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

Sandra Campo
University of North Carolina

Emmanuel Guerre
LSTA-University of Paris VI

Isabelle Perrigne
Pennsylvania State University

Quang Vuong
Pennsylvania State University

March 2006

* We are grateful to the Co-Editor and three referees for their comments and to Jingfeng Lu, Jim Powell and Mathew Shum for their suggestions. Previous versions were presented at the Malinvaud seminar, University of Cambridge, University of 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. This paper is based upon work supported by the National Science Foundation under Grant SES-0452154 awarded to the last two authors.
Correspondence to: I. Perrigne, Department of Economics, Pennsylvania State University, University Park, PA 16802. E-mail address : perrigne@psu.edu.

Abstract

This paper studies the identification and estimation of the first-price auction model with risk averse bidders within the private value paradigm. We show that the benchmark model is nonidentified in general from observed bids. We then consider various extensions including a binding reserve price, affiliation among private values and asymmetric bidders. In particular, we exploit heterogeneity across auctioned objects to establish semiparametric identification under a conditional quantile restriction and a parameterization of the bidders' utility function. Next we propose a semiparametric method for estimating the corresponding auction model. This method involves several steps and allows to recover the parameter(s) of the utility function as well as the bidders' private value distribution. We show that our semiparametric estimator of the utility function parameter(s) converges at the optimal rate, which is slower than the parametric one. The method is illustrated on U.S. Forest Service timber sales and a test of bidders' risk neutrality is performed.

Key words: Risk Aversion, Private Value, Nonparametric Identification, Semiparametric Estimation, Optimal Rate, Timber Auctions.

JEL classification: C14, D44, L70

Semiparametric Estimation of First-Price Auctions with Risk Averse Bidders

S. Campo, E. Guerre, I. Perrigne, and Q. Vuong

1 Introduction

Since the seminal unpublished work by Kenneth Arrow and its formalization by Pratt (1964), risk aversion has become a fundamental concept in economics whenever agents face various types of uncertainties such as in auctions. For instance, Goere, Holt and Palfrey (2002) maninly explain the deviations from the risk neutral Nash equilibrium by bidders' risk aversion, while Bajari and Hortacsu (2005) show that the risk aversion model provides the best fit to experimental bids over several competing models. On the other hand, many important results in auction theory crucially depend on risk neutrality such as the revenue equivalence theorem established by Vickrey (1961).[1] Despite its importance in auction modeling, very few empirical studies have assessed the extent of bidders' risk aversion on field data. Using US Forest Service auctions, Baldwin (1995) and Athey and Levin (2001) show that bidding diversification across species is consistent with bidders' risk aversion. The structural approach, which assumes that observed bids are the Bayesian Nash equilibrium outcomes, is well adapted to assess and test bidders' risk aversion.[2] Our paper adopts such an approach and focuses on the identification and estimation under nonparametric assumptions in the spirit of Laffont and Vuong (1996) and Guerre, Perrigne and Vuong (2000).

[1] The optimal auction mechanism under bidders' risk aversion requires complex transfers. See Maskin and Riley (1984) and Matthews (1987).

[2] See Paarsch (1992) and Laffont, Ossard and Vuong (1995) for early developments and Perrigne and Vuong (1999) and Athey and Haile (2006) for recent surveys.

Throughout, we consider first-price sealed-bid auctions with risk averse bidders within the private value paradigm. Under bidders' risk aversion, Maskin and Riley (1984) show that a first-price auction generates a larger revenue than an ascending auction, thereby providing a rationale for the use of the former mechanism. A first part of our paper briefly presents the benchmark model with independent private values (IPV) and reviews the existence, uniqueness and smoothness of the Bayesian Nash equilibrium strategy. In particular, existence and uniqueness follow from Maskin and Riley (2000b, 2003) among others. A second part is devoted to the identification of the benchmark model, i.e. whether its structural elements can be uniquely recovered from observed bids. The structural elements are the bidders' von Neuman Morgenstern (vNM) utility function and the bidders' private value distribution. First, we show that this model is nonidentified from observed bids even when the utility function is restricted to belong to well known families of risk aversion such as constant relative risk aversion (CRRA). Second, we show that any bid distribution can be rationalized by a CRRA model, a constant absolute risk aversion (CARA) model, and a fortiori a model with general risk aversion. Such a striking result implies that the game theoretical model does not impose testable restrictions on bids. 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 nonparametric. Again, we show that this model is not identified from observed bids.

In view of the preceding results, a third part of our paper seeks weak and palatable restrictions that can lead to identification. Specifically, we exploit heterogeneity across auctioned objects through a parametric quantile restriction of the private value distribution conditional upon object characteristics. We also restrict the bidders' vNM utility function to be parametric. Under these identifying conditions, we show that the utility function parameter(s) and the conditional private value distribution are semiparametrically identified. As a matter of fact, we show that dropping either one of these two conditions looses identification. In this sense, our semiparametric modeling is natural, while providing a new direction for the structural analysis of auction data. A fourth part of the paper provides an upper bound for the convergence rate, which can be attained by estimators of the utility function parameter(s). Specifically, we study the best (optimal) rate that an 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 a distribution upper boundary can be achieved at a faster rate than for any other quantile, we focus on a parametric restriction of 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 $N^{(R+1)/(2R+3)}$, which is independent of d though slower than $\sqrt{N}$.

A fifth part of the paper develops a multistep semiparametric estimation procedure. A first step consists in estimating nonparametrically the conditional bid density at its upper boundary. A second step uses (possibly weighted) nonlinear least squares (NLLS) with the nonparametric estimates obtained in the first step to estimate the utility function parameter(s). A third step allows us to recover the bidders' private values and their underlying 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 recently by Newey and McFadden (1994) and Powell (1994).[3] A notable feature of our estimation problem is that the variance of the error term in the nonlinear regression model diverges with N, thereby leading to a non standard convergence rate for our semiparametric estimator.

A sixth part of our paper studies extensions of the benchmark model. We show that our identification results extend to the case of 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 can be used to identify the model. Our estimation method can be readily adapted to the three first cases, while the 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 inversion of the equilibrium bidding strategy. This is especially convenient when there is no closed form solution to the differential equation defining the equilibrium strategy such as for general risk aversion and asymmetric bidders. We then illustrate our procedure on the US Forest Service timber auctions. In particular, a test of bidders' risk neutrality is performed and bidders are found to be risk averse.

The paper is organized as follows. Section 2 briefly presents the benchmark model

[3]Notable exceptions of semiparametric estimators converging at a slower rate than $\sqrt{N}$ are those proposed by Manski (1985), Horowitz (1992), Kyriazidou (1997) and Honoré and Kyriazidou (2000).

and the properties of the Bayesian Nash equilibrium strategy. Section 3 provides general nonidentification results of this model. Understanding of such results leads to additional identifying restrictions in Section 4. Section 5 provides an upper bound for the optimal convergence rate that can be attained by a semiparametric estimators of the utility function parameter(s). Section 6 presents our semiparametric estimation procedure with its various steps and statistical properties. Section 7 considers extensions of the benchmark model and discusses their identification and estimation. Section 8 is devoted to an illustration of our method to timber auction data. Section 9 concludes. Three appendices collect the proofs of our theoretical results.

2 The Benchmark Model

This section presents the IPV first-price sealed-bid auction model with risk averse bidders and properties of its equilibrium strategy. A single and indivisible object is sold through a first-price sealed-bid auction. Within the IPV paradigm, each bidder knows his own private value v_i for the auctioned object but not other bidders' private values. The private values are drawn independently from a distribution $F(\cdot)$, which is absolutely continuous with density $f(\cdot)$ on a support $[\underline{v}, \overline{v}] \subset I\!R_+$. The distribution $F(\cdot)$ and the number of potential bidders $I \geq 2$ are assumed to be common knowledge. 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 said to be symmetric. Bidder i then maximizes his expected utility

$$\mathrm{E}\Pi_i = U(v_i - b_i)\Pr(b_i \geq b_j, j \neq i) \tag{1}$$

with respect to his bid b_i, where b_j is the jth player's bid. This corresponds to the most studied case in the auction literature where the quality of the auctioned item is known and has equivalent monetary value. See Case 1 in Maskin and Riley (1984) and Krishna (2002).[4] In addition, 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.[5]

[4]Maskin and Riley (1984) consider a more general model where the utility of winning is of the form $u(-b_i, v_i)$ and the utility of loosing is equal to $w(\cdot)$. Here, $u(-b_i, v_i) = U(v_i - b_i)$ and $w(0) = U(0) = 0$.

[5]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

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.[6] Thus (1) becomes $\mathrm{E}\Pi_i = U(v_i - b_i)F^{I-1}(s^{-1}(b_i))$, 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

$$s'(v_i) = (I-1)\frac{f(v_i)}{F(v_i)}\lambda(v_i - b_i) \tag{2}$$

for all $v_i \in [\underline{v}, \overline{v}]$, where $\lambda(\cdot) = U(\cdot)/U'(\cdot)$. As shown by Maskin and Riley (1984), the boundary condition is $U(\underline{v} - s(\underline{v})) = 0$, i.e. $s(\underline{v}) = \underline{v}$ because $U(0) = 0$. Moreover, the second-order conditions are satisfied.

When the reserve price is nonbinding, existence of a pure equilibrium strategy follows from Maskin and Riley (2000b) and Athey (2001), while its uniqueness has been established by Maskin and Riley (2003) using an argument similar to Lebrun (1999). The main contribution of Theorem 1 is to derive the smoothness of the equilibrium strategy, which is used in the next sections. We assume that $U(\cdot)$ and $F(\cdot)$ belong to $\mathcal{U}_R$ and $\mathcal{F}_R$ defined as follows, respectively.

