, I) = \frac{h_g^{d+1}}{h_G^d} \frac{\sum_{\{\ell; I_\ell = I\}} \frac{1}{I_\ell} \sum_{i=1}^{I_\ell} \boldsymbol{I}(B_{i\ell} \le b) K_G\left(\frac{z - Z_\ell}{h_G}\right)}{\sum_{\{\ell; I_\ell = I\}} \frac{1}{I_\ell} \sum_{i=1}^{I_\ell} K_g\left(\frac{b - B_{i\ell}}{h_g}, \frac{z - Z_\ell}{h_g}\right)}$$

where $K_G(\cdot)$ and $K_g(\cdot)$ are kernels of order $R+1$ with bounded supports and h_G and h_g are bandwidths vanishing at the rates $(N/\log N)^{1/(2R+d+2)}$ and $(N/\log N)^{1/(2R+d+3)}$, respectively. The pseudo private values are then

$$\hat{V}_{i\ell} = B_{i\ell} + \lambda^{-1}\left(\frac{1}{I_\ell - 1}\hat{\Lambda}(B_{i\ell}, Z_\ell, I_\ell); \hat{\theta}_N\right),$$

if $(B_{i\ell}, Z_\ell) + \mathcal{S}(2h_G) \subset \hat{S}_{I_\ell}(G)$ and $(B_{i\ell}, Z_\ell) + \mathcal{S}(2h_g) \subset \hat{S}_{I_\ell}(G)$. Otherwise, we let $\hat{V}_{i\ell}$ be infinity, which corresponds to a trimming. The sets $\mathcal{S}(2h_G)$ and $\mathcal{S}(2h_g)$ are the supports of $K_G(\cdot/(2h_G))$ and $K_g(\cdot/(2h_g))$, respectively. The set $\hat{S}_I(G)$ is the estimated support of the conditional bid distribution $G_0(\cdot \mid \cdot, I)$. Specifically, $\hat{S}_I(G) = \{(b, z)\colon b \in [\hat{\underline{b}}_N(z, I), \hat{\overline{b}}_N(z, I)], z \in \mathcal{Z}\}$, where $\hat{\underline{b}}_N(\cdot, I)$ is defined similarly to $\hat{\overline{b}}_N(\cdot, I)$.

The N pseudo private values $\hat{V}_{i\ell}$ are then used in a standard kernel estimation of $f_0(\cdot \mid \cdot, \cdot)$. Namely, for an arbitrary (v, z, I), $f(v \mid z, I)$ is estimated by

$$\hat{f}(v \mid z, I) = \frac{h_Z^d}{h_f^{d+1}} \frac{\sum_{\ell; I_\ell = I} \frac{1}{I_\ell} \sum_{i=1}^{I_\ell} K_f\left(\frac{v - \hat{V}_{i\ell}}{h_f}, \frac{z - Z_\ell}{h_f}\right)}{\sum_{\ell; I_\ell = I} K_Z\left(\frac{z - Z_\ell}{h_Z}\right)},$$

where $K_f(\cdot)$ and $K_Z(\cdot)$ are kernels of order R and $R+1$ with bounded supports and h_f and h_Z are bandwidths vanishing at the rates $(N/\log N)^{1/(2R+d+3)}$ and $(L/\log L)^{1/(2R+d+2)}$, respectively. Because $\hat{\theta}_N$ converges at a faster rate, it follows from Guerre, Perrigne and Vuong (2000) that $\hat{f}(\cdot \mid \cdot)$ is uniformly consistent on compact subsets of its support at the rate $(N/\log N)^{R/(2R+d+3)}$, which is optimal for estimating $f_0(\cdot \mid \cdot)$ from observed bids.

5.2. *Asymptotic properties*

We make the next assumptions on δ, $(\theta_0, \overline{v}_0)$, the weights $\omega(\cdot, \cdot)$, the kernel $\Phi(\cdot)$, the bandwidth h_N, and the rate of uniform convergence a_N^{-1} of the boundary estimator $\hat{\overline{b}}_N(\cdot, \cdot)$.

Assumption A4.

(i) *δ is such that $0 < \delta < \overline{v}_0 - \sup_{(z,I) \in \mathcal{Z} \times \mathcal{I}} \overline{b}_0(z, I)$. Moreover, $(\theta_0, \overline{v}_0)$ belongs to $\Theta^o \times (0, \overline{v}_{\sup})$ for some $\overline{v}_{\sup} < \infty$, where Θ is a compact of $\boldsymbol{R}^p$, and*

$$\mathrm{Span}_{(z,I) \in \mathcal{Z} \times \mathcal{I}} \left\{ \frac{\partial \lambda(\overline{v}_0 - \overline{b}_0(z, I); \theta_0)}{\partial \beta} \right\} = \boldsymbol{R}^{p+1}.$$

(ii) *The weight functions $\omega(\cdot, \cdot)$ are uniformly bounded away from zero and infinity, i.e., $\inf_{(z,I) \in \mathcal{Z} \times \mathcal{I}} \omega(z, I) > 0$ and $\sup_{(z,I) \in \mathcal{Z} \times \mathcal{I}} \omega(z, I) < \infty$.*

(iii) *The kernel* $\Phi(\cdot)$ *is continuously differentiable on* $\boldsymbol{R}_-$ *with support* $[-1,0]$ *and satisfies* $\int \Phi(x)dx = 1$ *and* $\int x^j \Phi(x)dx = 0$ *for* $j = 1, \dots, R$.
(iv) $h_N = o(1)$ *with* $Nh_N \to \infty$.
(v) $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}} |\hat{\overline{b}}_N(z,I) - \overline{b}_0(z,I)| = O_P(a_N)$ *with* $a_N = o(\min\{h_N^{R+2}, \sqrt{h_N/N}\})$.

Regarding the first part of A4-(i), recall that $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}} \overline{b}_0(z,I) < \overline{v}_0$ by Theorem 1-(i), Lemma 1-(i), and the compactness of $\mathcal{Z}\times\mathcal{I}$. The second part of A4-(i) is standard in parametric estimation and strengthens A1-(iv). It implies that $\overline{b}_0(z,I)$ must have at least $p+1$ different values. As shown in Lemma B7, combined with A4-(ii), it ensures that

$$A(\beta) \equiv \frac{1}{\mathrm{E[I]}}\mathrm{E}\left[I\omega(z,I)\frac{\partial m(z,I;\beta)}{\partial\beta}\cdot\frac{\partial m(z,I;\beta)}{\partial\beta'}\right] \tag{14}$$

$$B(\beta) \equiv \frac{1}{\mathrm{E[I]}}\mathrm{E}\left[I\omega^2(z,I)m(z,I;\beta)\frac{\partial m(z,I;\beta)}{\partial\beta}\cdot\frac{\partial m(z,I;\beta)}{\partial\beta'}\right] \tag{15}$$

are full rank matrices in a neighbourhood of β_0. Though our kernel $\Phi(\cdot)$ is one-sided, A4-(iii, iv) are standard in kernel estimation. Assumption A4-(v) requires that $\hat{\overline{b}}_N(\cdot,\cdot)$ converges faster than $\hat{\theta}_N$ (see Theorem 3-(i) for the latter) so that estimation of the boundary does not affect the asymptotic distribution of $\hat{\theta}_N$. For instance, when $R=1$ and $d=1$, we have $a_N = (\log N/N)^{2/3}$ from Korostelev and Tsybakov (1993). If h_N is exactly of order $N^{-1/5}$, which gives the optimal convergence rate of $\hat{\theta}_N$ by Theorems 2 and 3, then A4-(v) is satisfied. More generally, when $d\geq 1$ and h_N is exactly of the optimal order $N^{-1/(2R+3)}$, $R\geq d$ is sufficient for the convergence rate $a_N^{-1} = (N/\log N)^{(R+1)/(R+1+d)}$ of $\hat{\overline{b}}_N(\cdot,\cdot)$ to satisfy A4-(v).

Analogously to equations (14) and (15), we introduce the following $(p+1)$-square matrices:

$$A_N(\beta) = \sum_{\ell=1}^{L} I_\ell\omega(Z_\ell,I_\ell)\frac{\partial m(Z_\ell,I_\ell;\beta)}{\partial\beta}\cdot\frac{\partial m(Z_\ell,I_\ell;\beta)}{\partial\beta'}, \tag{16}$$

$$B_N(\beta) = \sum_{\ell=1}^{L} I_\ell\omega^2(Z_\ell,I_\ell)m(Z_\ell,I_\ell;\beta)\frac{\partial m(Z_\ell,I_\ell;\beta)}{\partial\beta}\cdot\frac{\partial m(Z_\ell,I_\ell;\beta)}{\partial\beta'}, \tag{17}$$

which, when normalized by N, are consistent estimators of $A(\beta)$ and $B(\beta)$ as shown in Lemma B8. Since $m(\cdot,\cdot;\beta)$ is unknown, let $\hat{A}_N(\beta)$ and $\hat{B}_N(\beta)$ be defined as $A_N(\beta)$ and $B_N(\beta)$ with $m(\cdot,\cdot;\beta)$ replaced by $\hat{m}(\cdot,\cdot;\beta)$. Moreover, let

$$\mathbf{b}(\beta,g_0) = \frac{\int x^{R+1}\Phi(x)dx}{(R+1)!}\frac{1}{\mathrm{E}[I]}\mathrm{E}\left[I\omega(Z,I)\frac{\partial^{R+1}g_0(\overline{b}_0(Z,I)\mid Z,I)}{\partial b^{R+1}}\frac{\partial m(Z,I;\beta)}{\partial\beta}\right], \tag{18}$$

which gives the asymptotic bias of our estimator.

The next result establishes the consistency and asymptotic normality of $\hat{\beta}_N$. It also provides its rate of convergence and an estimator of its asymptotic variance.

Theorem 3. *Under A2–A4,*

(i) $\hat{\beta}_N \xrightarrow{P} \beta_0$ *with* $\hat{\beta}_N - \beta_0 = O_P(h_N^{R+1} + 1/\sqrt{Nh_N})$, *so the best rate of convergence of* $\hat{\beta}_N$ *is* $N^{(R+1)/(2R+3)}$, *which is achieved when* h_N *is of exact order* $N^{-1/(2R+3)}$.
(ii) *If* $\lim_{N\to\infty}\sqrt{Nh_N}h_N^{R+1} = \infty$, *then* $(1/h_N^{R+1})(\hat{\beta}_N - \beta_0) \xrightarrow{P} A(\beta_0)^{-1}\mathbf{b}(\beta_0,g_0)$.

(iii) *If* $\lim_{N\to\infty}\sqrt{Nh_N}h_N^{R+1}=c\geq 0$, *then*

$$\sqrt{Nh_N}(\hat{\beta}_N-\beta_0)\xrightarrow{d}\mathcal{N}\left(cA(\beta_0)^{-1}\mathbf{b}(\beta_0,g_0),A(\beta_0)^{-1}B(\beta_0)A(\beta_0)^{-1}\int\Phi^2(x)dx\right).$$

Moreover, consistent estimators of $A(\beta_0)$ *and* $B(\beta_0)$ *are* $N^{-1}\hat{A}_N(\hat{\beta}_N)$ *and* $N^{-1}\hat{B}_N(\hat{\beta}_N)$.

The proof of Theorem 3-(i) is complicated by the divergence of the error variance (11) in the non-linear model (10). In particular, omitting the estimation of the upper boundary $\overline{b}(\cdot,\cdot)$, which has no effect by A4-(v), $(1/N)Q_N(\beta)=O_{\mathrm{P}}(1/h_N)$ because of the diverging variance. Hence, $(1/N)Q_N(\beta)$ does not have a finite limit. This would lead to consider $h_NQ_N(\beta)/N$, but its limit is a constant. To overcome this difficulty, we show that $(Q_N(\beta)-Q_N(\beta_0)-\overline{Q}_N(\beta))/N$ vanishes asymptotically, where

$$\overline{Q}_N(\beta)=\sum_{\ell=1}^{L}I_\ell\omega(Z_\ell,I_\ell)[m(Z_\ell,I_\ell;\beta)-m(Z_\ell,I_\ell;\beta_0)]^2. \tag{19}$$

Consistency of $\hat{\beta}$ can then be established by standard arguments using the objective function $\overline{Q}_N(\beta)$ (see, *e.g.*, White, 1994).

Theorem 3-(ii, iii) gives the asymptotic distribution of $\hat{\beta}_N-\beta_0$ and its rate of convergence. In particular, our proof shows that $\hat{\beta}_N-\beta_0$ is approximately distributed as

$$h_N^{R+1}A^{-1}(\beta_0)\mathbf{b}(\beta_0,g_0)+\frac{1}{\sqrt{Nh_N}}A^{-1}(\beta_0)\,\mathcal{N}\left(0,B(\beta_0)\int\Phi^2(x)dx\right).$$

This expansion corresponds to the usual bias/variance decomposition of non-parametric estimators. When $Nh_N^{2R+3}\to 0$, the leading term is the second term, so that

$$\sqrt{Nh_N}(\hat{\beta}_N-\beta_0)\xrightarrow{d}\mathcal{N}\left(0,A(\beta_0)^{-1}B(\beta_0)A(\beta_0)^{-1}\int\Phi^2(x)dx\right).$$

When $Nh_N^{2R+3}\to\infty$, the leading term is the first term, *i.e.* the bias. Thus, the best convergence rate of $\hat{\beta}_N$ is achieved when the variance and the bias are of the same order, *i.e.* when h_N is exactly of order $N^{-1/(2R+3)}$, in which case $\hat{\beta}_N-\beta_0=O_P(N^{-(R+1)/(2R+3)})$.[21]

The best convergence rate of $\hat{\beta}_N$ is the optimal rate $N^{(R+1)/(2R+3)}$ established in Theorem 2. Moreover, this rate is independent of the dimension d of Z. Hence, our estimator does not suffer from the curse of dimensionality encountered in non-parametric estimation though it is slower than the parametric rate $\sqrt{N}$. The latter can be explained by the diverging variance of the error term in equation (11). Our proof shows that the average gradient $(1/N)\partial\hat{Q}_N(\beta_0)/\partial\beta=O_P(h^{R+1}+1/\sqrt{Nh_N})$, which is different from the usual $O_P(1/\sqrt{N})$. Alternatively, using the moment conditions (6) (see footnote 20), our semiparametric estimator does not satisfy Assumptions (iii)–(iv) of Theorem 8.1 in Newey and McFadden (1994). In particular, neglecting the estimation of $\overline{b}(\cdot,\cdot)$, which has no affect by A4-(v), these assumptions require that $\sqrt{L}\{\int[\hat{g}(\overline{b}_0(z,I)\mid z,I)-g_0(\overline{b}_0(z,I)\mid z,I)]W(z,I)dF_0(z,I)-(1/L)\sum_{\ell=1}^{L}\delta(Z_\ell,I_\ell)\}\xrightarrow{P}0$ for some influence function $\delta(\cdot,\cdot)$. This does not hold since $\sqrt{L}\int[\hat{g}(\overline{b}_0(z,I)\mid z,I)-g_0(\overline{b}_0(z,I)\mid z,I)]W(z,I)dF_0(z,I)$ is not an $O_P(1)$.