Definition 1: *For $R \geq 1$, let $\mathcal{U}_R$ be the set of utility functions $U(\cdot)$ satisfying*
(i) $U : [0, +\infty) \to [0, +\infty)$, $U(0) = 0$ and $U(1) = 1$,
(ii) $U(\cdot)$ is continuous on $[0, +\infty)$, and admits $R + 2$ continuous derivatives on $(0, +\infty)$ with $U'(\cdot) > 0$ and $U''(\cdot) \leq 0$ on $(0, +\infty)$,
(iii) $\lim_{x \downarrow 0} \lambda^{(r)}(x)$ is finite for $1 \leq r \leq R + 1$, where $\lambda^{(r)}(\cdot)$ is the rth derivative of $\lambda(\cdot)$.

Conditions (i) and (ii) have been discussed previously. Note that $\lim_{x \downarrow 0} \lambda(x) = 0$ since $U(0) = 0$ and $U'(\cdot)$ is nonincreasing. Thus, from (ii) and (iii) it follows that $\lambda(\cdot)$ admits $R + 1$ continuous derivatives on $[0, +\infty)$. These regularity assumptions are weak as they are satisfied by many vNM utility functions.

Definition 2: *For $R \geq 1$, let $\mathcal{F}_R$ be the set of distributions $F(\cdot)$ satisfying*
(i) $F(\cdot)$ is a c.d.f. with support of the form $[\underline{v}, \overline{v}]$, where $0 \leq \underline{v} < \overline{v} < +\infty$,

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 7, while the second case is beyond the scope of this paper. For multisignals, see Che and Gale (1998) for a model with budget constraints.

[6] Moreover, as noted by Maskin and Riley (1984, Remark 2.3), the only equilibria are symmetric when $F(\cdot)$ has bounded support, which is assumed below.

(ii) $F(\cdot)$ *admits* $R+1$ *continuous derivatives on* $[\underline{v}, \overline{v}]$,
(iii) $f(\cdot) > 0$ *on* $[\underline{v}, \overline{v}]$.

These restrictions are weak with the exception of the finite upper bound $\overline{v}$. Relaxing (i) is possible but would involve more technicalities, while allowing possibly asymmetric equilibria. Altogether (i)–(iii) imply that $f(\cdot)$ is bounded away from zero on $[\underline{v}, \overline{v}]$.

Theorem 1: *Let* $I \geq 2$ *and* $R \geq 1$. *Suppose that* $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$, *then there exists a unique (symmetric) equilibrium and its equilibrium strategy* $s(\cdot)$ *satisfies:*
(i) $\forall v \in (\underline{v}, \overline{v}]$, $s(v) < v$, *while* $s(\underline{v}) = \underline{v}$,
(ii) $\forall v \in [\underline{v}, \overline{v}]$, $s'(v) > 0$ *with* $s'(\underline{v}) = (I-1)\lambda'(0)/[(I-1)\lambda'(0)+1] < 1$,
(iii) $s(\cdot)$ *admits* $R+1$ *continuous derivatives on* $[\underline{v}, \overline{v}]$.

Theorem 1 is an immediate consequence of a more general theorem allowing for exogenous variables in Guerre and Vuong (2006).

3 General Nonidentification Results

In this section we study identification of the structure $[U, F]$ from observables. We assume that the number I of bidders is observed, as in a first-price sealed-bid auction with a nonbinding reserve price. We also assume that the distribution $G(\cdot)$ of an equilibrium bid is known. Knowledge of $G(\cdot)$ from observed bids is an estimation problem, which is addressed in Section 6. Thus the identification problem reduces to whether the structure $[U, F]$ can be recovered uniquely from the knowledge of (I, G). A related issue is whether any bid distribution $G(\cdot)$ can be rationalized by a structure $[U, F]$ given I. Such a question relates to the possibility of testing the validity of the auction model under consideration.

Following Guerre, Perrigne and Vuong (2000), we express (2) using the distribution $G(\cdot)$ of an equilibrium bid. For every $b \in [\underline{b}, \overline{b}] = [\underline{v}, s(\overline{v})]$, we have $G(b) = F(s^{-1}(b)) = F(v)$ with density $g(b) = f(v)/s'(v)$, where $v = s^{-1}(b)$. Thus (2) can be written as

$$1 = (I-1)\frac{g(b_i)}{G(b_i)}\lambda(v_i - b_i). \tag{3}$$

Since $U(\cdot) \geq 0$ and $U''(\cdot) \leq 0$, we have $\lambda'(\cdot) = 1 - U(\cdot)U''(\cdot)/U'^2(\cdot) \geq 1$. Thus $\lambda(\cdot)$ is strictly increasing. Solving (3) for v_i gives

$$v_i = b_i + \lambda^{-1}\left(\frac{1}{I-1}\frac{G(b_i)}{g(b_i)}\right) \equiv \xi(b_i, U, G, I), \tag{4}$$

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 equilibrium bid distribution $G(\cdot)$ satisfies some regularity properties implied by the smoothness of $s(\cdot)$ given in Theorem 1 and the regularity assumptions on $[U, F]$.

Definition 3: *For $R \geq 1$, let $\mathcal{G}_R$ be the set of distributions $G(\cdot)$ satisfying*
(i) $G(\cdot)$ is a c.d.f. with support of the form $[\underline{b}, \overline{b}]$, where $0 \leq \underline{b} < \overline{b} < +\infty$,
(ii) $G(\cdot)$ admits $R+1$ continuous derivatives on $[\underline{b}, \overline{b}]$,
(iii) $g(\cdot) > 0$ on $[\underline{b}, \overline{b}]$,
(iv) $g(\cdot)$ admits $R+1$ continuous derivatives on $(\underline{b}, \overline{b}]$,
(v) $\lim_{b \downarrow \underline{b}} d^r[G(b)/g(b)]/db^r$ exists and is finite for $r = 1, \ldots, R+1$.

The regularity properties (i)–(iii) are similar to those of Definition 2 for $F(\cdot)$. They imply that $g(\cdot)$ is bounded away from zero on $[\underline{b}, \overline{b}]$ and $\lim_{b \downarrow \underline{b}} G(b)/g(b) = 0$ so that $\lim_{b \downarrow \underline{b}} \xi(b, U, G, I) = \underline{b}$. Properties (iv) and (v) are specific to the auction model. In particular, (iv) says that $g(\cdot)$ is smoother than $f(\cdot)$, extending a similar property noted by Guerre, Perrigne and Vuong (2000) for the risk neutral model. Combined with (iii) and (iv), (v) implies that $G(\cdot)/g(\cdot)$ is $R+1$ continuously differentiable on $[\underline{b}, \overline{b}]$.

The following lemma provides necessary and sufficient conditions for rationalizing a distribution of observed bids by an IPV auction model with risk aversion. Hereafter, we say that a distribution is *rationalized* by an auction model with risk aversion if there exists a structure $[U, F]$ whose equilibrium bid distribution is identical to the given distribution.

Lemma 1: *Let $I \geq 2$, $R \geq 1$, and $\mathbf{G}(\cdot)$ be the joint distribution of $(b_1, \ldots, b_I)$. There exists an IPV auction model with risk aversion $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$ that rationalizes $\mathbf{G}(\cdot)$ if and only if*
(i) $\mathbf{G}(b_1, \ldots, b_I) = \prod_{i=1}^{I} G(b_i)$, with $G(\cdot) \in \mathcal{G}_R$,
(ii) $\exists \lambda : I\!R_+ \to I\!R_+$ with $R+1$ continuous derivatives on $[0, +\infty)$, $\lambda(0) = 0$ and $\lambda'(\cdot) \geq 1$ such that $\xi'(\cdot) > 0$ on $[\underline{b}, \overline{b}]$, where $\xi(b, U, G, I) = b + \lambda^{-1}\left[G(b)/((I-1)g(b))\right]$.

Condition (i) is related to the IPV paradigm and requires that bids be i.i.d., where $G(\cdot)$ satisfies the regularity properties of Definition 3. Condition (ii) arises from $\xi(\cdot, U, G, I)$ being the inverse of the equilibrium strategy, which is strictly increasing.[7]

The next proposition shows that any distribution $G(\cdot) \in \mathcal{G}_R$ can be rationalized by an IPV auction model with a utility function displaying risk aversion. Specifically, for $F(\cdot) \in \mathcal{F}_R$, we consider CRRA structures $[U, F]$ with $U(x) = x^{1-c}$ for $0 \leq c < 1$, where we have imposed $U(0) = 0$ and $U(1) = 1$. Similarly, CARA structures are of the form $[U, F]$ with $U(x) = (1 - \exp(-ax))/(1 - \exp(-a))$ for $a > 0$.[8]

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

Proposition 1 is striking. First, it implies that the restriction (ii) in Lemma 1 for rationalizing a bid distribution with risk averse bidders is redundant. Specifically, our proof indicates that we can always find a function $\lambda(\cdot)$ corresponding to either a CRRA or CARA utility function so that (ii) is satisfied whenever $G(\cdot) \in \mathcal{G}_R$. Alternatively, the IPV auction model with general risk aversion does not impose any restriction on observed bids beyond their independence and the weak regularity conditions embodied in $\mathcal{G}_R$. Second, because a structure $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$ leads to a bid distribution $G(\cdot) \in \mathcal{G}_R$ by Lemma 1, Proposition 1 implies that there always exist a CARA structure and a CRRA structure that are observationally equivalent to $[U, F]$. In other words, the game theoretic auction model with arbitrary risk aversion does not provide enough restrictions on observed bids to discriminate it from a CRRA or a CARA model.[9] Third, because a risk neutral model is a special risk averse model, 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

[7] As shown in the proof of Lemma 1, if condition (ii) is satisfied, then $G(\cdot)$ is rationalized by the structure $[U, F]$, where $U(x) = \exp \int_1^x (1/\lambda(t))dt$ and $F(\cdot)$ is the distribution of $\xi(b, U, G, I)$ with $b \sim G(\cdot)$. Because $\lambda(x) \sim \lambda'(0)x$ in the neighborhood of 0, $\int_1^0 (1/\lambda(t))dt = -\infty$ so that $U(0) = 0$, as required.