21. When h_N is optimally chosen, the estimator $\hat{\beta}_N$ is asymptotically biased. In a similar problem, Horowitz (1992) proposes a correction based on the estimation of the bias. See also Bierens (1987).

Theorem 3-(iii) is used to make inference on β_0 as it gives an estimate of the variance of $\hat{\beta}_N$, *i.e.*, $\left(\int \Phi^2(x)dx/h_N\right)\hat{A}_N^{-1}(\hat{\beta}_N)\hat{B}_N(\hat{\beta}_N)\hat{A}_N^{-1}(\hat{\beta}_N)$. Note that $\hat{\beta}_N$ depends on $\omega(\cdot,\cdot)$, which can be chosen optimally as in weighted NLLS. From equation (11), the optimal weight function $\omega^*(\cdot,\cdot)$ is inversely proportional to the variance, *i.e.*, $\omega^*(\cdot,\cdot) = 1/m(\cdot,\cdot;\beta_0)$. This optimal weighted NLLS estimator $\hat{\beta}_N^*$ can be implemented by a two-stage procedure, in which $\omega^*(\cdot,\cdot)$ is estimated by $1/\hat{m}(\cdot,\cdot;\hat{\beta}_N)$, where $\hat{\beta}_N$ is obtained in the first step by ordinary NLLS. The estimate of the variance of $\hat{\beta}_N^*$ then reduces to $\left(\int \Phi^2(x)dx/h_N\right)\hat{A}_N^{-1}(\hat{\beta}_N^*)$. This is the best variance achievable in the regression model (10) with $e_{i\ell} = 0$.

5.3. *Discussion*

To simplify the presentation and derivation of the asymptotic properties, Sections 4, 5.1, and 5.2 have focused on the case where the upper boundary of the private value distribution is constant, *i.e.*, $\overline{v}(z,I) = \overline{v}$. As noted by a referee, this specification has a particular feature as equation (5) leads to $X(1;z,I) \equiv 1/[(I-1)g(\overline{b}(z,I) \mid z,I)] = \lambda(\overline{v} - \overline{b}(z,I))$ thereby showing that $\lambda^{-1}(\cdot)$ is identified non-parametrically on $[\inf_{z,I} X(1;z,I), \sup_{z,I} X(1;z,I)]$ up to location as $\overline{v}$ is unknown.[22] In general, one chooses a non-constant specification of the upper boundary with parameters $\gamma \in \boldsymbol{R}^q$ such as a polynomial to capture non-linearities. In this case, the non-parametric identification up to location of $\lambda(\cdot)$ is lost. Moreover, the shape of $\lambda(\cdot)$ now depends on how the upper boundary varies with (z,I). Nevertheless, the optimal rate for estimating $\beta = (\theta,\gamma)$ remains $N^{(R+1)/(2R+3)}$, which is independent of the dimension of Z. The estimation procedure of Section 5.1 with its asymptotic properties of Section 5.2 still apply. The estimated variance–covariance matrix for $\hat{\beta}_N$ can then be used to test the degree of the polynomial for $\overline{v}(z,I)$.

Alternatively, one may prefer to specify a quantile in $(0,1)$ instead of the upper boundary to mitigate possible outliers in bid data. Our estimation procedure can be extended to this case. The first step becomes straightforward as the estimation of a quantile is computationally less involved than the estimation of an upper boundary. See Bhattacharya and Gangopadhyay (1990), Chaudhuri (1991), and Hall, Wolff and Yao (1999) for estimation of quantiles. The estimator of the second step is weighted NLLS as defined in equations (12) and (13) where $Y_{i\ell}$ and $\hat{Y}_{i\ell}$ are in equation (9) with a two-sided kernel $\Phi(\cdot)$, while $\overline{b}_0(Z_\ell,I_\ell)$ and $\hat{\overline{b}}_N(Z_\ell,I_\ell)$ are replaced by $b_0(\alpha;Z_\ell,I_\ell)$ and $\hat{b}_N(\alpha;Z_\ell,I_\ell)$, respectively. The third step remains the same. The difficulty lies in the derivation of asymptotic properties. Assumption A4-(v) is crucial to eliminate the estimation effect of the first step. This assumption can be satisfied because the estimation of the upper boundary $\overline{b}_0(z,I)$ converges sufficiently fast, namely at the uniform rate $(N/\log N)^{(R+1)/(R+1+d)}$. On the other hand, estimation of the quantile $b_0(\alpha;z,I)$ converges much slower, namely at the uniform rate $(N/\log N)^{(R+1)/(2R+2+d)}$. When A4-(v) is not satisfied, the first step introduces some non-negligible estimation error in the second step, which modifies the rate of convergence and the corresponding asymptotic covariance matrix. Our conjecture is that the optimal rate for estimating β under a quantile parametric specification remains $N^{(R+1)/(2R+3)}$, which is independent of d, while our estimator will achieve this rate. Such results could be obtained by using recent results on Bahadur representations of non-parametric conditional estimators to approximate the estimation error in $\hat{Y}_{i\ell}$. See Kong, Linton and Xia (2008) and Guerre and Sabbah (2008).

We now turn to the issue of overidentification. As noted by Florens, Marimoutou and Pequin-Feissolle (2007, Section 16.5.3), the concept of overidentification requires a reference model.

22. This remark applies to any other constant quantile specification $\alpha \in (0,1)$ leading to the non-parametric identification of $\lambda^{-1}(\cdot)$ up to location on $[\inf_{z,I} X(\alpha;z,I), \sup_{z,I} X(\alpha;z,I)]$. This also suggests that the model is overidentified and underlies the overidentification test discussed below.

Hereafter, the reference model is the semiparametric model $\{[U, F] \in \mathcal{U}_R \times \mathcal{F}_R(\mathcal{Z} \times \mathcal{I}); F(\cdot \mid \cdot, \cdot)$ satisfies A1-(iii)$\}$, *i.e.*, $U(\cdot)$ is left unspecified while the α-quantile of $F(\cdot \mid \cdot, \cdot)$ is parameterized, $\alpha \in (0, 1)$. As shown earlier, this model is not identified and leads to the condition

$$g(b(\alpha; z, I) \mid z, I) = \frac{1}{I-1} \frac{\alpha}{\lambda(v(\alpha; z, I, \gamma_0) - b(\alpha; z, I))},$$

where γ_0 is the true value. If γ_0 was known, $\lambda(\cdot)$ would be non-parametrically identified on $[\inf_{(z,I)} v(\alpha; z, I, \gamma_0) - b(\alpha; z, I), \sup_{(z,I)} v(\alpha; z, I, \gamma_0) - b(\alpha; z, I)]$. Provided the specification of the α-quantile is correct, which can be achieved by choosing a polynomial of sufficiently high degree to capture possible non-linearities, this suggests the following overidentification test, more precisely a specification test of $\lambda(\cdot; \theta)$. As indicated above with $\alpha \in (0, 1)$, weighted NLLS is used to estimate the parameters (θ, γ) giving

$$\hat{Y}_{i\ell} = \frac{\alpha}{(I_\ell - 1)\lambda[v(\alpha; Z_\ell, I_\ell, \hat{\gamma}) - \hat{b}(\alpha; Z_\ell, I_\ell); \hat{\theta}]} + \hat{u}_{i\ell} \equiv \hat{m}(Z_\ell, I_\ell; \hat{\beta}) + \hat{u}_{i\ell}. \tag{20}$$

Replacing γ_0 by $\hat{\gamma}$, one could then estimate non-parametrically $\lambda(\cdot)$ giving

$$\hat{Y}_{i\ell} = \frac{\alpha}{(I_\ell - 1)\tilde{\lambda}[v(\alpha; Z_\ell, I_\ell, \hat{\gamma}) - \hat{b}(\alpha; Z_\ell, I_\ell)]} + \tilde{u}_{i\ell} \equiv \tilde{m}(Z_\ell, I_\ell) + \tilde{u}_{i\ell}.$$

The overidentification test is based on the sum of squares $\sum_\ell^L [\hat{m}(Z_\ell, I_\ell; \hat{\beta}) - \tilde{m}(Z_\ell, I_\ell)]^2 w(Z_\ell, I_\ell)$, where the $w(Z_\ell, I_\ell)$s are some weights.[23] Lastly, the regression setting (20) also allows to select among competing parametric specifications of $\lambda(\cdot)$. This can be done by comparing the sum of squared errors (SSE) $\sum_{i=1}^{I_\ell} \sum_{\ell=1}^{L} \hat{u}_{i\ell}^2$ across specifications. Section 7 illustrates this selection procedure and the importance of the α-quantile specification.

6. EXTENSIONS

A number of extensions are useful in practice such as a binding reserve price, affiliated private values, and asymmetry among bidders. Guerre, Perrigne and Vuong (2009, Propositions 4–7) show that these models are not identified non-parametrically. In this section, we thus seek parametric restrictions to identify semiparametrically these models.

6.1. *Reserve price*

An announced binding reserve price $p_0 \in (\underline{v}(I), \overline{v}(I))$ acts as a screening device to bidders' participation.[24] Let $G^*(\cdot \mid I)$ be the observed truncated bid distribution and I^* the observed number of bidders with $I^* \leq I$. From the boundary condition $s(p_0) = p_0$, we have $G^*(b^* \mid I) = (F(v \mid I) - F(p_0 \mid I))/(1 - F(p_0 \mid I))$ with $b^* = s(v)$. Similarly to equation (3), equation (1) implies that

$$v_i = b_i^* + \lambda^{-1}\left(\frac{1}{I-1} \frac{G^*(b_i^* \mid I)}{g^*(b_i^* \mid I)} + \frac{1}{I-1} \frac{1}{g^*(b_i^* \mid I)} \frac{F(p_0 \mid I)}{1 - F(p_0 \mid I)}\right) \equiv \xi(b_i^*, G^*, I, F(p_0 \mid I)), \tag{21}$$

for $i = 1, \ldots, I^*$.

23. See, *e.g.*, Hart (1997) and the extensive literature on testing non-parametric versus parametric regression fits therein. Technical complications arise here from that quantile estimation has an effect on $\hat{Y}_{i\ell}$ as noted earlier and that $\hat{\theta}$ is not $\sqrt{N}$ consistent. The study of such a test is left for future research.

24. The case of a random or secret reserve price is treated in Perrigne (2003).

It can be shown that any smooth $G^*(\cdot \mid I)$ can be rationalized by some CRRA or CARA structure with $F(\cdot \mid I) \in \mathcal{F}_R$. Moreover, parameterizing the utility function is not sufficient to achieve identification. The proofs of these results follow those of Propositions 1 and 3, where $G(\cdot \mid I)/g(\cdot \mid I)$ is replaced by $G^*(\cdot \mid I)/g^*(\cdot \mid I) + F(p_0 \mid I)/[(1-F(p_0 \mid I))g^*(\cdot \mid I)]$ in view of equations (3) and (21).[25] As before, an increase in risk aversion can be compensated by lower quantiles of the private value distribution suggesting the need for a restriction on a single quantile to identify the model. With auction characteristics, the key identifying equation becomes

$$g^*(b(\alpha; z, I) \mid z, I) = \frac{1}{I-1} \frac{\alpha + \frac{F(p_0 \mid z, I)}{1-F(p_0 \mid z, I)}}{\lambda(v(\alpha; z, I, \gamma) - b(\alpha; z, I); \theta)}. \tag{22}$$

If $F(p_0 \mid Z, I)$ and I are known, then θ is semiparametrically identified under A1.

Estimation is performed by adapting the procedure in Section 5.1 following Guerre, Perrigne and Vuong (2000, Section 4.2). In particular, the first step requires estimators of I and $F(p_0 \mid z, I)$. A natural estimator for I is $\hat{I} = \max_\ell I^*_\ell$, while an estimator for $F(p_0 \mid z, I)$ is obtained from $\mathrm{E}(I^* \mid z) = I[1 - F(p_0 \mid z, I)]$. In the second step, the weighted NLLS is based on equation (22), where I and $F(p_0 \mid z, I)$ are replaced by their estimates. The convergence rate is optimal and as before, namely $N^{(R+1)/(2R+3)}$.

6.2. *Affiliated private values*

The private values $(v_1, \dots, v_I)$ are distributed as $\mathbf{F}(\cdot, \dots, \cdot \mid I)$, which is exchangeable and affiliated. Bidder i's expected profit is $U(v_i - b_i)G_{B_i|b_i}(b_i \mid b_i)$, where $B_i = \max_{j \neq i} b_j$, $b_i = s(v_i)$ and $G_{B_i|b_i}(b_i \mid b_i)$ is the probability that $b_i \geq B_i$ conditional on b_i. By symmetry, $G_{B_i|b_i}(\cdot \mid \cdot) = G_{B|b}(\cdot \mid \cdot)$, for $i = 1, \dots, I$, while equation (3) becomes

$$v_i = b_i + \lambda^{-1}\left(\frac{G_{B|b}(b_i \mid b_i)}{g_{B|b}(b_i \mid b_i)}\right) \equiv \xi(b_i, U, \mathbf{G}). \tag{23}$$

Note that $G_{B|b}(\cdot \mid \cdot)/g_{B|b}(\cdot \mid \cdot) = G_{B \times b}(\cdot, \cdot)/g_{Bb}(\cdot, \cdot)$, where $G_{B \times b}(\cdot, \cdot) = \partial G_{Bb}(\cdot, \cdot)/\partial b$ and $g_{Bb}(\cdot, \cdot)$ is the joint density. Let $\mathcal{G}_R$ be the set of exchangeable and affiliated distributions $\mathbf{G}(\cdot, \dots, \cdot \mid I)$ with R continuously differentiable densities such that $G_{B \times b}(b, b)/g_{Bb}(b, b)$ is $R+1$ continuously differentiable in $b \in [\underline{b}(I), \overline{b}(I)]$ and strictly positive on $(\underline{b}(I), \overline{b}(I)]$.

It can be shown that any $\mathbf{G}(\cdot, \dots, \cdot \mid I) \in \mathcal{G}_R$ can be rationalized by some CRRA or CARA structure with $\mathbf{F}(\cdot, \dots, \cdot \mid I) \in \mathcal{F}_R$. Moreover, parameterizing the utility function is not sufficient to achieve identification. The proofs of these results follow those of Propositions 1 and 3, where $G(\cdot \mid I)/[(I-1)g(\cdot \mid I)]$ is replaced by $G_{B \times b}(\cdot, \cdot)/g_{Bb}(\cdot, \cdot)$ in view of equations (3) and (23). With auction characteristics and under parameterization of $U(\cdot)$, the key identifying equation becomes

$$g_{Bb}(b(\alpha; z, I), b(\alpha; z, I) \mid z, I) = \frac{G_{B \times b}(b(\alpha; z, I), b(\alpha; z, I) \mid z, I)}{\lambda(v(\alpha; z, I, \gamma) - b(\alpha, z, I); \theta)}, \tag{24}$$

where $b(\alpha; z, I)$ is the α-quantile of the marginal distribution $G(\cdot \mid z, I)$. Under A1, the model $\mathcal{U}(\Theta) \times \mathcal{F}_R$ is semiparametrically identified.

Estimation is performed by adapting the procedure in Section 5.1 following Li, Perrigne and Vuong (2002). In particular, $G_{B \times b}(b, b \mid z, I)$ is estimated as the product of an indicator function for the first argument and kernels for the second and third arguments, while $g_{Bb}(b, b, z, I)$ is

25. Proposition 2 still holds as the example of a non-identified semiparametric model $\mathcal{U}_R \times \mathcal{F}(\Gamma)$ can be adapted with a truncation on the bid distribution at the reserve price.

estimated using a standard kernel density estimator. In the second step, weighted NLLS is based on equation (24). Because the bid distribution and density include an additional dimension, the optimal rate for estimating θ will be slower than in Theorem 1.

6.3. *Asymmetry*

Asymmetry among bidders can arise from differences in private value distributions or utility functions. The latter known as heterogeneous preferences in the literature includes different attitudes towards risk.

6.3.1. Asymmetry in private values. The joint distribution $\mathbf{F}(\cdot, \dots, \cdot)$ is equal to $\prod_i F_i(\cdot \mid I)$ with each $F_i(\cdot \mid I)$ defined on $[\underline{v}(I), \overline{v}(I)]$. Let $\mathcal{F}_R^I$ be the set of such distributions. Because of the boundary conditions $s_i(\underline{v}(I)) = \underline{v}(I)$ and $s_i(\overline{v}(I)) = s_j(\overline{v}(I))$, bidder i's distribution $G_i(\cdot \mid I)$ is defined on $[\underline{b}(I), \overline{b}(I)]$ for all $i = 1, \dots, I$. Following Campo, Perrigne and Vuong (2003), we have

$$v_i = b_i + \lambda^{-1}\left(\frac{1}{H_i(b_i)}\right) \equiv \xi_i(b_i, U, \mathbf{G}), \quad \text{where } H_i(\cdot) = \sum_{j \neq i} \frac{g_j(\cdot \mid I)}{G_j(\cdot \mid I)}, \tag{25}$$

for $i = 1, \dots, I$. Let $\mathcal{G}_R^I$ be the set of distributions $\mathbf{G}(\cdot, \dots, \cdot \mid I)$ such that each marginal distribution $G_i(\cdot) \in \mathcal{G}_R$.

It can be shown that any $\mathbf{G}(\cdot, \dots, \cdot \mid I) \in \mathcal{G}_R^I$ can be rationalized by some CRRA or CARA structure with $\mathbf{F}(\cdot, \dots, \cdot \mid I) \in \mathcal{F}_R^I$. Moreover, parameterizing $U(\cdot)$ is not sufficient to achieve identification. The proofs of these results follow those of Propositions 1 and 3, where $G(\cdot \mid I)/[(I-1)g(\cdot \mid I)]$ is replaced by $1/H_i(\cdot)$ in view of equations (3) and (25). With auction characteristics and parameterization of $U(\cdot)$, the key identifying equation becomes

$$\sum_{j \neq i} \frac{g_j(b_i(\alpha; z, I) \mid z, I)}{G_j(b_i(\alpha; z, I) \mid z, I)} = \frac{1}{\lambda(v_i(\alpha; z, I, \gamma_i) - b_i(\alpha; z, I); \theta)}, \tag{26}$$

for $i = 1, \dots, I$, where $b_i(\alpha; z, I)$ is the α-quantile of the marginal distribution $G_i(\cdot \mid z, I)$. Under A1, the model $\mathcal{U}(\Theta) \times \mathcal{F}_R^I$ is semiparametrically identified.

Comparing equations (5) and (26) at $\alpha = 1$ shows that $\varepsilon_{i\ell}$ in equation (10) is correlated across i since $Y_{i\ell}$ is replaced by $\hat{H}_{i\ell} = \sum_{j \neq i} Y_{j\ell}$. Thus, a generalized NLLS estimator is appropriate. We assume that the same I bidders are in the L auctions. Let $\hat{H}_\ell = (\hat{H}_{1\ell}, \dots, \hat{H}_{I\ell})'$ and $M_\ell(\beta) = (m(Z_\ell; \beta_1), \dots, m(Z_\ell; \beta_I))'$ with $m(Z_\ell; \beta) = 1/\lambda(\overline{v} - \overline{b}(Z_\ell); \theta)$ and $\beta = (\overline{v}, \theta)'$. The objective function is $\sum_{\ell=1}^L [\hat{H}_\ell - M_\ell(\beta)]' \Omega^*(Z_\ell) [\hat{H}_\ell - M_\ell(\beta)]$. Following Lemma B3, the optimal weight matrix $\Omega^*(Z_\ell)$ is $(RD_\ell R)^{-1}$, where R is an $(I \times I)$ matrix of ones with zeros on the diagonal and $D_\ell = \operatorname{diag}[R^{-1} M_\ell(\beta)]$. The resulting two-step estimator of θ converges at the optimal rate, namely $N^{(R+1)/(2R+3)}$.

6.3.2. Asymmetry in preferences. We consider structures of the form $[U_1, \dots, U_I, F] \in \mathcal{U}_R^I \times \mathcal{F}_R$ with $\mathcal{U}_R^I = \bigotimes_{i=1}^I \mathcal{U}_R$. For $i = 1, \dots, I$, we obtain

$$v_i = b_i + \lambda_i^{-1}\left(\frac{1}{H_i(b_i)}\right) \equiv \xi_i(b_i, U_i, \mathbf{G}), \tag{27}$$

where $\lambda_i(\cdot) = U_i(\cdot)/U_i'(\cdot)$ and $H_i(\cdot) = \sum_{j \neq i} g_j(\cdot \mid I)/G_j(\cdot \mid I)$. The boundary conditions $s_1(\underline{v}(I)) = \cdots = s_I(\underline{v}(I)) = \underline{v}(I)$ and $s_1(\overline{v}(I)) = \cdots = s_I(\overline{v}(I))$ give a common support $[\underline{b}(I), \overline{b}(I)]$ for the bid distributions. The set $\mathcal{G}_R^I$ is defined as before.

Because the α-quantiles $(b_1(\alpha; I), \ldots, b_I(\alpha; I))$ all correspond to the same α-quantile $v(\alpha; I)$, equation (27) evaluated at the α-quantile for an arbitrary pair (i, j) gives

$$b_j(\alpha; I) + \lambda_j^{-1}\left(\frac{1}{H_j(b_j(\alpha; I))}\right) = b_i(\alpha; I) + \lambda_i^{-1}\left(\frac{1}{H_i(b_i(\alpha; I))}\right). \tag{28}$$

The compatibility conditions (28) reduce the set of bid distributions that can be rationalized relative to the symmetric case and can help in identification as illustrated next.[26]

Proposition 5. *The semiparametric model* $\otimes_{i=1}^{I}\mathcal{U}^{CRRA} \times \mathcal{F}_R$ *is identified.*

This result first noted in Campo (2005) contrasts with Proposition 3 as it does not require any conditional quantile restriction. Considering a different parametric specification for $U_i(\cdot), i = 1, \ldots, I$, introduces non-linearities in equation (28). As such, only local identification of $\theta = (\theta_1, \ldots, \theta_I)$ can be achieved by using the implicit function theorem.

Considering $I = 2$ to simplify, estimation exploits the compatibility conditions (28). Specifically, using bid data $\{b_{1\ell}, b_{2\ell}, Z_\ell; \ell = 1, \ldots, L\}$ and some quantiles $\alpha_1, \ldots, \alpha_{n_L}$ with n_L increasing with L, we estimate non-parametrically the quantiles $\hat{b}_j(\alpha_n; z_\ell)$ and the bid densities at those quantiles $\hat{g}_j(\hat{b}_j(\alpha_n; z_\ell) \mid z_\ell)$ for $n = 1, \ldots, n_L$ and $j = 1, 2$. We can then estimate (θ_1, θ_2) by minimizing the objective function

$$\sum_{n=1}^{n_L}\left[\hat{b}_2(\alpha_n; z_\ell) + \lambda^{-1}\left(\frac{\alpha_n}{\hat{g}_2(\hat{b}_2(\alpha_n; z_\ell))}; \theta_2\right) - \hat{b}_1(\alpha_n; z_\ell) - \lambda^{-1}\left(\frac{\alpha_n}{\hat{g}_1(\hat{b}_1(\alpha_n; z_\ell))}; \theta_2\right)\right]^2.$$

This estimator is a semiparametric version of the non-parametric estimator suggested in Guerre, Perrigne and Vuong (2009).

7. EMPIRICAL APPLICATION

This section illustrates the previous methodology on USFS timber auctions which have been used in several empirical studies. Hansen (1985) tests the revenue equivalence theorem from ascending and sealed-bid auctions. Using the former, Baldwin, Marshall and Richard (1997) study collusion, while Haile (2001) analyses bidding behaviour with resale opportunities. Athey and Levin (2001) study skewed bidding on species when payments are based on harvested value and find that diversification across species is consistent with bidders' risk aversion.[27] Athey, Levin and Seira (2004) study entry and bidding patterns in sealed-bid and ascending auctions with asymmetric bidders. Each of these papers focuses on a specific economic issue. None of the latter studies have measured the extent of risk aversion. The objective of our application is to assess bidders' risk aversion.

We focus on the first-price sealed-bid auctions in 1979 for the Western half of the U.S. (Regions 1 to 6). The data set contains 378 auctions involving a total of 1,400 bids and a set of variables characterizing each timber tract including the estimated volume in thousand board feet (mbf), the tract location, the reserve price, and the appraisal value per mbf. The latter is an estimated value of timber taking into account its quality and quantity. The data also provide the sealed bids in dollars and the bidders' identities. Table 1 gives some summary statistics. The auctioned tracts display important heterogeneity in quality and size. Though several variables

26. The non-parametric model is still not identified as shown in Guerre, Perrigne and Vuong (2009, Proposition 7). Moreover, parameterizing $F(\cdot \mid I)$ does not help as the compatibility conditions are not exploited.

27. Empirical results in Baldwin (1995) also suggest the existence of bidders' risk aversion.

TABLE 1
Summary statistics

Variable	Mean	Standard deviation
Bids ($)	202,564	494,178
Winning bids ($)	211,639	520,178
Appraisal value ($ per mbf)	57·07	45·41
Volume (mbf)	1,625	3,153
Number of bidders	3·72	1·81

can explain bids' variability, the tract appraisal value, denoted Z_ℓ hereafter, best captures the heterogeneity in both volume and quality across tracts. Thus, $d = 1$. Because of the important dispersion in volume, our empirical analysis considers the 300 auctions for which Z_ℓ is smaller than $300,000. As in the previous empirical studies, we consider a non-binding reserve price.

For a sensitivity analysis, we estimate the model for two quantiles corresponding to $\alpha = 1$ and $\alpha = 0{\cdot}5$, *i.e.*, upper bound and median, respectively. For each quantile, we consider four functional forms, namely constant, linear, quadratic, and cubic polynomials in Z_ℓ. Moreover, risk aversion is specified as either CRRA or CARA giving $U(x) = x^\theta$ and $U(x) = [1 - \exp(-\theta x)]/[1 - \exp(-\theta)]$, respectively.[28] We let $R = 1$. The first step consists in estimating nonparametrically the upper boundary $\overline{b}_0(Z_\ell, I_\ell)$ or the median $b_0(0{\cdot}5; Z_\ell, I_\ell)$ and the bid density $g_0(\cdot \mid Z_\ell, I_\ell)$ at these values for $\ell = 1, \ldots, L$.[29] For the density at the upper boundary, we use the one-sided kernel $\Phi(x) = (6x + 4)\boldsymbol{I}(-1 \le x \le 0)$ thereby satisfying A4-(iii). A standard triweight kernel is used everywhere else a kernel is needed. The second step consists in estimating the risk aversion parameter θ. The CRRA specification allows us to test for risk neutrality corresponding to $\theta = 1$. For $\alpha = 1$, the optimal weights $\omega^*(Z_\ell, I_\ell)$ are equal to $(I_\ell - 1)(\overline{v}(Z_\ell; \gamma_0) - \overline{b}_0(Z_\ell))$. The standard errors are computed using Theorem 3. For $\alpha = 0{\cdot}5$, standard errors are those provided in the third step without adjustment. See also the discussion in Section 5.3.