[8] See Gollier (2001) for other families of vNM utility functions.

[9] Proposition 1 also implies that any auction model with some wealth $w > 0$, as defined in footnote 5, is observationally equivalent to some CRRA/CARA model with zero wealth. Note that $\lambda(\cdot)$ is independent of w with 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

relative to risk neutrality. As a matter of fact, Proposition 1 says that allowing for CRRA or CARA rationalizes any bid distribution in $\mathcal{G}_R$.

A model is a set of structures $[U, F]$. Hereafter, a structure $[U, F]$ is *nonidentified* 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. As suggested by Proposition 1, auction models with risk averse bidders are nonidentified.

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

As $\mathcal{U}_R$ and $\mathcal{F}_R$ do not impose any parametric restrictions, Proposition 2 shows that the auction model with risk averse bidders is nonparametrically nonidentified. This contrasts with Guerre, Perrigne and Vuong (2000) who show that the auction model with risk neutral bidders is nonparametrically identified. Thus the nonidentification of the general risk aversion model $\mathcal{U}_R \times \mathcal{F}_R$ arises from the unknown utility function $U(\cdot)$, which is restricted to the identity function under risk neutrality.

In view of Proposition 2, additional restrictions must be imposed to identify the model. Two natural identifying strategies are available. First, since little is known in practice about $U(\cdot)$, while $F(\cdot)$ is approximatively log-normal in some empirical studies, we can parameterize the private value distribution as $F(\cdot;\gamma)$ giving the semiparametric model $\mathcal{U}_R \times \mathcal{F}(\Gamma)$. Alternatively, following the previous discusion, we can parameterize the utility function as $U(\cdot;\theta)$ leading to the semiparametric model $\mathcal{U}(\Theta) \times \mathcal{F}_R$. The next proposition states that parameterizing $F(\cdot)$ is not sufficient to achieve identification.

Proposition 3: *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 nonidentified semiparametric model. Let $I = 2$ and $\mathcal{F}(\Gamma) =$

---

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 satisfies Definition 3 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 $2/5 < c < 1$ by Lemma 1-(iii). 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.

$\{F(v;\gamma) = v/\gamma, v \in [0,\gamma], \gamma \in \Gamma = (1,2]\}$. Any structure $[U,F]$ with $U(x) = x^{\gamma-1}$ then leads to a uniform bid distribution $G(\cdot)$ on $[0,1]$ as the solution to (2) 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, exploiting the monotonicity of the equilibrium strategy, (4) evaluated at the α-quantile b_α of $G(\cdot)$ gives $F^{-1}(\alpha;\gamma) - b_\alpha = \lambda^{-1}(\alpha/g(b_\alpha))$ for $\alpha \in [0,1]$. This equation does not contain enough information to identify both γ and $\lambda(\cdot)$.[11]

As noted above, a second identifying strategy is to consider the semiparametric model $\mathcal{U}(\Theta) \times \mathcal{F}_R$. We define the CARA model (with regularity R) as the set of structures $[U,F] \in \mathcal{U}^{CARA} \times \mathcal{F}_R$. The CRRA model is similarly defined as $\mathcal{U}^{CRRA} \times \mathcal{F}_R$. The next proposition shows that parameterizing $U(\cdot)$ is not sufficient to achieve identification.

Proposition 4: *Let $I \geq 2$ and $R \geq 1$. Any structure $[U,F]$ in $\mathcal{U}^{CARA} \times \mathcal{F}_R$ or $\mathcal{U}^{CRRA} \times \mathcal{F}_R$ is not identified.*

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

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

where $b \sim G(\cdot)$. Because the above function is strictly increasing in b when $c < \tilde{c} < 1$, then $G(\cdot)$ can also be rationalized by $[\tilde{U},\tilde{F}]$. 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 $\tilde{F}(\cdot)$ and $F(\cdot)$ to have the same known support. At $b = \overline{b}$ the above equation indicates that the support of $\tilde{F}(\cdot)$ must shrink, i.e. $\overline{\tilde{v}} < \overline{v}$, to compensate for the increase in the constant relative risk aversion parameter $\tilde{c} > c$. More generally, all the quantiles of $\tilde{F}(\cdot)$ are smaller than the corresponding ones for $F(\cdot)$ as $\tilde{v}_\alpha = [(\tilde{c}-c)/(1-c)]b_\alpha + [(1-\tilde{c})/(1-c)]v_\alpha$ with $b_\alpha < v_\alpha$ for $\alpha \in (0,1]$. Intuitively, an increase in risk aversion can be compensated by a shrinkage in the private value distribution.

[11]In contrast, if $F(\cdot)$ is known, this equation shows that $\lambda(\cdot)$ and hence $U(\cdot)$ are nonparametrically identified on $[0, \max_v(v - s(v))]$. This property is exploited in Lu and Perrigne (2005) who rely on ascending auction data to estimate $F(\cdot)$ and hence to identify $U(\cdot)$ nonparametrically from first-price sealed-bid auction data.

4 Semiparametric Identification

Propositions 3 and 4 show that parameterizing either $F(\cdot)$ or $U(\cdot)$ alone is not sufficient for identification. Additional identifying restrictions need to be imposed. In the first case, assuming that the upper bound $\overline{v}$ is known is not sufficient to identify $U(\cdot)$ nonparametrically. Thus, parameterizing $U(\cdot)$ would be the most natural additional restriction. This would lead to a full parametric model $\mathcal{U}(\Theta) \times \mathcal{F}(\Gamma)$. In the second case, as the example after Proposition 4 suggests, imposing a single quantile v_α to be the same though unknown across $F(\cdot)$ may help toward identification, while still providing flexibility about $F(\cdot)$. For instance, we may require that all $F(\cdot)$ have the same unknown median or upper bound. Such a quantile restriction is still insufficient as (4) leads to a single equation in the unknown quantile v_α and risk aversion parameter(s) θ.

Hereafter, we exploit heterogeneity across auctions, which arises from variations in observed characteristics Z and/or the number of bidders $I \in \mathcal{I}$. Thus we consider that private values are drawn from the conditional distribution $F(\cdot|Z, I)$, where Z can be discrete or continuous with values z in $\mathcal{Z} \subset I\!\!R^d$.[12] The support of $F(\cdot|z, I)$ is denoted $[\underline{v}(z, I), \overline{v}(z, I)]$, while $G(\cdot|z, I)$ is the corresponding equilibrium bid distribution defined on $[\underline{b}(z, I), \overline{b}(z, I)]$. The α-quantiles of $F(\cdot|z, I)$ and $G(\cdot|z, I)$ are denoted $v_\alpha(z, I)$ and $b_\alpha(z, I)$, respectively. Our identifying assumptions are as follows.

Assumption A1: *For $\mathcal{I}$ a subset of $\{2, 3, \ldots\}$ and $R \geq 1$,*
(i) $U(\cdot) = U(\cdot; \theta) \in \mathcal{U}_R$ for every $\theta \in \Theta \subset I\!\!R^p$,
(ii) $F(\cdot|\cdot, \cdot) \in \mathcal{F}_R(\mathcal{Z} \times \mathcal{I}) \equiv \{F(\cdot|\cdot, \cdot) : F(\cdot|z, I) \in \mathcal{F}_R, \forall (z, I) \in \mathcal{Z} \times \mathcal{I}\}$.
(iii) For some $\alpha \in (0, 1]$, $v_\alpha(z, I) = v_\alpha(z, I; \gamma)$ for all $(z, I) \in \mathcal{Z} \times \mathcal{I}$ and some $\gamma \in \Gamma \subset I\!\!R^q$,
(iv) The function $\phi_\alpha(z, I; \theta, \gamma) \equiv \lambda(v_\alpha(z, I; \gamma) - b_\alpha(z, I); \theta)$ for $(z, I) \in \mathcal{Z} \times \mathcal{I}$ determines uniquely $(\theta, \gamma) \in \Theta \times \Gamma$.

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 $F(\cdot|z, I)$ satisfies the regularity conditions of Definition 2 for every $(z, I) \in \mathcal{Z} \times \mathcal{I}$. Condition (iii) is a parametric conditional quantile specification on $v_\alpha(z, I)$. See Powell (1994). For instance, $v_\alpha(\cdot, \cdot; \gamma)$ can be

[12] Such a specification allows for unobserved heterogeneity across objects provided I is a sufficient statistic for such unobserved heterogeneity conditional upon Z. See Campo, Perrigne and Vuong (2003).

[13] If wealth w is unknown, then w is included in θ. See also footnote 5.

chosen to be a constant or a polynomial, where γ is the vector of unknown coefficients. If $\alpha = 1$, a parametric specification of the upper bound $\overline{v}(z, I)$ is considered. The case of a common support then corresponds to a constant specification of $\overline{v}(z, I)$. On the other hand, no assumption is made on the lower bound $\underline{v}(z, I)$, i.e. $\alpha = 0$, as it is nonparametrically identified from the boundary condition $\underline{v}(z, I) = \underline{b}(z, I)$. An alternative identifying assumption to (iii) would specify parametrically the difference $v_\alpha(z, I) - \underline{v}(z, I)$. This would be equivalent to imposing a restriction on the α-quantile as $\underline{v}(z, I)$ can be recovered from $\underline{b}(z, I)$. In particular when $\alpha = 1$, the latter corresponds to a parametric specification of the length $\overline{v}(z, I) - \underline{v}(z, I)$ of the support of $F(\cdot|z, I)$.

Condition (iv) is the crucial identifying condition. For instance, consider a CRRA model with a constant α-quantile, i.e. $v_\alpha(z, I; \gamma) = \gamma$. This gives

$$\lambda(\gamma - b_\alpha(z, I); \theta) = \frac{\gamma - b_\alpha(z, I)}{\theta} = \frac{\alpha}{(I-1)g(b_\alpha(z, I)|z, I)}$$

from (4) evaluated at the α-quantile, where $\theta = 1 - c$. This equation determines uniquely (θ, γ) as soon as there exist two different values for $b_\alpha(z, I)$. This is possible from a variation in Z and/or a variation in I. Similarly, a CARA model with a constant conditional quantile restriction is identified. More generally, condition (iv) implies the "order" condition Card $(\mathcal{Z} \times \mathcal{I}) \geq p + q$. Furthermore, it exploits variations in the shading $v_\alpha(z, I) - b_\alpha(z, I)$ across (z, I) arising from different effects of Z and I on v_α and b_α. For instance, when I is exogenous so that $F(\cdot|\cdot, I)$ does not depend on I, an increase in I reduces the shading.[14] Consequently, a shift between $F(\cdot|z, I)$ and $G(\cdot|z, I)$ is excluded. Altogether, Assumption A1 is not as restrictive since a parameterization of a single quantile is needed, while this specification can include a polynomial of sufficiently high degree to capture possible nonlinearities. In particular, $F(\cdot|\cdot, \cdot)$ is left almost entirely unspecified.

The next proposition establishes the semiparametric identification of the first-price auction model with risk averse bidders. It relies upon the key equation (4) evaluated at the α-quantile $b_\alpha(z, I)$, which gives

$$g(b_\alpha(z, I)|z, I) = \frac{1}{I-1} \frac{\alpha}{\lambda(v_\alpha(z, I; \gamma) - b_\alpha(z, I); \theta)}, \tag{5}$$

[14] This idea and more generally exclusion restrictions are exploited in Guerre, Perrigne and Vuong (2005) to identify nonparametrically $U(\cdot)$ and $F(\cdot|\cdot)$.

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

Proposition 5: *The semiparametric model defined as the set of structures $[U, F]$ satisfying Assumption A1 is identified.*

Proposition 3 provides a semiparametric identification result since $U(\cdot)$ is parametrically identified through θ while $F(\cdot|\cdot)$ is nonparametrically identified subject to a parametric conditional quantile restriction. It is worthnoting that dropping either one of these parameterizations would lead to a nonidentified model. For instance, assume that $v_\alpha(z, I)$ is left unspecified, while a parametric specification for $U(\cdot)$ is retained, leading to the semiparametric model composed of structures $[U, F]$ satisfying A1-(i,ii). Such a model would not be necessarily identified. An example is the CRRA model with $F(\cdot|\cdot,\cdot) \in \mathcal{F}_R(\mathcal{Z} \times \mathcal{I})$. The argument is similar to that given after Proposition 2, where $G(\cdot)$ and $g(\cdot)$ are replaced by $G(\cdot|\cdot,\cdot)$ and $g(\cdot|\cdot,\cdot)$, respectively. Hence, restricting the utility function to be parametric does not achieve identification, despite that $U(\cdot)$ does not vary with (Z, I). Likewise, suppose that the utility function is left unspecified while a parametric conditional quantile restriction is retained, leading to the semiparametric model composed of structures $[U, F]$ satisfying A1-(ii,iii) with $U(\cdot) \in \mathcal{U}_R$. This model is not necessarily identified. Specifically, let $[U, F]$ be such a structure and consider the structure $[\tilde{U}, \tilde{F}]$, where $\tilde{U}(x) = c_1[U(x/\delta)]^\delta$ if $0 \leq x < \delta^2$ and $\tilde{U}(x) = c_2 U(x + \delta(1-\delta))$ if $x \geq \delta^2$, where $0 < \delta < 1$, $c_1 = c_2[U(\delta)]^{1-\delta}$, and $c_2 = 1/U(1 + \delta(1-\delta))$.[15] Let $\tilde{F}(\cdot|z, I)$ be the distribution of $\tilde{\xi}(b|z, I) = b + \tilde{\lambda}^{-1}\left(G(b|z, I)/[(I-1)g(b|z, I)]\right)$, where $b \sim G(\cdot|z, I)$. It can be shown that $[\tilde{U}, \tilde{F}]$ rationalizes $G(\cdot|\cdot,\cdot)$ and that $\tilde{F}(\cdot|\cdot,\cdot)$ satisfies A1-(ii,iii). Hence, the parameterization of the conditional quantile of $F(\cdot|z, I)$ is not sufficient for identification.

5 Optimal Convergence Rate

Given the nonstandard nature of our model, it is especially useful to derive the optimal (best) convergence rate that can be attained by semiparametric estimators of the risk averison parameter(s) θ. This is the purpose of this section. The optimal convergence rate

[15] Note that $\tilde{U}(0) = 0$, and $\tilde{U}(\cdot)$ has $R+2$ continuous derivatives on $(0, \delta) \cup (\delta, +\infty)$. Thus $\tilde{U}(\cdot) \in \mathcal{U}_R$ if $\tilde{U}(\cdot)$ has $R+2$ continuous derivatives at $x = \delta^2$. Hence $\tilde{U}(\cdot)$ should be smoothed out in the neighborhood of $x = \delta^2$ to be $R+2$ continuously differentiable on $(0, +\infty)$. We omit this smoothing requirement.

for estimating the conditional density $f(\cdot|\cdot,\cdot)$ will follow from Guerre, Perrigne and Vuong (2000). We first need to strengthen our regularity assumptions on $F(\cdot|\cdot,\cdot)$ and $U(\cdot;\cdot)$ with respect to (z,I) and θ. Regarding $F(\cdot|\cdot,\cdot)$, we introduce the following definition, which parallels Definition 2 taking into account the conditioning variables (Z,I).

Definition 4: *For $R \geq 1$ and some unknown $\underline{v}$ and $\overline{v}$, $0 \leq \underline{v} < \overline{v} < +\infty$, let $\mathcal{F}_R^* \equiv \mathcal{F}_R^*(\mathcal{Z} \times \mathcal{I})$ be the set of conditional distributions $F(\cdot|\cdot,\cdot)$ satisfying*
(i) $\forall (z,I) \in \mathcal{Z} \times \mathcal{I}$, $\underline{v}(z,I) = \underline{v}$ and $\overline{v}(z,I) = \overline{v}$,
(ii) $\forall I \in \mathcal{I}$, $F(\cdot|\cdot,I)$ admits $R+1$ continuous derivatives on $[\underline{v},\overline{v}] \times \mathcal{Z}$,
(iii) $\forall I \in \mathcal{I}$, $\inf_{(v,z)\in[\underline{v},\overline{v}]\times\mathcal{Z}} f(v|z,I) > 0$.

While conditions (ii) and (iii) are straightforward extensions of Definition 2-(ii)-(iii), condition (i) needs further discussion. Assuming a constant lower boundary $\underline{v}(z,I) = \underline{v}$ simplifies the proof of the smothness of $s(\cdot,\cdot,I)$ in Guerre and Vuong (2006), which is needed to obtain the smoothness of the bid distribution. Such a restriction, however, is not used in estimation as $\underline{v}(z,I)$ can be recovered from $\underline{b}(z,I)$. Regarding the upper boundary restriction, our estimation procedure relies on (5), which requires an estimate for $b_\alpha(z,I)$. There is an important difference between estimating a quantile with $\alpha \in (0,1)$ and estimating the upper boundary ($\alpha = 1$). In particular, the convergence rate for estimating the latter is faster than for estimating the former. This suggests that the optimal convergence rate for estimating θ cannot be faster when considering an α-quantile restriction with $\alpha \in (0,1)$ than when considering the upper boundary. Hereafter, we thus focus on $\alpha = 1$, and for sake of simplicity, we consider a constant upper boundary so that $q = 1$. In other words, we assume a common but unknown support for the distributions $F(\cdot|z,I)$, where $(z,I) \in \mathcal{Z} \times \mathcal{I}$. Such an assumption is implicitely made in parametric estimation.

We next derive the smoothness properties of $G(\cdot|\cdot,\cdot)$ corresponding to a structure $[U,F]$ in $\mathcal{U}_R \times \mathcal{F}_R^*$ and hence of the implied statistical model for the observables. Lemma 2 extends Lemma 1-(i) to the case with exogenous variables (Z,I) and is proved in Guerre and Vuong (2006).