Tables 2 and 3 provide the estimated results with standard errors in parentheses.[30] The tables also provide a lack-of-fit measure through the ratio SSE over total sum of squares (TSS) in the NLLS estimation of equation (10). The constant quantile specification is clearly inadequate whether $\alpha = 1$ or $0{\cdot}5$ as adding Z_ℓ to control for heterogeneity dramatically reduces the SSE. For $\alpha = 1$, the cubic specification is retained for the CRRA utility function, while the quadratic specification is retained for the CARA utility function in view of the significance/insignificance of the coefficient $\hat{\gamma}_3$. Similarly, for $\alpha = 0{\cdot}5$ the quadratic specification is retained for both CRRA and CARA utility functions. For the retained CRRA specifications, we observe that $\hat{\theta}$ takes a similar value whether $\alpha = 1$ or $\alpha = 0{\cdot}5$, namely 0·7061 versus 0·7223 leading to a CRRA coefficient of 0·2939 and 0·2777, respectively. In particular, risk aversion is significant though weak for $\alpha = 1$. Using experimental auction data, Goeree, Holt and Palfrey (2002) and Bajari and Hortacsu (2005) find a larger value for relative risk aversion in the [0·50; 0·85] range. Similarly, the order of magnitude for the CARA coefficient is the same across the retained specifications. This robustness suggests that the choice of the quantile functional form is more important than

28. Results for a polynomial of degree 4 provided insignificant estimates for the coefficient associated with Z_ℓ^4 in the quantile specification. Similarly, a CRRA specification with common wealth w, namely $U(x) = (x + w)^\theta - w^\theta$, did not provide a significant estimate for the parameter w.

29. Since the data do not provide enough auctions for four and more bidders, we pool across the number of bidders in the estimation of the upper boundary and median.

30. Estimation was performed without imposing the restriction $\hat{v}(\alpha, Z_\ell, I_\ell; \hat{\gamma}) > \hat{b}(\alpha; Z_\ell, I_\ell)$. We observe a few violations for the CRRA specification when $\alpha = 1$ only.

TABLE 2
Estimation results for $\alpha = 1$

	Constant	Linear	Quadratic	Cubic
CRRA	$\hat{\theta} = 0.2864$	$\hat{\theta} = 0.6813$	$\hat{\theta} = 0.7783$	$\hat{\theta} = 0.7061$
	(0.0771)	(0.1393)	(0.2552)	(0.1517)
	$\hat{\gamma}_0 = 89.8849$	$\hat{\gamma}_0 = 10.6114$	$\hat{\gamma}_0 = 15.8472$	$\hat{\gamma}_0 = 15.8431$
	(0.4977)	(1.0467)	(3.0077)	(2.3706)
		$\hat{\gamma}_1 = 5.2266$	$\hat{\gamma}_1 = -5.5372$	$\hat{\gamma}_1 = -5.6692$
		(0.1386)	(2.0275)	(1.6059)
			$\hat{\gamma}_2 = 2.3064$	$\hat{\gamma}_2 = 2.5279$
			(0.3499)	(0.3208)
				$\hat{\gamma}_3 = -0.1290$
				(0.0146)
	SSE/TSS = 0.9736	SSE/TSS = 0.7333	SSE/TSS = 0.6425	SSE/TTS = 0.6140
CARA	$\hat{\theta} = 0.000002$	$\hat{\theta} = 0.00002$	$\hat{\theta} = 0.000001$	$\hat{\theta} = 0.00006$
	(0.0004)	(0.0002)	(0.00002)	(0.0002)
	$\hat{\gamma}_0 = 296.6368$	$\hat{\gamma}_0 = 12.7502$	$\hat{\gamma}_0 = 20.0384$	$\hat{\gamma}_0 = 19.9201$
	(0.4196)	(1.2093)	(1.1350)	(1.4084)
		$\hat{\gamma}_1 = 10.7390$	$\hat{\gamma}_1 = -9.9618$	$\hat{\gamma}_1 = -9.8521$
		(0.3739)	(1.0891)	(1.7714)
			$\hat{\gamma}_2 = 3.6621$	$\hat{\gamma}_2 = 3.6433$
			(0.2101)	(0.5903)
				$\hat{\gamma}_3 = 0.0000003$
				(0.0621)
	SSE/TSS = 0.9770	SSE/TSS = 0.6998	SSE/TSS = 0.5706	SSE/TSS = 0.5682

the choice of which quantile to use. In other words, as long as the parametric specification of the quantile is correct, which can be achieved by a sufficiently high-order polynomial, the researcher may choose any quantile to estimate risk aversion in the second step of the estimation procedure.

For $\alpha = 1$, the CARA utility function provides marginally a better fit over CRRA, while for $\alpha = 0.5$ both provide a similar fit. Following the discussion of Section 5.3 on overidentification, we can assess the correct specification of either utility function by comparing the non-parametric estimate $\tilde{\lambda}(\cdot)$ with the parametric estimate $\lambda(\cdot;\hat{\theta})$.[31] It is interesting to compare our results with those obtained by Lu and Perrigne (2008), where $\lambda(\cdot)$ was estimated non-parametrically from ascending auction data. Their estimated $\hat{\lambda}(\cdot)$ function can be approximated by a CRRA specification with $\hat{\theta} = 0.65$ or $\hat{c} = 0.35$, which is not far from the results we obtain here using first-price sealed-bid auctions only. Their CARA parameter estimate is also very small of the same order as here. In their paper, results suggest that a CRRA specification provides a better fit though the latter has a tendency to overestimate the utility function especially when the rent $v - b$ is small.

In the third step, we estimate the conditional density of private values, which can be used to simulate the optimal posted reserve price. Because a closed form for the equilibrium strategy can be derived for the CRRA model only, we focus on this case.[32] Figures 1 and 2 display the estimated conditional densities of private values at the sample mean $Z_\ell = \overline{Z} = 36{,}773$ and $I = 4$ across the four specifications of the quantiles $\alpha = 1$ and $\alpha = 0.5$. For $\alpha = 1$, the estimated

31. Graphs representing both estimated functions for $\alpha = 1$ or 0.5 with $U(\cdot)$ equal to the CRRA or CARA utility functions are available upon request from the authors.

32. Li, Perrigne and Vuong (2003) propose a method to estimate the optimal reserve price in an affiliated private value model using the observed bid distribution and density. Their proof relies on the closed form of the equilibrium strategy, which is not available for the CARA specification.

TABLE 3
Estimation results for $\alpha = 0{\cdot}5$

	Constant	Linear	Quadratic	Cubic
CRRA	$\hat{\theta} = 0{\cdot}5866$	$\hat{\theta} = 0{\cdot}6006$	$\hat{\theta} = 0{\cdot}7223$	$\hat{\theta} = 0{\cdot}5560$
	(0·1150)	(0·1502)	(0·3586)	(0·2788)
	$\hat{\gamma}_0 = 114{\cdot}9999$	$\hat{\gamma}_0 = 25{\cdot}0037$	$\hat{\gamma}_0 = 28{\cdot}6067$	$\hat{\gamma}_0 = 22{\cdot}1137$
	(1·0379)	(6·1497)	(12·7786)	(9·7024)
		$\hat{\gamma}_1 = 1{\cdot}3339$	$\hat{\gamma}_1 = 1{\cdot}4245$	$\hat{\gamma}_1 = 1{\cdot}4978$
		(0·0658)	(0·0899)	(0·1245)
			$\hat{\gamma}_2 = -0{\cdot}0016$	$\hat{\gamma}_2 = -0{\cdot}00346$
			(0·0004)	(0·0016)
				$\hat{\gamma}_3 = 0{\cdot}000007$
				(0·000006)
	SSE/TSS = 0·8585	SSE/TSS = 0·3261	SSE/TSS = 0·3208	SSE/TSS = 0·3206
CARA	$\hat{\theta} = 0{\cdot}0000005$	$\hat{\theta} = 0{\cdot}0000002$	$\hat{\theta} = 0{\cdot}000003$	$\hat{\theta} = 0{\cdot}000002$
	(0·002)	(0·007)	(0·018)	(0·0183)
	$\hat{\gamma}_0 = 249{\cdot}0001$	$\hat{\gamma}_0 = 40{\cdot}2831$	$\hat{\gamma}_0 = 38{\cdot}3498$	$\hat{\gamma}_0 = 38{\cdot}0018$
	(4·5300)	(3·7615)	(11·0000)	(11·8562)
		$\hat{\gamma}_1 = 1{\cdot}2898$	$\hat{\gamma}_1 = 1{\cdot}4662$	$\hat{\gamma}_1 = 1{\cdot}5103$
		(0·1230)	(0·2020)	(0·3131)
			$\hat{\gamma}_2 = -0{\cdot}0015$	$\hat{\gamma}_2 = -0{\cdot}0023$
			(0·0009)	(0·0037)
				$\hat{\gamma}_3 = 0{\cdot}000003$
				(0·00001)
	SSE/TSS = 0·7854	SSE/TSS = 0·3233	SSE/TSS = 0·3211	SSE/TSS = 0·3211

densities are similar with the exception of the constant specification. For $\alpha = 0{\cdot}5$, all estimated densities are very similar. This suggests that the parametric specification of the chosen quantile does not affect much the estimated private value density. Figure 3 compares the estimated conditional density for the risk-neutral case, *i.e.*, $\theta = 1$ with those obtained from the retained specifications (cubic and quadratic) for the upper quantile and the median. The estimated density for the risk-neutral model is slightly shifted to the right with a longer tail. This can be explained by the overestimation of private values when imposing $\theta = 1$. Moreover, the estimated densities for the CRRA model are similar confirming that the choice of the quantile does not matter much.

From a policy perspective, risk aversion implies that bidders bid more aggressively relative to risk neutrality as they shade less their private values. In particular, a CRRA model is equivalent to having more competition in the auctions. For instance, with $I = 4$, a relative risk aversion $c = 0{\cdot}29$ is equivalent of having five bidders in an auction with risk neutrality. Another interpretation is that bidders' rents decrease by $100c\%$. Measuring risk aversion is also important for policy recommendations. Though the optimal mechanism with risk-averse bidders involves complex transfers (see Maskin and Riley, 1984; Matthews, 1987), an optimal posted reserve price can generate more revenue for the seller. The optimal reserve price p_0^* is solution of $p_0^* = v_0 + [((1-c)/(1-cI))[F^{(I-1)c/(1-c)}(p_0^* \mid z, I) - F(p_0^* \mid z, I)]]/f(p_0^* \mid z, I)$, where v_0 is the seller's value of the auctioned object. We set z and v_0 at the average appraisal value and $I = 4$. The estimated optimal reserve prices for the CRRA model using the upper bound and the median are \$74,270 and \$77,611, respectively, which are very close. The estimated optimal reserve price for risk-neutral bidders is significantly larger at \$95,818. Because risk-averse bidders bid more aggressively, the precommitment effect need not be as important thereby reducing the level of the reserve price that generates the maximum profit for the seller.

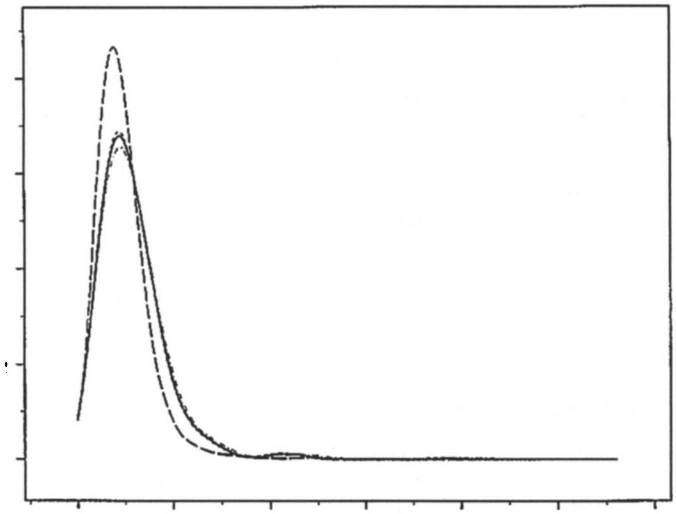

FIGURE 1
Conditional densities, $\alpha = 1$

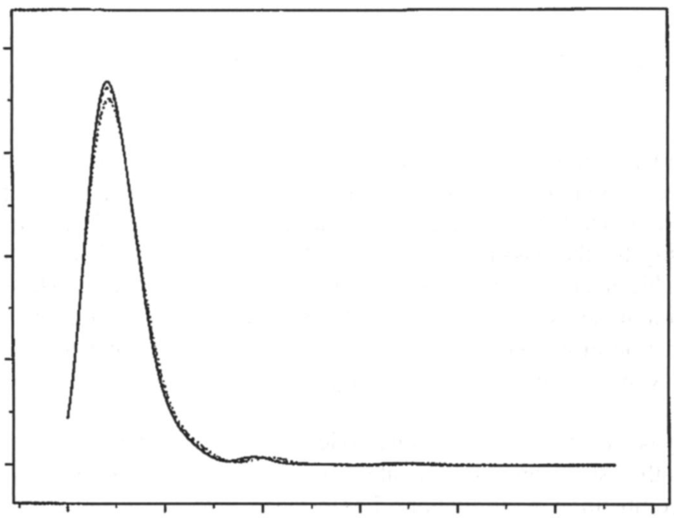

FIGURE 2
Conditional densities, $\alpha = 0{\cdot}5$

8. CONCLUSION

This paper extends the structural analysis of auction data to risk-averse bidders. In particular, our method allows one to estimate and test for bidders' risk aversion in first-price auctions within the private value paradigm. This represents an important extension as various experiments have shown that bidders are risk averse even when the financial stakes are small, suggesting that risk aversion is a natural component of agents' behaviour. On econometric grounds, the paper proposes a semiparametric method for estimating the structure of the model, namely bidders' risk

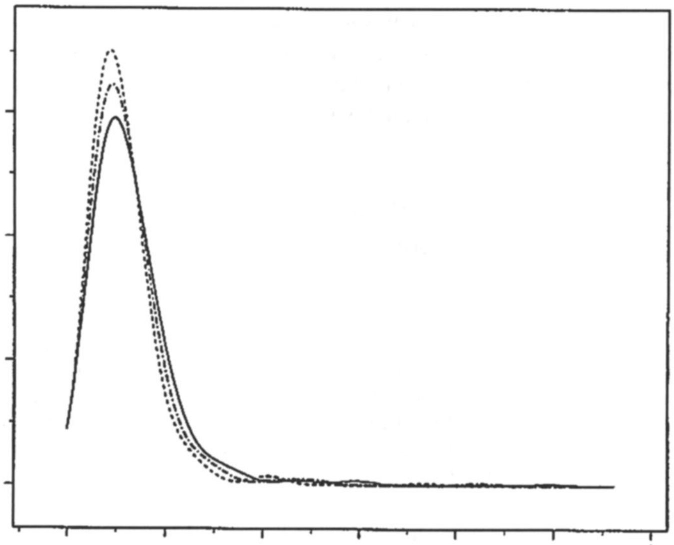

FIGURE 3
Conditional densities, risk aversion versus risk neutrality

aversion parameter(s) and the density of their private values. While previous papers have considered either fully parametric or non-parametric methods, this paper is the first one proposing a semiparametric estimator that arises naturally from the identification of the auction model.