Lemma 2: *Let $\mathcal{I} \subset \{2,3,\ldots\}$, $R \geq 1$ and $\mathcal{Z}$ be a rectangular compact of $I\!R^d$ with nonempty interior.*[16] *For every $I \in \mathcal{I}$, the conditional distribution $G(\cdot|\cdot,I)$ corresponding*

[16] To simplify, discrete exogenous variables are excluded. If not, our next results continue to hold with suitable modifications.

to a structure $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R^*$ *satisfies*
(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* $\underline{b}(z, I) = \underline{v}$*,*
(ii) $G(\cdot|\cdot, I)$ *admits* $R + 1$ *continuous partial derivatives on* $S_I(G) \equiv \{(b, z); z \in \mathcal{Z}, b \in [\underline{b}(z, I), \overline{b}(z, I)]\}$*,*
(iii) $g(b|z, I) > c_g > 0$ *for all* $(b, z) \in S_I(G)$*,*
(iv) $g(\cdot|\cdot, I)$ *admits* $R + 1$ *continuous partial derivatives on* $S_I^u(G) \equiv \{(b, z); z \in \mathcal{Z}, b \in (\underline{b}(z, I), \overline{b}(z, I)]\}$*,*
(v) $\lim_{b\downarrow\underline{b}(z,I)} \partial^r[G(b|z, I)/g(b|z, I)]/\partial b^r$ *exists and is finite for* $r = 1, \ldots, R+1$ *and* $z \in \mathcal{Z}$*.*

We then consider the semiparametric model composed of structures $[U, F]$ satisfying:

Assumption A2: *Let* $\mathcal{I} \subset \{2, 3, \ldots\}$*,* $R \geq 1$ *and* $\mathcal{Z}$ *be a rectangular compact of* $\mathbb{R}^d$ *with nonempty interior.*
(i) In addition to A1-(i), $U(\cdot; \cdot)$ *is* $R + 2$ *continuously differentiable on* $(0, +\infty) \times \Theta$*,*
(ii) $F(\cdot|\cdot, \cdot) \in \mathcal{F}_R^*$*,*
(iii) The function $\phi_1(z, I; \theta, \overline{v}) \equiv \lambda(\overline{v} - \overline{b}(z, I); \theta)$ *for* $(z, I) \in \mathcal{Z} \times \mathcal{I}$ *determines uniquely* $(\theta, \overline{v}) \in \Theta \times (0, +\infty)$*.*

Conditions (i) and (ii) strengthen A1-(i,ii,iii). Condition (iii) simply expresses A1-(iv) at the upper boundary under a constant restriction. Thus (5) becomes

$$g(\overline{b}(z, I)|z, I) = \frac{1}{I - 1} \frac{1}{\lambda(\overline{v} - \overline{b}(z, I); \theta)}, \tag{6}$$

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.[17]

Assumption A3:
(i) The variables (Z_ℓ, I_ℓ)*,* $\ell = 1, 2 \ldots$ *are i.i.d. with support* $\mathcal{Z} \times \mathcal{I}$*, where* $\mathcal{I}$ *is finite, and with density* $f_{ZI}(\cdot, \cdot)$ *satisfying* $0 < \inf_{(z,I)\in\mathcal{Z}\times\mathcal{I}} f_{ZI}(z, I) \leq \sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}} f_{ZI}(z, I) < +\infty$*,*

[17]Assumption 3-(i) can be weakened allowing Z_ℓs not to be i.i.d. distributed as Theorem 3 is derived conditionally upon $(Z_1, I_1, \ldots, Z_\ell, I_\ell)$.

(ii) For every ℓ, the private values $V_{i\ell}, i = 1, \ldots, I_\ell$ are i.i.d. conditionally upon (Z_ℓ, I_ℓ) as $F_0(\cdot|Z_\ell, I_\ell)$,
(iii) The semiparametric model is correctly specified, i.e. the true utility function $U_0(\cdot)$ and conditional distribution $F_0(\cdot|\cdot,\cdot)$ satisfy A2 for some $\theta_0 \in \Theta$ and $0 \leq \underline{v}_0 < \overline{v}_0 < +\infty$.

As in Horowitz (1993), 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 following norms

$$||\beta||_\infty = \max(|\theta_1|, \ldots, |\theta_p|, |\overline{v}|), \quad ||f(\cdot|\cdot,\cdot)||_\infty = \sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}} \sup_{v\in[\underline{v},\overline{v}]} |f(v|z,I)|$$

and define the set of conditional densities

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

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

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

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}) with f_{ZI} the joint density of the $(Z_\ell, I_\ell)s$. 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*

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

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 nonparametric 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 6 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 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 $g_N(b|z, I) = g_0(b|z, I) + [m(z, I; \beta_N) - m(z, I; \beta_0)]\, \psi\left(\kappa\sqrt{\rho_N}(b - \overline{b}_0(z, I))\right)$, where $\psi : I\!R^- \rightarrow I\!R$ is compactly supported with $\psi(0) = 1$, and $\int \psi(x)dx = 0$, while

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

$\kappa > 0$, and $\|\beta_N - \beta_0\|_\infty = O(1/\rho_N)$. Using Lemmas 1 and 2, we first establish that each such density can be rationalized by an auction model with $(\beta_N, f_N(\cdot|\cdot, \cdot)) \in \mathcal{V}_\epsilon(\beta_0, f_0)$ for ρ_N sufficiently large. We then show that the probability distributions of the $B_{i\ell}$s under $g_N(\cdot|\cdot, \cdot)$ and $g_0(\cdot|\cdot, \cdot)$ cannot be distinguished statistically from each other.

6 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 $f(\cdot|\cdot, \cdot)$. Because $f(\cdot|\cdot, \cdot)$ is not parameterized, the estimation problem is semiparametric. A first subsection presents the different steps of our semiparametric procedure, while a second subsection establishes the asymptotic properties of our estimator of θ.

6.1 A Semiparametric Procedure

Our semiparametric procedure relies on the identifying relation obtained from (6) and (7)

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

where the subscript 0 indicates the truth. If one knew the upper boundary $\overline{b}_0(\cdot, \cdot)$ and the density $g_0(\cdot|\cdot, \cdot)$, one could recover $\beta_0 = (\theta_0, \overline{v}_0)$ from (8) given the parametric form for $\lambda(\cdot;\cdot)$. From the knowledge of $G_0(\cdot|\cdot,\cdot)$ and θ_0, one could then recover bidders' private values v_i from (4) to estimate $f_0(\cdot|\cdot,\cdot)$. This suggests the following three steps procedure:

- Step 1: From observed bids, estimate nonparametrically $\overline{b}_0(\cdot,\cdot)$ and $g_0(\overline{b}_0(\cdot,\cdot)|\cdot,\cdot)$ at the observed values (Z_ℓ, I_ℓ),

- Step 2: Using (8), where $g_0(\overline{b}_0(Z_\ell,I_\ell)|Z_\ell,I_\ell)$ 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 (4), where $G_0(\cdot|\cdot,\cdot)$, $g_0(\cdot|\cdot,\cdot)$ and $\lambda(\cdot;\theta_0)$ are replaced by their nonparametric estimators and $\lambda(\cdot;\hat{\theta}_N)$, recover the pseudo private values $\hat{v}_i$ to estimate nonparametrically $f_0(\cdot|\cdot,\cdot)$.

Nonparametric Boundary Estimation

This step consists in estimating the upper boundary $\overline{b}_0(\cdot,\cdot)$ and the conditional density $g_0(\cdot|\cdot,\cdot)$ at the upper boundary. We first discuss the estimation of $\overline{b}_0(\cdot,\cdot)$. Fix $I \in \mathcal{I}$. By Lemma 2-(i), $\overline{b}_0(\cdot, I)$ is $R+1$ continuously differentiable on $\mathcal{Z}$. Following Korostelev and Tsybakov (1993), one introduces 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 $z \in I\!R^d$ 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 $d=1$, 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

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

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

Turning to the estimation of $g(\cdot|\cdot,\cdot)$, it is well-known that standard kernel density estimators suffer from bias at boundary points. 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

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

where h_N is a bandwidth. Lemma B3 shows that $Y_{i\ell}$ is an asymptotically unbiased estimator of $g_0(\overline{b}_0(Z_\ell, I_\ell)|Z_\ell, I_\ell)$ given (Z_ℓ, I_ℓ) as h_N vanishes.[18] 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 $\hat{\overline{b}}_N(\cdot,\cdot)$.

Semiparametric Estimation of β_0

Let $\mathcal{F}_L$ be the σ-field generated by $Z_\ell, \ell = 1,\dots,L$. In view of (8)-(9) we consider

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

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

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

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. The latter feature raises some technical difficulties when deriving the asymptotic properties of $\hat{\beta}_N$. In particular, the diverging variance is the main reason why our estimator does not achieve the parametric rate. Specifically, its rate $N^{(R+1)/(2R+3)}$ is smaller than $\sqrt{N}$ but is optimal in the minimax sense.

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 $\overline{v}_{\sup} > 0$. The set $\mathcal{B}_\delta$ is introduced to bound $m(\cdot,\cdot;\beta)$ away from 0. Because $\overline{b}_0(\cdot,\cdot)$ in $m(\cdot,\cdot;\beta)$ 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

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

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

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

Nonparametric Estimation of $f(\cdot|\cdot)$

This step is similar to the second step in Guerre, Perrigne and Vuong (2000) with the difference that $\lambda(\cdot;\theta_0)$ in (4) 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 $G_0(\cdot|\cdot,\cdot)/g_0(\cdot|\cdot,\cdot)$ evaluated at $(B_{i\ell},Z_\ell,I_\ell)$. For an arbitrary (b,z,I), the ratio $G_0(b|z,I)/g_0(b|z,I)$ is estimated by

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

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 $(B_{i\ell},Z_\ell)+\mathcal{S}(2h_G)\subset\hat{\mathcal{S}}(G_{I_\ell})$ and $(B_{i\ell},Z_\ell)+\mathcal{S}(2h_g)\subset\hat{\mathcal{S}}(G_{I_\ell})$. 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 $\hat{\mathcal{S}}_I(G)$ is the estimated support of the conditional bid distribution $G_0(\cdot|\cdot,I)$. Specifically, $\hat{\mathcal{S}}_I(G)=\{(b,z):b\in[\hat{\underline{b}}_N(z,I),\hat{\overline{b}}_N(z,I)],z\in\mathcal{Z}\}$, 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 used in a standard kernel estimation of $f_0(\cdot|\cdot,\cdot)$. Namely, for an arbitrary pair (v,z,I), $f(v|z,I)$ is estimated by

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

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 $(L/\log L)^{1/(2R+d+2)}$. Because $\hat{\theta}_N$ converges at a faster rate, it follows from Guerre, Perrigne and Vuong (2000) that $\hat{f}(\cdot|\cdot)$ is uniformly consistent on compact subsets of its support at the rate $(N/\log N)^{R/(2R+d+3)}$, which is optimal for estimating $f_0(\cdot|\cdot)$ from observed bids.