We show that a model with CARA or CRRA is not identified, though either one can explain any observed bid distribution. We then propose some parametric restrictions to achieve semiparametric identification through a parameterization of the utility function and a conditional quantile restriction on the latent private value distribution. We show that our method extends to more general auction models such as affiliated private values and asymmetric bidders. Our semiparametric estimation method involves non-parametric quantile estimation, kernel estimators, and weighted NLLS. Our estimator converges at the optimal rate $N^{(R+1)/(2R+3)}$, which is smaller than $\sqrt{N}$. An important feature is that this rate is independent of the number of exogenous variables thereby avoiding the curse of dimensionality. An illustration of the method is proposed on USFS auction data showing bidders' risk aversion.

Given that little is known on the bidders' utility function, Guerre, Perrigne and Vuong (2009) exploit some exclusion restrictions to achieve non-parametric identification of the utility function. In particular, the latent distribution of private values is assumed to be independent of the number of bidders. The corresponding estimation method is yet to be studied. The non-parametric estimator of the utility function is likely to be not as friendly as the one studied here and will suffer from the curse of dimensionality. Thus, there is a trade-off. Either one is willing to make an exclusion restriction, in which case the utility function can be estimated non-parametrically or one restricts the utility function to be parametric, in which case one does not encounter the curse of dimensionality.

Several extensions can be entertained. The first interesting extension relates to the practice of random reserve prices, which may dominate posted reserve prices by accentuating overbidding under risk aversion. Perrigne (2003) studies random reserve prices and assesses empirically the gain for the seller of keeping the reserve price secret. Campo (2005) considers an auction model with heterogenous bidders for analysing construction procurements and finds that bidder's risk aversion decreases with experience. Stochastic private values due to *ex post* uncertainties as studied by Eso and White (2004) represent another interesting application. Perrigne and Vuong (2009) study precautionary bidding in construction procurements. More generally, estimating risk aversion is an important issue in the analysis of microeconomic data. See Cohen and Einav (2007) for a recent contribution in insurance.

APPENDIX A

Appendix A gathers the proofs of Propositions 1 and 3–5. A supplementary to Appendix A can be found in the Supplementary Material on the web site of the review.

Proof of Proposition 1.

(i) Consider a bid distribution $G(\cdot \mid I) \in \mathcal{G}_R$. We show that there exists a structure $[U, F]$, where $U(x) = x^{1-c}$, $0 \le c < 1$, and $F(\cdot \mid I) \in \mathcal{F}_R$, that rationalizes $G(\cdot \mid I)$. Note that $\lambda(x) = x/(1-c)$ with $\lambda(0) = 0$ and $\lambda'(\cdot) \ge 1$. From Guerre, Perrigne and Vuong (2009, Lemma 1), it suffices to show that there exists a value $c \in [0, 1)$ such that $\xi(b, c, G) = b + [(1-c)G(b \mid I)]/[(I-1)g(b \mid I)]$ has a strictly positive derivative on $[\underline{b}(I), \overline{b}(I)]$. Differentiating gives $[G(b \mid I)/g(b \mid I)]' > -(I-1)/(1-c)$ for all $b \in [\underline{b}(I), \overline{b}(I)]$, *i.e.*,

$$\inf_{b \in [\underline{b}, \overline{b}]} \left[\frac{G(b \mid I)}{g(b \mid I)} \right]' > -\frac{I-1}{1-c}. \tag{A.1}$$

The L.H.S. is finite because $G(\cdot \mid I)/g(\cdot \mid I)$ is $R+1$ continuously differentiable on $[\underline{b}(I), \overline{b}(I)]$. If $\inf_b [G(b \mid I)/g(b \mid I)]' \ge 0$, then any value $c \in (0, 1)$ satisfies inequality (A.1). If $\inf_b [G(b \mid I)/g(b \mid I)]' < 0$, inequality (A.1) can be written as $c > 1 - (I-1)/(-\inf_b [G(b \mid I)/g(b \mid I)]')$, where the R.H.S. is less than one. Thus, we can always find a $c \in (0, 1)$ satisfying inequality (A.1) and hence a CRRA model that rationalizes $G(\cdot \mid I)$.

(ii) The proof for the CARA case is similar. Consider $U(\cdot) \in \mathcal{U}_R^{\mathrm{CARA}}$. This gives $U(x) = (1-e^{-ax})/(1-e^{-a})$ with $a > 0$. Hence, $\lambda(x) = (e^{ax} - 1)/a$ and $\lambda^{-1}(x) = (1/a)\log(1+ax)$. The inverse bidding strategy is $\xi(b) = b + (1/a)\log\{1 + [aG(b \mid I)/(I-1)g(b \mid I)]\}$. We show that there exists $a > 0$ such that $\xi'(b) > 0$ on $[\underline{b}(I), \overline{b}(I)]$. Differentiating gives

$$a\frac{G(b \mid I)}{g(b \mid I)} > -\left[(I-1) + \left(\frac{G(b \mid I)}{g(b \mid I)} \right)' \right], \quad \forall b \in [\underline{b}(I), \overline{b}(I)].$$

Note that $\lim_{b \downarrow \underline{b}(I)} [G(b \mid I)/g(b \mid I)]' = \lim_{b \downarrow \underline{b}(I)} 1 - G(b \mid I)g'(b \mid I)/g^2(b \mid I) = 1$ because $R \ge 1$ and $g(b \mid I) > 0$. Hence, the preceding inequality holds at $\underline{b}(I)$ for any $a > 0$. Thus, it becomes

$$a > \sup_{b \in (\underline{b}(I), \overline{b}(I)]} -\frac{g(b \mid I)}{G(b \mid I)} \left[(I-1) + \left(\frac{G(b \mid I)}{g(b \mid I)} \right)' \right] \tag{A.2}$$

This is satisfied for an infinity of values for $a > 0$ provided the supremum is not $+\infty$. We know that $-[g(b \mid I)/G(b \mid I)]\{I - 1 + [G(b \mid I)/g(b \mid I)]'\}$ is R continuously differentiable and hence continuous on $(\underline{b}(I), \overline{b}(I)]$ because $R \ge 1$. Moreover, $\lim_{b \downarrow \underline{b}(I)} -[g(b \mid I)/G(b \mid I)]\{I - 1 + [G(b \mid I)/g(b \mid I)]'\} = -\infty$ because $g(b \mid I)/G(b \mid I)$ tends to $+\infty$ and $[G(b \mid I)/g(b \mid I)]'$ tends to 1. Thus, we can always find an $a > 0$ satisfying inequality (A.2) and hence a CARA model that rationalizes $G(\cdot \mid I)$. ||

Proof of Proposition 3. We first consider the CRRA case. Let $[U, F] \in \mathcal{U}^{\mathrm{CRRA}} \times \mathcal{F}_R$ with $c \in [0, 1)$ generating a bid distribution $G(\cdot \mid I) \in \mathcal{G}_R$. The proof of Proposition 1 shows that there exists a CRRA utility function $\tilde{U}(\cdot)$ with $0 \le c < \tilde{c} < 1$ and a distribution $\tilde{F}(\cdot \mid I) \in \mathcal{F}_R$ leading to the same $G(\cdot \mid I)$. Thus, the CRRA model is not identified. We can use a similar argument to show that the CARA model is not identified from the proof of Proposition 1. ||

Proof of Proposition 4. Let $[U, F]$ satisfy Assumption A1 with parameters (θ, γ) and $G(\cdot \mid \cdot, \cdot)$ be the corresponding equilibrium bid distribution given (Z, I). Suppose that there exists another structure $[\tilde{U}, \tilde{F}]$ satisfying A1 with parameters $(\tilde{\theta}, \tilde{\gamma})$ and leading to the same $G(\cdot \mid \cdot, \cdot)$. We first show that (θ, γ) is identified, *i.e.*, $(\theta, \gamma) = (\tilde{\theta}, \tilde{\gamma})$. Writing equation (4) for each structure gives

$$\frac{1}{I-1} \frac{\alpha}{g[b(\alpha; z, I) \mid z, I]} = \lambda[v(\alpha; z, I, \gamma) - b(\alpha; z, I); \theta] = \lambda[v(\alpha; z, I, \tilde{\gamma}) - b(\alpha; z, I); \tilde{\theta}],$$

for every $(z, I) \in \mathcal{Z} \times \mathcal{I}$. Hence, A1-(iv) implies that $(\tilde{\theta}, \tilde{\gamma}) = (\theta, \gamma)$. From A1-(i), $\tilde{U}(\cdot) = U(\cdot; \tilde{\theta}) = U(\cdot; \theta) = U(\cdot)$, which establishes the identification of $U(\cdot)$. Moreover, from equation (3), we have $v = b + \lambda^{-1}[G(b \mid z, I)/((I-1)g(b \mid z, I)); \theta] = \tilde{v}$, for every $b \in [\underline{b}(z, I), \overline{b}(z, I)]$ and $(z, I) \in \mathcal{Z} \times \mathcal{I}$. This shows that $\tilde{F}(\cdot \mid \cdot, \cdot) = F(\cdot \mid \cdot, \cdot)$, *i.e.*, that the latter is identified. ||

Proof of Proposition 5. Consider any pair (i, j) of individuals such that $c_i \neq c_j$. The compatibility condition (28) for a CRRA model is

$$b_j(\alpha; I) - b_i(\alpha; I) = (1 - c_i)/H_i(b_i(\alpha; I)) - (1 - c_j)/H_j(b_j(\alpha; I)) \quad \text{for all } \alpha \in [0, 1] \tag{A.3}$$

We first show that there exists $(\alpha, \tilde{\alpha}) \in [0, 1]^2$ such that $H_i(b_i(\alpha; I))/H_j(b_j(\alpha; I)) \neq H_i(b_i(\tilde{\alpha}; I))/H_j(b_j(\tilde{\alpha}; I))$. Suppose not, then $H_i(b_i(\alpha; I))/H_j(b_j(\alpha; I))$ is a constant. Because equation (A.3) at $\overline{b}(I)$ gives $H_i(\overline{b}(I))/H_j(\overline{b}(I)) = (1 - c_i)/(1 - c_j)$, then $H_i(b_i(\alpha; I))/H_j(b_j(\alpha; I)) = (1 - c_i)/(1 - c_j)$ for all $\alpha \in [0, 1]$. Using this in equation (A.3) gives $b_i(\alpha; I) = b_j(\alpha; I)$, which implies that $G_i(\cdot \mid I) = G_j(\cdot \mid I)$ and hence $H_i(\cdot) = H_j(\cdot)$ because $H_i(\cdot) = [g_j(\cdot \mid I)/G_j(\cdot \mid I)] + \sum_{k \neq i,j}[g_k(\cdot \mid I)/G_k(\cdot \mid I)]$ and $H_j(\cdot) = [g_i(\cdot \mid I)/G_i(\cdot \mid I)] + \sum_{k \neq i,j}[g_k(\cdot \mid I)/G_k(\cdot \mid I)]$. From equation (A.3) we then have $c_i = c_j$, which contradicts $c_i \neq c_j$. Thus, there exists $(\alpha, \tilde{\alpha}) \in [0, 1]^2$ such that $H_i(b_i(\alpha; I))/H_j(b_j(\alpha; I)) \neq H_i(b_i(\tilde{\alpha}; I))/H_j(b_j(\tilde{\alpha}; I))$. This guarantees that equation (A.3) written at α and $\tilde{\alpha}$ has a unique solution (c_i, c_j). Thus, $(c_1, \ldots, c_I)$ are identified. The identification of $F(\cdot \mid I)$ follows from equation (26). ||

APPENDIX B

Appendix B proves Theorem 2 and Theorem 3. The proofs use some lemmas, which are proved in Appendix C. The latter can be found in the Supplementary Material available on the web site of the review. Hereafter, let $\xi(\cdot; z, I) = s^{-1}(\cdot; z, I)$ and $\mathcal{F}_L$ be the σ-field generated by $\{(Z_\ell, I_\ell), 1 \leq \ell \leq L\}$. Moreover, let $a \asymp b$ mean that $a/b \to c$ with $0 < c < \infty$, and for $u = (u_{i\ell}) \in \boldsymbol{R}^N$, define the norms $\|u\|_p = \left(\sum_{\ell=1}^{L} \sum_{i=1}^{I_\ell} |u_{i\ell}|^p\right)^{1/p}$ and $\|u\|_\infty = \max_{1 \leq \ell \leq L} \max_{1 \leq i \leq I_\ell} |u_{i\ell}|$.

Proof of Theorem 1. The proof is in three steps.