6.2 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 $I\!R^p$, and

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

(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 $I\!R_-$ with support $[-1,0]$ and satisfies $\int \Phi(x)dx = 1$, $\int x^j \Phi(x)dx = 0$ for $j = 1,\ldots,R$,

(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 $a_N = o\left(\min\left\{h_N^{R+2}, \sqrt{h_N/N}\right\}\right)$.

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 2-(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

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

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

are full rank matrices in a neighborhood of β_0. Though our kernel $\Phi(\cdot)$ is one-sided, A4-(iii,iv) are standard in kernel estimation when using higher order kernels. 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 $N^{-1/(2R+3)}$, it is easily checked that $R \geq d$ 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 (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

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

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* $\hat{\beta}_N - \beta_0 = O_P\left(h_N^{R+1} + 1/\sqrt{N h_N}\right)$*, 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 $\lim_{N\to\infty} \sqrt{N h_N} h_N^{R+1} = \infty$*, then* $(1/h_N^{R+1})\left(\hat{\beta}_N - \beta_0\right) \xrightarrow{P} A(\beta_0)^{-1} \mathbf{b}(\beta_0, g_0)$.
(iii) If $\lim_{N\to\infty} \sqrt{N h_N} h_N^{R+1} = c \geq 0$*, then*

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

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

On technical grounds, the proof of Theorem 3-(i) is complicated by the divergence of the error variance (11) in the nonlinear model (10). In particular, omitting the estimation of the upper boudary $\overline{b}(\cdot, \cdot)$, 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 $h_N Q_N(\beta)/N$, 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

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

This expansion corresponds to the usual bias/variance decomposition of nonparametric estimators (see e.g. Härdle and Linton (1994)). When $Nh_N^{2R+3} \to 0$, the leading term is the second term, and we obtain

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

When $Nh_N^{2R+3} \to \infty$, it 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 $\hat{\beta}_N - \beta_0 = O_P(N^{-(R+1)/(2R+3)})$.[19]

The best convergence rate $N^{(R+1)/(2R+3)}$ of $\hat{\beta}_N$ corresponds to the optimal rate for estimating an univariate density with $R+1$ bounded derivatives. Moreover, it is independent of the dimension d of Z and hence avoids the curse of dimensionality encountered in nonparametric estimation. This seems surprising in view of (8), which suggests that β_0 is as difficult to estimate as the conditional density $g_0(\cdot|\cdot,\cdot)$, while the latter cannot be estimated faster than $N^{(R+1)/(2R+3+d)}$ from Stone (1982) given the $(R+1)$ bounded derivatives of $g_0(\cdot|\cdot, I)$. The faster rate $N^{(R+1)/(2R+3)}$ can be explained by noting that (8) leads to the moment conditions $\mathrm{E}[\{g_0(\overline{b}_0(Z,I)|Z,I) - m(Z,I;\beta_0)\}W(Z,I)] = 0$ for any vector function $W(\cdot)$. Integrating with respect to Z intuitively improves the rate of convergence by eliminating the Z dimension. This is similar to Newey and McFadden (1994) though Assumptions (iii)-(iv) of their Theorem 8.1 are not satisfied here. In fact, because the variance (11) is diverging, our proof shows that the average gradient

[19] 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). Another bias correction using a modification of the $Y_{i\ell}$s could be based on Hengartner (1997). From Liu and Brown (1993), however, such a bias correction cannot hold in the minimax sense of Theorem 2. Because the limit results used in the proof hold uniformly with respect to (β, f) in a neighborhood of (β_0, f_0), $\hat{\beta}_N$ is rate efficient in the sense of Theorem 2.

$(1/N)\partial\hat{Q}_N(\beta_0)/\partial\beta = O_{\rm P}(h^{R+1} + 1/\sqrt{Nh_N})$, which is different from the usual $O_{\rm P}(1/\sqrt{N})$. Hence, $\hat{\beta}_N$ converges at a slower rate than $\sqrt{N}$.

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. $(\int \Phi^2(x)dx/h_N)\hat{A}_N^{-1}(\hat{\beta}_N)\hat{B}_N(\hat{\beta}_N)\hat{A}_N^{-1}(\hat{\beta}_N)$. Note that $\hat{\beta}_N$ depends on $\omega(\cdot,\cdot)$, which can be chosen optimally as in weighted NLLS. From (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 $(\int \Phi^2(x)dx/h_N)\hat{A}_N^{-1}(\hat{\beta}_N^*)$. This is the best variance achievable in the regression model (10) with $e_{i\ell} = 0$.

7 Extensions

The benchmark model highlights the identification issues arising from bidders' risk aversion in first-price sealed-bid auctions. A number of extensions can be useful in practice such as a binding reserve price, affiliated private values and asymmetries among bidders.[20]

7.1 Reserve Price

An announced binding reserve price $p_0 \in (\underline{v}, \overline{v})$ acts as a screening device to bidders' participation.[21] Let $G^*(\cdot)$ the observed bid distribution and I^* the observed number of active bidders with $I^* \leq I$. From the boundary condition $s(p_0) = p_0$, we have $G^*(b^*) = (F(v) - F(p_0))/(1 - F(p_0))$ with $b^* = s(v)$. Hence, similarly to (4), (2) implies

$$v_i = b_i^* + \lambda^{-1}\left(\frac{1}{I-1}\frac{G^*(b_i^*)}{g^*(b_i^*)} + \frac{1}{I-1}\frac{1}{g^*(b_i^*)}\frac{F(p_0)}{1-F(p_0)}\right) \equiv \xi(b_i^*, G^*, I, F(p_0)), \qquad (20)$$

for $i = 1, \ldots, I_*$. In contrast to (4), I and $F(p_0)$ are unknown. Definitions 1, 2 and 3 remain the same with the exception that p_0 replaces $\underline{b}$ in Definition 3 and $\lim_{b\downarrow p_0} g^*(b) = +\infty$. Given that I^* is Binomial distributed with parameters $[I, 1 - F(p_0)]$, the observed bid distribution $\mathbf{G}^*(\cdot,\ldots,\cdot)$ conditionally on I^* is rationalized if and only if Lemma 1

[20]For every case, we assume that Theorem 1 extends.

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

is satisfied with $\mathcal{G}_R$ and $\xi(\cdot)$ as defined above. This result is shown using Lemma 1 and Guerre, Perrigne and Vuong (2000, Theorem 4). Note that I and $F(p_0)$ are identified from the distribution of I^*.

From this rationalization result, any $G^*(\cdot) \in \mathcal{G}_R$ can be rationalized by a CRRA or CARA structure with $F(\cdot) \in \mathcal{F}_R$. Hence, any structure $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$ is not identified. Moreover, parameterizing the utility function such as CRRA or CARA is not sufficient to achieve identification. The proofs of these results follow those of Propositions 1, 2 and 4, where $G(\cdot)/g(\cdot)$ is replaced by $G^*(\cdot)/g^*(\cdot) + F(p_0)/[(1-F(p_0))g^*(\cdot)]$ in view of (4) and (20).[22] As for a nonbinding reserve price, an increase in the aversion parameter 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 and a binding reserve price, the key identifying equation (5) becomes

$$g^*(b_\alpha(z,I)|z,I) = \frac{1}{I-1}\frac{\alpha + \frac{F(p_0|z,I)}{1-F(p_0|z,I)}}{\lambda(v_\alpha(z,I;\gamma) - b_\alpha(z,I);\theta)}. \tag{21}$$

Under A1, θ is semiparametrically identified.

Estimation is performed by adapting the procedure in Section 6.1 following Guerre, Perrigne and Vuong (2000, Section 4.2). In particular, the first step requires estimators of I and $F(p_0|z,I)$. Assuming that I is constant across auctions, a natural estimator for I is $\hat{I} = \max_\ell I^*_\ell$, while an estimator for $F(p_0|z)$ is obtained from $\mathrm{E}(I^*|z) = I[1 - F(p_0|z)]$. In the second step, the weighted NLLS is based on (21), where I and $F(p_0|z)$ are replaced by their estimates. The consistency rate is optimal and as before, namely $N^{(R+1)/(2R+3)}$, which is independent of the dimension of Z.

7.2 Affiliated Private Values

Following Milgrom and Weber (1982), the private values $(v_1, \ldots, v_I)$ are distributed as $\mathbf{F}(\cdot, \ldots, \cdot)$, which is exchangeable and affiliated. Bidder i's expected profit is $U(v_i - b_i)G_{B_i|b_i}(b_i|b_i)$, where $B_i = \max_{j\neq i} b_j$, $b_i = s(v_i)$ and $G_{B_i|b_i}(b_i|b_i)$ is the probability that $b_i \geq B_i$ conditional on b_i. By symmetry, $G_{B_i|b_i}(\cdot|\cdot) = G_{B|b}(\cdot|\cdot)$, for $i = 1, \ldots, I$, while (4)

[22] Note that Proposition 3 still holds as the example of a nonidentified semiparametric model $\mathcal{U}_R \times \mathcal{F}(\Gamma)$ can be adapted with a truncation on the bid distribution at the reserve price.

becomes

$$v_i = b_i + \lambda^{-1}\left(\frac{G_{B|b}(b_i|b_i)}{g_{B|b}(b_i|b_i)}\right) \equiv \xi(b_i, U, \mathbf{G}). \tag{22}$$

Definitions 1 and 2 remain the same except that $\mathbf{F}(\cdot,\ldots,\cdot)$ is $R+I$ continuously differentiable following Li, Perrigne and Vuong (2000, 2002). Note that $G_{B|b}(\cdot|\cdot)/g_{B|b}(\cdot|\cdot) = G_{B\times b}(\cdot,\cdot)/g_{Bb}(\cdot,\cdot)$, 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 $\mathbf{G}(\cdot,\ldots,\cdot)$ with R continuously differentiable densities such that $G_{B\times b}(b,b)/g_{Bb}(b,b)$ is $R+1$ continuously differentiable in $b \in [\underline{b}, \overline{b}]$ and strictly positive on $(\underline{b}, \overline{b}]$. The observed bid distribution $\mathbf{G}(\cdot,\ldots,\cdot)$ is then rationalized if and only if Lemma 1 is satisfied with $\mathcal{G}_R$ and $\xi(\cdot)$ as defined above. This result is shown using Lemma 1 and Li, Perrigne and Vuong (2002, Proposition 1).