Step 1. *Smoothness of $m(z, I; \beta)$. We have the following lemma.*

Lemma B1. *Let (U_0, F_0) satisfy A2-(i,ii) for some $\beta_0 = (\theta_0', \overline{v}_0)' \in \Theta^o \times (0, \infty)$ and $\mathcal{I}$ finite. Then, for every $I \in \mathcal{I}$, the function $m(\cdot, I; \cdot)$ defined in equation (7) is $R+1$ continuously differentiable on $\mathcal{Z} \times \mathcal{B}$, where $\mathcal{B} = \Theta \times (\sup_{(z,I) \in \mathcal{Z} \times \mathcal{I}} \overline{b}_0(z, I), +\infty)$ with $\sup_{(z,I) \in \mathcal{Z} \times \mathcal{I}} \overline{b}_0(z, I) < \overline{v}_0$.*

Step 2. *Perturbated model. Let $t > 0$ and $\psi(\cdot)$: $\boldsymbol{R}_- \to \boldsymbol{R}$ be an infinitely differentiable function on $\boldsymbol{R}_-$ with support $[-1, 0]$ such that $\psi(0) = 1$ and $\int \psi(x)dx = 0$. Let $\boldsymbol{I}_p = (1, \ldots, 1)' \in \boldsymbol{R}^p$. For a fixed constant $\kappa > 0$, consider the following perturbations of θ_0 and $g_0(b \mid z, I)$, where $I \in \mathcal{I}$,*

$$\beta_N = (\theta_N', \overline{v}_0)' = (\theta_0' + 2t\boldsymbol{I}_p'/\rho_N, \overline{v}_0)' = \beta_0 + (2t\boldsymbol{I}_p'/\rho_N, 0)',$$

$$g_N(b \mid z, I) = g_0(b \mid z, I) + \pi_N(z, I)\psi[\kappa\rho_N^{1/(R+1)}(b - \overline{b}_0(z, I))],$$

$$\pi_N(z, I) = m(z, I; \beta_N) - m(z, I; \beta_0) = \frac{\partial m(z, I; \beta_0)}{\partial \beta}(\beta_N - \beta_0) + o(\|\beta_N - \beta_0\|) = O(1/\rho_N).$$

Without loss of generality, we can assume that $\{\beta_N; N = 1, 2, \ldots\}$ is in a compact subset $\mathcal{B}_c \subset \mathcal{B}\,\overline{v} = \overline{v}_0\}$ as $\theta_0 \in \Theta^o$ and $\overline{v}_0 > \sup_{(z,I) \in \mathcal{Z} \times \mathcal{I}} \overline{b}_0(z, I)$. Thus, the reminder term is uniform in z because $\partial m(\cdot, I; \cdot)/\partial\beta$ is continuous on $\mathcal{Z} \times \mathcal{B}$ by Lemma B1 and hence uniformly continuous on $\mathcal{Z} \times \mathcal{B}_c$.

From Lemma 1-(iii), it follows that $g_N(\cdot \mid z, I)$ is a conditional density with support $[\underline{b}_0(z, I), \overline{b}_0\ (z, I)]$ for N large enough. Moreover, it is crucial to verify that such a density corresponds to a structure $[U(\cdot; \theta_N), F_N]$ in our semiparametric model.

Lemma B2. *Let (U_0, F_0) satisfy A2-(i,ii) for some $\beta_0 = (\theta_0, \overline{v}_0) \in \Theta^o \times (0, \infty)$, $f_0 \in \mathcal{F}_R^*(M)$, and $\mathcal{I}$ finite. For $\kappa > 0$ small enough and N large enough, we have*

(i) *for every $(z, I) \in \mathcal{Z} \times \mathcal{I}$, $G_N(\cdot \mid z, I)$ is rationalized by the IPV auction structure with risk aversion $[U(\cdot; \theta_N), F_N(\cdot \mid z, I)]$, where $F_N(\cdot \mid \cdot, I) \in \mathcal{F}_R^*$ with support $[\underline{v}_0(z, I), \overline{v}_0]$,*

(ii) *the conditional distribution function* $F_N(\cdot \mid \cdot,\cdot)$ *is such that* $(\beta_N, f_N) \in \mathcal{V}_\varepsilon(\beta_0, f_0)$.

Step 3. *Lower bound. Using the triangular inequality we have for any estimator* $\tilde{\beta}$,

$$\Pr_{\beta_N,f_N}(\|\rho_N(\tilde{\beta}-\beta_N)\|_\infty \geq t) \geq \Pr_{\beta_N,f_N}(\|\rho_N(\beta_N-\beta_0)\|_\infty - \|\rho_N(\tilde{\beta}-\beta_0)\|_\infty \geq t)$$

$$\geq \Pr_{\beta_N,f_N}(\|\rho_N(\tilde{\beta}-\beta_0)\|_\infty < t)$$

since $\|\rho_N(\beta_N-\beta_0)\|_\infty = 2t$. *Therefore, because* (β_0, f_0) *and* (β_N, f_N) *are in* $\mathcal{V}_\varepsilon(f_0,\beta_0)$ *for L large enough by Lemma B2-(ii), we have*

$$\sup_{(\beta,f)\in\mathcal{V}_\varepsilon(\beta_0,f_0)} \Pr_{\beta,f}(\|\rho_N(\tilde{\beta}-\beta)\|_\infty \geq t)$$

$$\geq \frac{1}{2}\left[\Pr_{\beta_0,f_0}(\|\rho_N(\tilde{\beta}-\beta_0)\|_\infty \geq t) + \Pr_{\beta_N,f_N}(\|\rho_N(\tilde{\beta}-\beta_N)\|_\infty \geq t)\right]$$

$$\geq \frac{1}{2}\mathrm{E}\left[\Pr_{\beta_0,f_0}(\|\rho_N(\tilde{\beta}-\beta_0)\|_\infty \geq t \mid \mathcal{F}_L) + \Pr_{\beta_N,f_N}(\|\rho_N(\tilde{\beta}-\beta_0)\|_\infty < t \mid \mathcal{F}_L)\right] \tag{B.1}$$

Let $\Pr_e(\mathcal{F}_L)$ *denote the term within brackets and* $\Pr_j$ *be the joint probability of the* $B_{i\ell}$*s given* $\mathcal{F}_L$ *under* $g_j(\cdot \mid \cdot,\cdot)$, $j = 0, N$. *Standard relations between the distance in variation, the* L_1 *norm, and the Hellinger distance (see, e.g., Bickel, Klaassen, Ritov and Wellner (1993, p. 464)) yield*

$$\Pr_e(\mathcal{F}_L) = 1 - (\Pr_0(\|\rho_N(\tilde{\beta}-\beta_0)\|_\infty < t) - \Pr_N(\|\rho_N(\tilde{\beta}-\beta_0)\|_\infty < t))$$

$$\geq 1 - \sup_A |\Pr_0(A) - \Pr_N(A)| = 1 - \frac{1}{2}\int |d\Pr_0 - d\Pr_N|$$

$$\geq 1 - \left[\int (\sqrt{d\Pr_0} - \sqrt{d\Pr_N})^2\right]^{1/2} = 1 - \sqrt{2}\left(1 - \int \sqrt{d\Pr_0 d\Pr_N}\right)^{1/2}$$

$$= 1 - \sqrt{2}\left(1 - \prod_{\ell=1}^{L}\prod_{i=1}^{I_\ell}\int_{\underline{b}_0}^{\overline{b}_0(Z_\ell,I_\ell)} \sqrt{g_0(b_{i\ell} \mid Z_\ell, I_\ell) g_N(b_{i\ell} \mid Z_\ell, I_\ell)}\, db_{i\ell}\right)^{1/2}. \tag{B.2}$$

But, because $g_j(\cdot \mid \cdot,\cdot)$, $j = 0, N$, *are bounded away from zero and* $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}} \pi_N(z, I) = O(1/\rho_N)$, *we obtain from the definition of* $g_N(\cdot \mid \cdot,\cdot)$ *and a Taylor expansion*

$$\int_{\underline{b}_0(Z_\ell,I_\ell)}^{\overline{b}_0(Z_\ell,I_\ell)} \sqrt{g_0(b \mid Z_\ell, I_\ell) g_N(b \mid Z_\ell, I_\ell)}\, db$$

$$= \int_{\underline{b}_0(Z_\ell,I_\ell)}^{\overline{b}_0(Z_\ell,I_\ell)} g_0(b \mid Z_\ell, I_\ell)\sqrt{1 + \frac{\pi_N(Z_\ell, I_\ell)}{g_0(b \mid Z_\ell, I_\ell)}\psi\left(\kappa\rho_N^{\frac{1}{R+1}}(b - \overline{b}_0(Z_\ell, I_\ell))\right)}\, db$$

$$= \int_{\underline{b}_0(Z_\ell,I_\ell)}^{\overline{b}_0(Z_\ell,I_\ell)} g_0(b \mid Z_\ell, I_\ell)\left[1 + \frac{\pi_N(Z_\ell, I_\ell)}{2g_0(b \mid Z_\ell, I_\ell)}\psi\left(\kappa\rho_N^{\frac{1}{R+1}}(b - \overline{b}_0(Z_\ell, I_\ell))\right)\right.$$

$$\left. - \frac{\pi_N^2(Z_\ell, I_\ell)}{8g_0^2(b \mid Z_\ell, I_\ell)}\psi^2\left(\kappa\rho_N^{\frac{1}{R+1}}(b - \overline{b}_0(Z_\ell, I_\ell))\right)\right] db + O\left(\frac{1}{\rho_N^3}\right)$$

$$= 1 + \frac{\pi_N(Z_\ell, I_\ell)}{2}\int_{\underline{b}_0(Z_\ell,I_\ell)}^{\overline{b}_0(Z_\ell,I_\ell)} \psi\left(\kappa\rho_N^{\frac{1}{R+1}}(b - \overline{b}_0(Z_\ell, I_\ell))\right) db$$

$$-\frac{\pi_N^2(Z_\ell,I_\ell)}{8\kappa\rho_N^{\frac{1}{R+1}}}\int_{-1}^{0}\frac{\psi^2(x)}{g_0\left(\overline{b}_0(Z_\ell,I_\ell)+\rho_N^{-\frac{1}{R+1}}x/\kappa\right)}dx+O(\rho_N^{-3})$$

$$=1+0+O\left(\rho_N^{-\frac{1}{R+1}-2}\right)=1+O\left(\rho_N^{-\frac{2R+3}{R+1}}\right),$$

uniformly in ℓ, since $\int\psi(x)dx=0$. Consequently, since $N\rho_N^{-(2R+3)/(R+1)}\to 0$, we have

$$\prod_{\ell=1}^{L}\prod_{i=1}^{I_\ell}\int_{\underline{b}_0(Z_\ell,I_\ell)}^{\overline{b}_0(Z_\ell,I_\ell)}\sqrt{g_0(b_{i\ell}\mid Z_\ell,I_\ell)g_N(b_{i\ell}\mid Z_\ell,I_\ell)}db_{i\ell}$$

$$=\left[1+O\left(\rho_N^{-\frac{2R+3}{R+1}}\right)\right]^N=\exp[N\log(1+O(\rho_N^{-(2R+3)/(R+1)}))]=1+0(1).$$

Hence, equation (B.2) implies that $\Pr_e(\mathcal{F}_L)\geq 1-o(1)$. *Thus, equation (B.1) yields*

$$\inf_{\tilde{\beta}}\sup_{(\beta,f)\in\mathcal{V}_\varepsilon(\beta_0,f_0)}\Pr_{\beta,f}(\|\rho_N(\tilde{\beta}-\beta)\|_\infty\geq t)\geq\frac{1}{2}[1-o(1)]=\frac{1}{2}-o(1).$$

The desired result follows by taking limits as $N\to\infty$ and then $\varepsilon\to 0$. ||

Proof of Theorem 2. The proof is in three steps. All limits are taken as $N\to\infty$.

Step 1. *Some lemmas. The first lemma studies the bias and error terms of equation (10).*

Lemma B3. *Let A2–A3 and A4-(iii, iv) hold.*

(i) *The variables $Y_{i\ell}$ (or $\varepsilon_{i\ell}$), $1\leq i\leq I_\ell$, $1\leq\ell\leq L$, are independent given $\mathcal{F}_L$.*
(ii) *Uniformly in (i,ℓ),*

$$\mathrm{E}[Y_{i\ell}\mid\mathcal{F}_L]=g_0(\overline{b}_0(Z_\ell,I_\ell)\mid Z_\ell,I_\ell)+\frac{h_N^{R+1}}{(R+1)!}\left(\frac{\partial^{R+1}g_0(\overline{b}_0(Z_\ell,I_\ell)\mid Z_\ell,I_\ell)}{\partial b^{R+1}}\int x^{R+1}\Phi(x)dx+o(1)\right),$$

$$e_{i\ell}=\frac{h_N^{R+1}}{(R+1)!}\left(\frac{\partial^{R+1}g_0(\overline{b}_0(Z_\ell,I_\ell)\mid Z_\ell,I_\ell)}{\partial b^{R+1}}\int x^{R+1}\Phi(x)dx+o(1)\right).$$

(iii) *Uniformly in (i,ℓ),*

$$\mathrm{Var}[\varepsilon_{i\ell}\mid\mathcal{F}_L]=\frac{g_0(\overline{b}_0(Z_\ell,I_\ell)\mid Z_\ell,I_\ell)+o(1)}{h_N}\int\Phi^2(x)dx=\frac{m(Z_\ell,I_\ell;\beta_0)+o(1)}{h_N}\int\Phi^2(x)dx,$$

$$\max_{1\leq\ell\leq L,1\leq i\leq I_\ell}|\varepsilon_{i\ell}|\leq\frac{2\sup_{x\in\boldsymbol{R}}|\Phi(x)|}{h_N}.$$

The second lemma is a central limit theorem, which is useful for weighted averages of $\varepsilon_{i\ell}$.

Lemma B4. *Let A2–A3 and A4-(iii, iv) hold.*
For any $u\in\boldsymbol{R}^N\setminus\{0\}$ that is $\mathcal{F}_L$-measurable with $\|u\|_\infty/(\|u\|_2\sqrt{h_N})=o_P(1)$, then $\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}u_{i\ell}\,\varepsilon_{i\ell}/\mathrm{Var}^{1/2}\left[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}u_{i\ell}\mid\mathcal{F}_L\right]\xrightarrow{d}\mathcal{N}(0,1)$ conditionally on $\mathcal{F}_L$ and thus unconditionally.