From this rationalization result, any $\mathbf{G}(\cdot,\ldots,\cdot) \in \mathcal{G}_R$ can be rationalized by a CRRA or CARA structure with $\mathbf{F}(\cdot,\ldots,\cdot) \in \mathcal{F}_R$. Hence, any structure $[U,F] \in \mathcal{U}_R \times \mathcal{F}_R$ is not identified. Moreover, parameterizing the utility function such as CRRA or CARA is not sufficient to achieve identification. The proofs of these results follow those of Propositions 1, 2 and 4, where $G(\cdot)/[(I-1)g(\cdot)]$ is replaced by $G_{B\times b}(\cdot,\cdot)/g_{Bb}(\cdot,\cdot)$ in view of (4) and (22). With auction characteristics and affiliation and under parameterization of $U(\cdot)$, the key identifying equation (5) becomes

$$g_{Bb}(b_\alpha(z,I), b_\alpha(z,I)|z,I) = \frac{G_{B\times b}(b_\alpha(z,I), b_\alpha(z,I)|z,I)}{\lambda(v_\alpha(z,I;\gamma) - b_\alpha(z,I);\theta)}, \tag{23}$$

where $b_\alpha(z,I)$ is an α-quantile of the marginal distribution $G(\cdot|z,I)$ as all bidders are symmetric. Under A1, the model $\mathcal{U}(\Theta) \times \mathcal{F}_R(\mathcal{Z} \times \mathcal{I})$ is semiparametrically identified.

Estimation is performed by adapting the procedure in Section 6.1 following Li, Perrigne and Vuong (2002). In particular, $G_{B\times b}(b,b,z,I)$ is estimated as the product of an indicator function for the first term and kernels for the second and third arguments, while $g_{Bb}(b,b,z,I)$ is estimated using a standard kernel density estimator. In the second step, the weighted NLLS is based on (23). Because the bid distribution and density include an additional dimension, the optimal rate for estimating θ will be slower than in Theorem 2.

7.3 Asymmetry

Asymmetry among bidders can arise from (i) different private value distributions or (ii) different utility functions. The latter known as heterogeneous preferences in the literature includes different attitudes toward risk and/or different wealth levels.[23] The major difficulty with asymmetric auctions is that the equilibrium strategies are the solutions of an intractable system of I differential equations. See Maskin and Riley (2000).

Asymmetry in Private Values

The joint distribution $\mathbf{F}(\cdot, \ldots, \cdot)$ is equal to $\prod_i F_i(\cdot)$ with each $F_i(\cdot)$ defined on $[\underline{v}, \overline{v}]$ and satisfying Definition 2. Let $\mathcal{F}_R^I$ be the set of such distributions. Because of the boundary conditions $s_i(\underline{v}) = \underline{v}$ and $s_i(\overline{v}) = s_j(\overline{v})$, bidder i's distribution $G_i(\cdot)$ is defined on $[\underline{b}, \overline{b}]$ for all $i = 1, \ldots, I$. Following Campo, Perrigne and Vuong (2003), we have

$$v_i = b_i + \lambda^{-1}\left(\frac{1}{H_i(b_i)}\right) \equiv \xi_i(b_i, U, \mathbf{G}), \text{ where } H_i(\cdot) = \sum_{j \neq i} \frac{g_j(\cdot)}{G_j(\cdot)}, \tag{24}$$

for $i = 1, \ldots, I$. Let $\mathcal{G}_R^I$ be the set of distributions $\mathbf{G}(\cdot, \ldots, \cdot)$ such that each marginal distribution $G_i(\cdot)$ satisfies Definition 3 with $G(b)/g(b)$ replaced by $1/H_i(b)$ in (v). The bid distribution $\mathbf{G}(\cdot, \ldots, \cdot)$ is then rationalized if and only if Lemma 1 is satisfied with $\mathcal{G}_R^I$ and $\xi_i(\cdot), i = 1, \ldots, I$ as defined above. The proof is similar to that of Lemma 1.

Hence, any $\mathbf{G}(\cdot, \ldots, \cdot) \in \mathcal{G}_R^I$ can be rationalized by a CRRA or CARA structure with $\mathbf{F}(\cdot, \ldots, \cdot) \in \mathcal{F}_R^I$. Thus any structure $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R^I$ is not identified. Moreover, parameterizing $U(\cdot)$ such as CRRA or CARA is not sufficient to achieve identification. The proofs of these results follow those of Propositions 1, 2 and 4, where $G(\cdot)/[(I-1)g(\cdot)]$ is replaced by $1/H_i(\cdot)$ in view of (4) and (24). With auction characteristics, bidders' asymmetry and parameterization of $U(\cdot)$, the key identifying equation (5) becomes

$$\sum_{j \neq i} \frac{g_j(b_{i\alpha}(z, I)|z, I)}{G_j(b_{i\alpha}(z, I)|z, I)} = \frac{1}{\lambda(v_{i\alpha}(z, I; \gamma_i) - b_{i\alpha}(z, I); \theta)}, \tag{25}$$

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

Comparing (6) and (25) at $\alpha = 1$ shows that $\epsilon_{i\ell}$ in (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 GNLLS estimator is more appropriate. To simplify, we assume that the same I bidders are in the L auctions. Let $\hat{H}_\ell = (\hat{H}_{1\ell}, \ldots, \hat{H}_{I\ell})'$ and $M_\ell(\beta) = (m(Z_\ell; \beta_1), \ldots, m(Z_\ell; \beta_I))'$ with $m(Z_\ell; \beta) = 1/\lambda(\overline{v} - \overline{b}(Z_\ell); \theta)$, $\beta = (\overline{v}, \theta)'$. 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

[23] See also footnote 5, where $U(v_i - b_i)$ is replaced by $U_i(v_i - b_i) \equiv U(v_i - b_i + w_i) - U(w_i)$.

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 $D_\ell = \text{diag}[R^{-1}M_\ell(\beta)]$. The resulting two-step estimator of θ converges at the optimal rate, namely $N^{(R+1)/(2R+3)}$.

Asymmetry in Preferences

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

$$v_i = b_i + \lambda_i^{-1}\left(\frac{1}{H_i(b_i)}\right) \equiv \xi_i(b_i, U_i, \mathbf{G}), \tag{26}$$

where $\lambda_i(\cdot) = U_i(\cdot)/U_i'(\cdot)$ and $H_i(\cdot) = \sum_{j \neq i} g_j(\cdot)/G_j(\cdot)$. The boundary conditions $s_1(\underline{v}) = \ldots = s_I(\underline{v}) = \underline{v}$ and $s_1(\overline{v}) = \ldots = s_I(\overline{v})$ give a common support $[\underline{b}, \overline{b}]$ for the bid distributions. Let $\mathcal{G}_R^I$ be the set of distributions $\mathbf{G}(\cdot, \ldots, \cdot)$ such that each marginal distribution $G_i(\cdot)$ satisfies Definition 3 with $G(b)/g(b)$ replaced by $1/H_i(b)$ in (v). Because the α-quantiles $(b_{1\alpha}, \ldots, b_{I\alpha})$ all correspond to the same α-quantile v_α, (26) evaluated at an α-quantile for an arbitrary pair (i, j) gives

$$b_{j\alpha} + \lambda_j^{-1}\left(\frac{1}{H_j(b_{j\alpha})}\right) = b_{i\alpha} + \lambda_i^{-1}\left(\frac{1}{H_i(b_{i\alpha})}\right). \tag{27}$$

The bid distribution $\mathbf{G}(\cdot, \ldots, \cdot)$ is then rationalized if and only if (i) Lemma 1 is satisfied with $\mathcal{G}_R^I$ and $\xi_i(\cdot), i = 1, \ldots, I$ as defined above and (ii) the *compatibility* condition (27) is satisfied for any pair (i, j) and $\alpha \in [0, 1]$. The latter reduces the set of bid distributions that can be rationalized relative to the symmetric case and may help in identification. Despite these conditions, the nonparametric model is still not identified.[24]

Proposition 6: *Any structure* $[U_1, \ldots, U_I, F] \in \mathcal{U}_{\mathcal{R}}^{\mathcal{I}} \times \mathcal{F}_{\mathcal{R}}$ *is not identified.*

On the other hand, if (say) $U_1(\cdot)$ is known, the nonparametric model $\mathcal{U}_R^I \times \mathcal{F}_R$ becomes identified as (26) for $i = 1$ allows to identify $F(\cdot)$. Thus, evaluated at the α-quantile, (26) for $i \neq 1$ allows to identify $\lambda_i(\cdot)$ on $[0, \max_\alpha(v_\alpha - b_{i\alpha})]$. This result can be used when bidders differ by their sizes, in which case "large" ones can be assumed to be risk neutral. The next proposition shows that the semiparametric model $\mathcal{U}^{CRRA} \times \mathcal{F}_R$ is identified without additional identifying conditions.

Proposition 7: *The semiparametric model* $\mathcal{U}^{CRRA} \times \mathcal{F}_R$ *is identified.*

[24] Parameterizing $F(\cdot)$ does not help as it does not make use of the compatibility conditions.