The third and fourth lemmas control the estimation errors $|\hat{Y}_{i\ell}-Y_{i\ell}|$ and $|\hat{m}(\cdot,\cdot;\beta)-m(\cdot,\cdot;\beta)|$ arising from estimating the upper boundary $\overline{b}_0(\cdot,\cdot)$.

Lemma B5. *Let A2–A3 and A4-(iii–v) hold.*

For any $u \in \boldsymbol{R}^N$ that is $\mathcal{F}_L$-measurable,

$$\left|\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell} u_{i\ell}(\hat{Y}_{i\ell}-Y_{i\ell})\right| \le \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}|u_{i\ell}||\hat{Y}_{i\ell}-Y_{i\ell}| = O_P\left[\max\left(\|u\|_1\frac{a_N}{h_N}, \|u\|_2\frac{\sqrt{a_N}}{h_N}\right)\right],$$

$$\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}|u_{i\ell}|(\hat{Y}_{i\ell}-Y_{i\ell})^2 = \|u\|_1 O_P\left(\frac{a_N}{h_N^2}\right).$$

Lemma B6. *Let A2-(i, ii), A3-(i), and A4-(i, v) hold. Then,* $\sup_{\beta\in\mathcal{B}_\delta}\max_{1\le\ell\le L}|\hat{m}(Z_\ell, I_\ell;\beta)-m(Z_\ell, I_\ell;\beta)|$ *and* $\sup_{\beta\in\mathcal{B}_\delta}\max_{1\le\ell\le L}\|\partial\hat{m}(Z_\ell, I_\ell;\beta)/\partial\beta-\partial m(Z_\ell, I_\ell;\beta)/\partial\beta\|_\infty$ *are both* $O_P(a_N)$.

The next two lemmas study the properties of the limit and convergence of the approximate objective function $\overline{Q}_N(\cdot)$ defined in equation (19).

Lemma B7. *Let A2–A3 and A4-(i, ii) hold. Let* $\overline{Q}(\beta)=\mathrm{E}[I\omega(Z,I)(m(Z,I;\beta)-m(Z,I;\beta_0))^2]$. *Then, for any* $\varepsilon>0$, *there exists* $C_\varepsilon>0$ *such that* $\inf_{\beta\in\mathcal{B}_\delta;\|\beta-\beta_0\|_\infty\ge\varepsilon}\overline{Q}(\beta)>C_\varepsilon$. *Moreover, the matrix* $A(\beta)$ *and* $B(\beta)$ *defined in equations (14) and (15) are of full rank in a neighbourhood of* β_0.

Lemma B8. *Let A2–A3 and A4-(i, ii) hold. Then,* $\sup_{\beta\in\mathcal{B}_\delta}|(1/L)\overline{Q}_N(\beta)-\overline{Q}(\beta)| = O_P(1/\sqrt{L}) = o_P(1)$. *Moreover, for any* $\beta\in\mathcal{B}_\delta$

$$\frac{A_N(\beta)}{N}=A(\beta)+O_P(1/\sqrt{N}),\quad \frac{B_N(\beta)}{N}=B(\beta)+O_P(1/\sqrt{N}),\quad \frac{\mathbf{b}_N(\beta,g_0)}{N}=\mathbf{b}(\beta,g_0)+O_P(1/\sqrt{N}),$$

where $A(\beta)$, $B(\beta)$, $A_N(\beta)$, $B_N(\beta)$, *and* $\mathbf{b}(\beta,g_0)$ *are defined in equations (14)–(18), and*

$$\mathbf{b}_N(\beta,g_0)=\frac{\int x^{R+1}\Phi(x)dx}{(R+1)!}\sum_{\ell=1}^{L}I_\ell\omega(Z_\ell,I_\ell)\frac{\partial^{R+1}g_0(\bar{b}_0(Z_\ell,I_\ell)\mid Z_\ell,I_\ell)}{\partial b^{R+1}}\frac{\partial m(Z_\ell,I_\ell;\beta)}{\partial\beta}.$$

The last lemma deals with the following processes:

$$W_N(\beta)=\frac{\sqrt{h_N}}{\sqrt{N}}\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)\varepsilon_{i\ell}m(Z_\ell,I_\ell;\beta),$$

$$W_N^{(1)}(\beta)=\frac{\sqrt{h_N}}{\sqrt{N}}\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)\varepsilon_{i\ell}\left(\frac{\partial m(Z_\ell,I_\ell;\beta)}{\partial\beta}-\frac{\partial m(Z_\ell,I_\ell;\beta_0)}{\partial\beta}\right).$$

Lemma B9. *Let A2–A3 and A4-(i–iv) hold. If* $\tilde{\beta}_N=\beta_0+o_P(1)$, *then* $\sup_{\beta\in\mathcal{B}_\delta}|W_N(\beta)|=O_P(1)$ *and* $W_N^{(1)}(\tilde{\beta}_N)=o_P(1)$.

Step 2. *Consistency. Note that* $|\max_{1\le\ell\le L}\hat{\bar{b}}_N(Z_\ell,I_\ell)-\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\bar{b}_0(z,I)|<\delta/4$ *with probability approaching one by A4-(v) and A3-(i), where the latter implies that* $\{Z_\ell,\ell=1,2,\ldots\}$ *is a.s. dense in* $\mathcal{Z}$ *by the Glivenko–Cantelli Theorem. Thus,* $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\bar{b}_0(z,I)+\delta/4<\max_{1\le\ell\le L}\hat{\bar{b}}_N(Z_\ell,I_\ell)+\delta/2<\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\bar{b}_0(z,I)+3\delta/4<\overline{v}_0<\overline{v}_{\sup}$ *with probability approaching one, using A4-(i). That is,* $\overline{v}_0\in\mathcal{B}_N\subset\mathcal{B}_{\delta/4}$ *with probability approaching one.*

Now, equations (12) and (13) and the triangular inequality give

$$|\hat{Q}_N^{1/2}(\beta)-Q_N^{1/2}(\beta)|$$
$$=\left|\left[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(\hat{Y}_{i\ell}-\hat{m}(Z_\ell,I_\ell;\beta))^2\right]^{1/2}-\left[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(Y_{i\ell}-m(Z_\ell,I_\ell;\beta))^2\right]^{1/2}\right|$$

$$\leq \left[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(\hat{Y}_{i\ell}-Y_{i\ell}+m(Z_\ell,I_\ell;\beta)-\hat{m}(Z_\ell,I_\ell;\beta))^2\right]^{1/2}$$

$$\leq \left[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(\hat{Y}_{i\ell}-Y_{i\ell})^2\right]^{1/2}+\left[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(m(Z_\ell,I_\ell;\beta)-\hat{m}(Z_\ell,I_\ell;\beta))^2\right]^{1/2}.$$

Thus, Lemmas B5 and B6 together with A4-(ii) yield

$$\sup_{\beta\in\mathcal{B}_{\delta/4}}|\hat{Q}_N^{1/2}(\beta)-Q_N^{1/2}(\beta)|=\sqrt{N}O_P\left(\frac{\sqrt{a_N}}{h_N}\right)+\sqrt{N}O_P(a_N)=\sqrt{N}O_P\left(\frac{\sqrt{a_N}}{h_N}\right), \tag{B.3}$$

since $a_N=o(\sqrt{a_N}/h_N)$ *by A4-(iv). On the other hand, equations (10) and (12) and the inequality* $(x_1+x_2+x_3)^2\leq 3(x_1^2+x_2^2+x_3^2)$ *yield*

$$Q_N(\beta)\leq 3\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)[(m(Z_\ell,I_\ell;\beta)-m(Z_\ell,I_\ell;\beta_0))^2+e_{i\ell}^2+\varepsilon_{i\ell}^2]$$

$$=O_P(N)+O_P(Nh_N^{2(R+1)})+O_P(N/h_N)=O_P(N/h_N), \tag{B.4}$$

uniformly in $\beta\in\mathcal{B}_{\delta/4}$, *where the first equality follows from A4-(ii, iii), Lemmas 1-(i, iv), B1, B3-(ii), and* $\sum_\ell\sum_i\varepsilon_{i\ell}^2=O_p(1/h_N)$, *which follows from the Markov inequality and* $\mathrm{E}[\varepsilon_{i\ell}^2]=\mathrm{E}\{\mathrm{Var}[\varepsilon_{i\ell}^2\mid\mathcal{F}_L]\}=O(1/h_N)$ *using* $\mathrm{E}[\varepsilon_{i\ell}\mid\mathcal{F}_L]=0$ *and Lemma B3-(iii). The second equality then follows from A4-(iv). Thus, combining (B.3) and (B.4) gives*

$$\sup_{\beta\in\mathcal{B}_{\delta/4}}|\hat{Q}_N(\beta)-Q_N(\beta)|=\sup_{\beta\in\mathcal{B}_{\delta/4}}|(\hat{Q}_N^{1/2}(\beta)-Q_N^{1/2}(\beta))(2\hat{Q}_N^{1/2}(\beta)+Q_N^{1/2}(\beta)-\hat{Q}_N^{1/2}(\beta))|$$

$$\leq 2\sup_{\beta\in\mathcal{B}_{\delta/4}}Q_N^{1/2}(\beta)\sup_{\beta\in\mathcal{B}_{\delta/4}}|\hat{Q}_N^{1/2}(\beta)-Q_N^{1/2}(\beta)|+\sup_{\beta\in\mathcal{B}_{\delta/4}}|\hat{Q}_N^{1/2}(\beta)-Q_N^{1/2}(\beta)|^2$$

$$=NO_P\left(\sqrt{\frac{a_N}{h_N^3}}\right)+NO_P\left(\frac{a_N}{h_N^2}\right)=o_P(N), \tag{B.5}$$

since $a_N=o(h_N^3)$ *by A4-(v).*

Next, consider $Q_N(\beta)-Q_N(\beta_0)-\overline{Q}_N(\beta)$, *where* $\overline{Q}_N(\beta)$ *is defined by equation (19). We have*

$$Q_N(\beta)-Q_N(\beta_0)=\overline{Q}_N(\beta)-2\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(e_{i\ell}+\varepsilon_{i\ell})(m(Z_\ell,I_\ell;\beta)-m(Z_\ell,I_\ell;\beta_0))$$

using equations (10) and (12). Hence,

$$\sup_{\beta\in\mathcal{B}_{\delta/4}}|Q_N(\beta)-Q_N(\beta_0)-\overline{Q}_N(\beta)|=O_P(Nh_N^{R+1})+O_P(\sqrt{N/h_N})=o_P(N), \tag{B.6}$$

using Lemmas 1-(i, iv), B1, B3-(ii), B9, and A4-(iv). Thus,

$$\sup_{\beta\in\mathcal{B}_{\delta/4}}\left|\frac{1}{L}\hat{Q}_N(\beta)-\frac{1}{L}\hat{Q}_N(\beta_0)-\overline{Q}(\beta)\right|\leq\sup_{\beta\in\mathcal{B}_{\delta/4}}\left|\frac{1}{L}\hat{Q}_N(\beta)-\frac{1}{L}Q_N(\beta)\right|+\left|\frac{1}{L}\hat{Q}_N(\beta_0)-\frac{1}{L}Q_N(\beta_0)\right|$$

$$+\sup_{\beta\in\mathcal{B}_{\delta/4}}\left|\frac{1}{L}Q_N(\beta)-\frac{1}{L}Q_N(\beta_0)-\frac{1}{L}\overline{Q}_N(\beta)\right|$$

$$+\sup_{\beta\in\mathcal{B}_{\delta/4}}\left|\frac{1}{L}\overline{Q}_N(\beta)-\overline{Q}(\beta)\right|$$

$$=o_P(1)$$

using (B.5), (B.6), Lemma B8, and $L\asymp N$. *Combining this with Lemma B7 and recalling that* $\overline{v}_0\in\mathcal{B}_N\subset\mathcal{B}_{\delta/4}$ *with probability approaching one show that the usual consistency conditions of M-estimators are satisfied (see, e.g., White, 1994). Hence,* $\hat{\beta}_N$ *converges in probability to* β_0.

Step 3. *Asymptotic normality. Given A4-(i), we have* $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\overline{b}_0(z,I)+\delta<\overline{v}_0<\overline{v}_{\sup}$. *Thus,* β_0 *is an inner point of* $\mathcal{B}_N$ *with probability approaching one. Hence, because* $\hat{\beta}_N \xrightarrow{P} \beta_0$, $\hat{\beta}_N$ *solves with probability approaching one the first-order conditions*

$$0=\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(\hat{Y}_{i\ell}-\hat{m}(Z_\ell,I_\ell;\hat{\beta}_N))\frac{\partial\hat{m}(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}.$$

Taking a Taylor expansion with integral remainder of $\hat{m}(Z_\ell,I_\ell;\hat{\beta}_N)$ *around* β_0 *and solving give*

$$\hat{\beta}_N-\beta_0=\left[\sum_{\ell=1}^{L}I_\ell\omega(Z_\ell,I_\ell)\frac{\partial\hat{m}(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}\int_0^1\frac{\partial\hat{m}(Z_\ell,I_\ell;\beta_0+t(\hat{\beta}_N-\beta_0))}{\partial\beta'}dt\right]^{-1}$$
$$\times\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(\hat{Y}_{i\ell}-\hat{m}(Z_\ell,I_\ell;\beta_0))\frac{\partial\hat{m}(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}. \tag{B.7}$$

Let $\hat{J}_N$ *be the term within brackets. Lemmas B1, B6, B8, and the consistency of* $\hat{\beta}_N$ *yield*

$$\begin{aligned}\hat{J}_N&=\sum_{\ell=1}^{L}I_\ell\omega(Z_\ell,I_\ell)\frac{\partial m(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}\int_0^1\frac{\partial m(Z_\ell,I_\ell;\beta_0+t(\hat{\beta}_N-\beta_0))}{\partial\beta'}dt+O_P(Na_N)\\&=A_N(\beta_0)+o_P(N)=NA(\beta_0)+o_P(N),\end{aligned} \tag{B.8}$$

where $A(\beta_0)$ *is non-singular by Lemma B7. Next, we study the second term in equation (B.7), i.e.,*

$$\begin{aligned}\hat{S}_N&=\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(Y_{i\ell}-m(Z_\ell,I_\ell;\beta_0))\frac{\partial m(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}\\&+\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)\left(m(Z_\ell,I_\ell;\beta_0)\frac{\partial m(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}-\hat{m}(Z_\ell,I_\ell;\beta_0)\frac{\partial\hat{m}(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}\right)\\&+\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)\left[(\hat{Y}_{i\ell}-Y_{i\ell})\frac{\partial\hat{m}(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}+Y_{i\ell}\left(\frac{\partial\hat{m}(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}-\frac{\partial m(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}\right)\right]\end{aligned}$$

From equation (10), we have $\sum_\ell\sum_i|Y_{i\ell}|=O_P(N/\sqrt{h_N})$ *by Lemmas B1 and B3, using the Markov inequality and* $E|\varepsilon_{i\ell}|\le[E(\varepsilon_{i\ell}^2)]^{1/2}$ *to get* $\sum_\ell\sum_i|\varepsilon_{i\ell}|=O_P(N/\sqrt{h_N})$, *which is the leading term given A4-(iv). Therefore, using Lemmas B1, B5, and B6 together with A4-(ii), we obtain*

$$\begin{aligned}\hat{S}_N&=\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(Y_{i\ell}-m(Z_\ell,I_\ell;\beta_0))\frac{\partial m(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}\\&+NO_P(a_N)+O_P\left[\max\left(N\frac{a_N}{h_N},\left(\frac{Na_N}{h_N^2}\right)^{1/2}\right)\right]+O_P\left(\frac{Na_N}{\sqrt{h_N}}\right)\\&=\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(Y_{i\ell}-m(Z_\ell,I_\ell;\beta_0))\frac{\partial m(Z_\ell,I_\ell;\beta_0)}{\partial\beta}\\&+\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)(Y_{i\ell}-m(Z_\ell,I_\ell;\beta_0))\left(\frac{\partial m(Z_\ell,I_\ell;\hat{\beta}_N)}{\partial\beta}-\frac{\partial m(Z_\ell,I_\ell;\beta_0)}{\partial\beta}\right)\\&+O_P\left[\max\left(N\frac{a_N}{h_N},\left(\frac{Na_N}{h_N^2}\right)^{1/2}\right)\right].\end{aligned}$$

Using equation (10), the consistency of $\hat{\beta}_N$, Lemmas B1, B3, and B9 with A2-(ii) implies that the second term is an $o_P(Nh_N^{R+1}) + o_P(\sqrt{N/h_N})$. *Note that* $Na_N/h_N = o(Nh_N^{R+1})$ *and* $Na_N/h_N^2 = o(N^{1/2}h_N^{-3/2}) = o(N/h_N)$ *under A4-(iv, v). Hence, equation (10) and Lemmas B3 and B8 imply that*

$$\hat{S}_N = \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell, I_\ell)(e_{i\ell}+\varepsilon_{i\ell})\frac{\partial m(Z_\ell, I_\ell;\beta_0)}{\partial\beta} + o_P\left(Nh_N^{R+1}+\sqrt{\frac{N}{h_N}}\right)$$
$$= Nh_N^{R+1}\mathbf{b}(\beta_0, g_0) + \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell, I_\ell)\varepsilon_{i\ell}\frac{\partial m(Z_\ell, I_\ell;\beta_0)}{\partial\beta} + o_P\left(Nh_N^{R+1}+\sqrt{\frac{N}{h_N}}\right). \tag{B.9}$$

Let $u_{i\ell} = \omega(Z_\ell, I_\ell)\partial m(Z_\ell, I_\ell;\beta_0)/\partial\beta$. *Using equation (17), Lemmas B1, B3-(iii), B8, and A4-(ii) gives*

$$\mathrm{Var}\left(\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}u_{i\ell}\varepsilon_{i\ell}\,|\mathcal{F}_L\right) = \frac{B_N(\beta_0)}{h_N}\int\Phi^2(x)dx + o_P\left(\frac{N}{h_N}\right) = \frac{N}{h_N}\left(B(\beta_0)\int\Phi^2(x)dx + o_p(1)\right).$$

Because $\|u\|_\infty/(\|u\|_2\sqrt{h_N}) = O_P(1/\sqrt{Nh_N}) = o_P(1)$ *by A4-(iv), Lemma B4 implies that*

$$\sqrt{\frac{h_N}{N}}\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell, I_\ell)\varepsilon_{i\ell}\frac{\partial m(Z_\ell, I_\ell;\beta_0)}{\partial\beta} \xrightarrow{d} \mathcal{N}\left(0, B(\beta_0)\int\Phi^2(x)dx\right). \tag{B.10}$$

Collecting equations (B.8)–(B.10) and using $\hat{\beta}_N - \beta_0 = \hat{J}_N^{-1}\hat{S}_N$ *from equation (B.7) give*

$$\hat{\beta}_N - \beta_0 = h_N^{R+1}A(\beta_0)^{-1}\mathbf{b}(\beta_0, g_0)$$
$$+\frac{1}{\sqrt{Nh_N}}A(\beta_0)^{-1}\sqrt{\frac{h_N}{N}}\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell, I_\ell)\varepsilon_{i\ell}\frac{\partial m(Z_\ell, I_\ell;\beta_0)}{\partial\beta} + o_P\left(h_N^{R+1}+\frac{1}{\sqrt{Nh_N}}\right),$$

showing that $\hat{\beta}_N - \beta_0 = O_P(h_N^{R+1} + 1/\sqrt{Nh_N})$. *This also gives the limits in probability and in distribution of Theorem 3-(ii, iii). Moreover,* $N^{-1}\hat{A}_N(\hat{\beta}_N) = A(\beta_0) + o_P(1)$ *and* $N^{-1}\hat{B}_N(\hat{\beta}_N) = B(\beta_0) + o_P(1)$ *can be established arguing as in equation (B.8).* ||

The proofs of Lemmas B1–B9 stated in this appendix can be found in Appendix C in the Supplementary Material on the web site of the review.

Acknowledgement. We are grateful to Jingfeng Lu, Jim Powell, and Mathew Shum for their suggestions as well as to the co-editor and three referees for their constructive comments. Previous versions were presented at the Malinvaud seminar, University of Cambridge, University College London, Oxford University, University of Toulouse, University of California at Los Angeles and Berkeley, University of Rochester, University of Pittsburgh, Penn State University, Princeton University, Rutgers University, New York University, Harvard-MIT, Cornell University, the Tow Conference on Auctions at the University of Iowa, the Cowles Foundation Conference on Strategy and Decision Making at Yale University, the SITE Meeting at Stanford University, the World Congress of the Econometric Society 2000, and the European Congress of the Econometric Society 2002. We thank Phil Haile and Jason Cummins for providing the data set used in this paper. Financial support from the National Science Foundation grant SES-0452154 is gratefully acknowledged by the last two authors. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

SUPPLEMENTARY DATA

Supplementary data are available at *Review of Economic Studies* online

REFERENCES

ACKERBERG, D., HIRANO, K. and SHAHRIAR, Q. (2006), "The Buy-it-now Option, Risk Aversion and Impatience in an Empirical Model of eBay Bidding" (Working Paper, University of Arizona).

ATHEY, S. and HAILE, P. (2007), "Nonparametric Approaches to Auctions", in Heckman, J. and Leamer, E. (eds) *Handbook of Econometrics*, vol. VI (Amsterdam: North Holland) 3847–3966.

ATHEY, S. and LEVIN, J. (2001), "Information and Competition in U.S. Forest Service Timber Auctions", *Journal of Political Economy*, **109**, 375–417.

ATHEY, S., LEVIN, J. and SEIRA, E. (2004), "Comparing Open and Sealed Bid Auctions: Theory and Evidence from Timber Auctions" (Working Paper, Stanford University).

BAJARI, P. and HORTACSU, A. (2005), "Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data", *Journal of Political Economy*, **113**, 703–741.

BALDWIN, L. (1995), "Risk Averse Bidders at Oral Ascending-Bid Forest Service Timber Sales" (Working Paper, RAND Corporation).

BALDWIN, L., MARSHALL, R. and RICHARD, J. F. (1997), "Bidder Collusion at Forest Service Timber Sales", *Journal of Political Economy*, **105**, 657–699.

BHATTACHARYA, P. and GANGOPADHYAY, A. (1990), "Kernel and Nearest-Neighbor Estimation of a Conditional Quantile", *Annals of Statistics*, **18**, 1400–1415.

BICKEL, P. J., KLAASSEN, C. A. J., RITOV, Y. and WELLNER, J. A. (1993), *Efficient and Adaptive Estimation for Semiparametric Models* (Baltimore: Johns Hopkins University Press).

BIERENS, H. J. (1987), "Kernel Estimators of Regression Functions", in Bewley, T. (ed) *Advances in Econometrics* (Cambridge: Cambridge University Press) 99–144.

CAMPO, S. (2005), "Asymmetry and Risk Aversion within the Independent Private Paradigm: The Case of the Construction Procurements" (Working Paper, University of North Carolina).

CAMPO, S., PERRIGNE, I. and VUONG, Q. (2003), "Asymmetry in First-Price Auctions with Affiliated Private Values", *Journal of Applied Econometrics*, **18**, 197–207.

CHAUDHURI, P. (1991), "Global Nonparametric Estimation of Conditional Quantile Functions and Their Derivatives", *Journal of Multivariate Analysis*, **39**, 246–269.

CHE, Y. K. and GALE, I. (1998), "Standard Auctions with Financially Constrained Bidders", *Review of Economic Studies*, **65**, 1–21.

CHERNOZHUKOV, V., HONG, H. and TAMER, E. (2007), "Estimation and Confidence Regions for Parameter Sets in Econometric Models", *Econometrica*, **75**, 1243–1284.

COHEN, A. and EINAV, L. (2007), "Estimating Risk Preferences from Deductible Choices", *American Economic Review*, **97**, 745–788.

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

ESO, P. and WHITE, L. (2004), "Precautionary Bidding in Auctions", *Econometrica*, **72**, 77–92.

FLORENS, J. P., MARIMOUTOU, V. and PEQUIN-FEISSOLLE, A. (2007), *Econometric Modeling and Inference* (New York: Cambridge University Press).

GOEREE, J., HOLT, C. and PALFREY, T. (2002), "Quantal Response Equilibrium and Overbidding in Private Value Auctions", *Journal of Economic Theory*, **104**, 247–272.

GUERRE, E., PERRIGNE, I. and VUONG, Q. (2000), "Optimal Nonparametric Estimation of First-Price Auctions", *Econometrica*, **68**, 525–574.

GUERRE, E., PERRIGNE, I. and VUONG, Q. (2009), "Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions", *Econometrica*, **77**, 1193–1227.

GUERRE, E. and SABBAH, C. (2008), "Uniform Bahadur Approximation for Local Polynomial Estimators of Conditional Quantiles" (Working Paper, Queen Mary University London).

HAILE, P. (2001), "Auctions with Resale Markets: An Application to U.S. Forest Service Timber Sales", *American Economic Review*, **91**, 399–427.

HALL, P., WOLFF, R. C. L. and YAO, Q. W. (1999), "Methods for Estimating a Conditional Distribution Function", *Journal of the American Statistical Association*, **94**, 154–163.

HANSEN, R. (1985), "Empirical Testing of Auction Theory", *American Economic Review*, **75**, 156–159.

HART, J. (1997), *Nonparametric Smoothing and Lack-of-Fit Tests* (New York: Springer Verlag).

HONORE, B. E. and KYRIAZIDOU, E. (2000), "Panel Data Discrete Choice Models with Lagged Dependent Variable", *Econometrica*, **68**, 839–874.

HOROWITZ, J. L. (1992), "A Smoothed Maximum Score Estimator for the Binary Response Model", *Econometrica*, **60**, 505–531.

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

KONG, E., LINTON, O. and XIA, Y. (2008), "Polynomial Estimates of M-regression and Its Application to the Additive Model" (Working Paper, London School of Economics).

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

KYRIAZIDOU, E. (1997), "Estimation of a Panel Data Sample Selection Model", *Econometrica*, **65**, 1335–1364.

LI, T., PERRIGNE, I. and VUONG, Q. (2002), "Structural Estimation of the Affiliated Private Value Auction Model", *Rand Journal of Economics*, **33**, 171–193.

LI, T., PERRIGNE, I. and VUONG, Q. (2003), "Semiparametric Estimation of the Optimal Reserve Price in First-Price Auctions", *Journal of Business and Economic Statistics*, **21**, 53–64.

LU, J. and PERRIGNE, I. (2008), "Estimating Risk Aversion from Ascending and Sealed-Bid Auctions: The Case of Timber Auction Data", *Journal of Applied Econometrics*, **21**, 1221–1247.

MASKIN, E. and RILEY, J. (1984), "Optimal Auctions with Risk Averse Buyers", *Econometrica*, **52**, 1473–1518.

MATTHEWS, S. (1987), "Comparing Auctions for Risk Averse Buyers: A Buyer's Point of View", *Econometrica*, **55**, 633–646.

NEWEY, W. K. and MCFADDEN, D. L. (1994), "Large Sample Estimation and Hypothesis Testing", in Engle, R. F. and McFadden, D. L. (eds) *Handbook of Econometrics*, vol. IV (Amsterdam: North Holland) 2111–2245.

PERRIGNE, I. (2003), "Random Reserve Prices and Risk Aversion in Timber Auctions" (Working Paper, Pennsylvania State University).

PERRIGNE, I. and VUONG, Q. (1999), "Structural Econometrics of First-Price Auctions: A Survey of Methods", *Canadian Journal of Agricultural Economics*, **47**, 203–223.

PERRIGNE, I. and VUONG, Q. (2009), "Cost Uncertainty in Procurements" (Working Paper, Pennsylvania State University).

POWELL, J. L. (1994), "Estimation of Semiparametric Models", in Engle, R. F. and McFadden, D. L. (eds) *Handbook of Econometrics*, vol. IV (Amsterdam: North Holland) 2443–2521.

PRATT, J. (1964), "Risk Aversion in the Small and in the Large", *Econometrica*, **32**, 122–136.

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

WHITE, H. (1994), *Estimation, Inference and Specification Analysis.* Econometric Society Monograph, vol. 22 (Cambridge University Press).