This result first noted in Campo (2002) contrasts with Proposition 4 and relies heavily on the presence of asymmetry in preferences as it exploits the compatibility conditions (27). Considering a different parametric specification for $U_i(\cdot), i = 1, \ldots, I$ introduces nonlinearities in (27). 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 can be conducted as follows. For some values of α, we estimate nonparametrically the bid quantiles $b_{i\alpha}(Z_\ell)$ and the functions $H_{i\ell}(\cdot)$, $i = 1, 2, \ell = 1, \ldots, L$. We then exploit (27) to contruct the following nonlinear model

$$\hat{b}_{2\alpha}(Z_\ell) - \hat{b}_{1\alpha}(Z_\ell) = \lambda_1^{-1}\left(\frac{1}{\hat{H}_{1\ell}(\hat{b}_{1\alpha}(Z_\ell))}; \theta_1\right) - \lambda_2^{-1}\left(\frac{1}{\hat{H}_{2\ell}(\hat{b}_{2\alpha}(Z_\ell))}; \theta_2\right) + e_{\alpha\ell} + \epsilon_{\alpha\ell},$$

for any $\alpha \in (0, 1)$, similarly to (10). As in Section 6.1, a weighted NLLS can be used for an arbitrary α. Asymptotic efficiency can be improved by integrating $[\hat{b}_{2\alpha}(Z_\ell) - \hat{b}_{1\alpha}(Z_\ell)]^2\omega(\alpha, Z_\ell)$ with respect to α. Properties of this estimator are left for future research.

8 Empirical Application

This section illustrates the previous methodology on US Forest Service (USFS) timber auctions. The USFS timber auction data have been used in several empirical studies. Comparing revenues generated from ascending and sealed-bid auctions, Hansen (1985) tests the revenue equivalence theorem. Using ascending auction data, Baldwin, Marshall and Richard (1997) study collusion, while Haile (2001) analyzes bidding behavior when there are resale opportunities after the auction. Athey and Levin (2001) study bidders' skewed bidding on species when payments are based on actual harvested value. Bidding diversification across species is consistent with bidders' risk aversion.[25] 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. While bidders' risk aversion is suspected in two of them, the extent of risk aversion has not been measured. The objective of our application is to shed some light on bidders' risk aversion.

We focus on the first-price sealed-bid auctions in 1979 for the Western half of the US (Regions 1 to 6). The data set contains 378 auctions involving a total of 1,400 bids.

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

The data contain a set of variables characterizing each timber tract varying from the various species, the estimated volume measured in thousand board feet (mbf), the logging costs, the tract acreage, the term of the contract in months, the month during which the auction was held, the tract location, the reserve price and the appraisal value. The latter is an estimated value of timber provided by the USFS taking into account its quality and quantity. In addition, the data provide the sealed bids in dollars and the bidders' identities. Table 1 gives some summary statistics on the total bids, the total winning bids, the appraisal value per mbf, the tract volume and the number of bidders. The auctioned tracts display important heterogeneity in quality and especially size. Though several variables can explain the bid variability, the tract appraisal value seems to be the best candidate to capture the heterogeneity in both volume and quality across tracts. In view of our nonparametric estimators for the first and third steps, we thus let Z_ℓ be the tract appraisal value. In addition, we consider that the reserve price is nonbinding based on empirical evidence provided by Haile (1996).

We follow the estimation procedure of Section 6.1 with $R = 1$ and $d = 1$. The first step consists in estimating nonparametrically the upper boundary $\overline{b}_0(Z_\ell, I_\ell)$ and the bid density at this upper boundary $g_0(\overline{b}_0(Z_\ell, I_\ell)|Z_\ell, I_\ell)$ for $\ell = 1, \ldots, L$.[26] Our one-sided kernel $\Phi(\cdot)$ is defined on $[-1, 0]$ with $\int_{-1}^{0} \Phi(x)dx = 1$ and $\int_{-1}^{0} x\Phi(x)dx = 0$. The linear kernel $\Phi(x) = (6x + 4)\mathbb{1}(-1 \leq x \leq 0)$ satisfies such requirements. The second step consists in estimating the risk aversion parameter θ. We experiment with three different functional forms: a CRRA specification, i.e. $U(x) = x^\theta$, a CRRA specification with common wealth w, i.e. $U(x) = (x + w)^\theta - w^\theta$ and a CARA specification, i.e. $U(x) = [1 - \exp(-\theta x)]/[1 - \exp(-\theta)]$.[27] The CRRA specification allows us to test for risk neutrality corresponding to $\theta = 1$ or equivalently the relative risk aversion $c = 0$. Morover, we also consider three different specifications for the upper boundary of the private distribution, namely a constant $\overline{v}(Z_\ell; \gamma) = \gamma_0$, a linear function $\overline{v}(Z_\ell; \gamma) = \gamma_0 + \gamma_1 Z_\ell$ and a quadratic function $\overline{v}(Z_\ell; \gamma) = \gamma_0 + \gamma_1 Z_\ell + \gamma_2 Z_\ell^2$. This makes a total of 9 different

[26] Because of the important dispersion in volume (see Table 1), our empirical analysis considers the 300 auctions for which the tract appraisal value Z_ℓ is smaller than \$300,000. Moroever, since the data do not provide enough auctions for four and more bidders, we pool all the data in the estimation of the upper boundary. Additional details on the estimation is available upon request to the authors.

[27] Wealth then becomes an additional parameter to be estimated.

specifications. For every specification, $m(Z_\ell, I_\ell; \beta)$ in (12) takes a different expression. The optimal weights $\omega^*(Z_\ell, I_\ell)$ are equal to $(I_\ell - 1)(\overline{v}(Z_\ell; \gamma_0) - \overline{b}_0(Z_\ell))$. This estimator is implemented by a standard two-step procedure in which the optimal weights are first estimated by ordinary NLLS. The standard errors are computed using Theorem 3. Table 2 provides the estimated results with standard errors in parentheses.[28]

For CRRA with or without wealth, a linear upper bound provides superior results to a constant upper bound, while the quadratic term is not significant. The estimated risk aversion parameter is about 0.30 and is stable across the linear and quadratic upper bounds though significantly larger than for the constant upper bound. Using experimental auction data, Goeree, Holt and Palfrey (2002) and Bajari and Hortacsu (2005) found a larger value for relative risk aversion in the $[0.50; 0.85]$ range. Nevertheless, risk neutrality ($\theta = 1$ or $c = 0$) is rejected in all six specifications suggesting that bidders are risk averse. We also find that the wealth parameter is insignificant suggesting no wealth effect.[29] The CARA specification provides some estimated coefficients for the upper bound quite close to those found for the CRRA specifications except for the quadratic term, which appears to be significant. The SSE is somewhat lower for the CARA specification though a display of both specifications in a graph representing the pairs $(Z_\ell, \hat{g}(\hat{\overline{b}}(Z_\ell)|Z_\ell))$ shows little difference in goodness of fit.

From a policy perpective, risk aversion implies that bidders bid more aggressively relative to the risk neutral case as they shade less their private values. In particular, a CRRA model is equivalent of having more competition in the auctions. For instance, with $I = 4$, a relative risk aversion $c = 0.32$ is roughly equivalent of having 5 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. Ignoring risk aversion, i.e. $\theta = 1$, leads to larger estimated private values thereby shifting to the right the estimated private value distribution as shown in Figure 1 for

[28] Estimation was performed without imposing the restriction $\hat{v}(Z_\ell, I_\ell; \hat{\gamma}) > \hat{b}(Z_\ell, I_\ell)$. We observe 83 and 81 violations for the constant case for the CRRA and CRRA with wealth specifications, respectively, and only 3 violations for the linear and quadratic cases for both specifications. The CARA specificition does not lead to any violation.

[29] The gain at the auction $v - b$ does not add directly to the firm's wealth measured as the firm's capital. This could explain why wealth is not significant. A more general (unidentified) form $U(w, v - b)$ may better capture the wealth effect.

$Z_\ell = \overline{Z} = 36,773$ and $I = 4$. Moreover, though the optimal mechanism with risk averse bidders involves some complex transfers among bidders (see Maskin and Riley (1984) and Matthews (1987)), an optimal posted reserve price can be set to generate more revenue for the seller. For a CRRA specification, the optimal reserve price p_0^* is solution of $p_0^* = v_0 + [((1-c)/(1-cI))[F^{(I-1)c/(1-c)}(p_0^* \mid z) - F(p_0^* \mid z)]]/f(p_0^* \mid z)$, where v_0 is the auctioned object value for the seller. For $\hat{c} = 0.3187$, $I = 4$, $Z_\ell = \overline{Z}$ and $v_0 = \overline{Z}$, we find $\hat{p}_0^*$ equal to \$54,905. The estimate with risk neutral bidders ($c = 0$) gives a significantly larger optimal reserve price at \$68,418. 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.

9 Conclusion

This paper extends the structural analysis of auction data to risk averse bidders. In particular, our methods allow researchers 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' behavior. On econometric grounds, the paper proposes a semiparametric method for estimating the structure of the model, namely bidders' risk aversion parameter(s) and the density of their private values. While previous papers have considered either fully parametric or nonparametric methods, this paper is the first one proposing a semiparametric estimator that arises naturally from the identification of the theoretical auction model.

Specifically, we show that any bid distribution can be rationalized by some auction model with risk averse bidders. This implies that the auction model with risk averse bidders is not testable in view of bids only. Moreover, the model is not identified and a model with constant absolute or relative risk aversion can be considered without loss of explanatory power. We then propose minimal restrictions to achieve semiparametric identification, namely 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 nonparametric boundary estimation, kernel estimators and weighted nonlinear least squares. We show that our estimator converges at the optimal rate, which is smaller than $\sqrt{N}$ and 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.

Many extensions can be entertained based on our methodology. A 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) extends the present method to random reserve prices and assesses empirically the gain for the seller of keeping the reserve price secret. Second, Campo (2005) considers an auction model with heterogeneous bidders for analyzing contruction procurements and shows that bidders' risk aversion decreases with bidders' experience. A third extension is conducted by Lu (2002) relying on Eso and White (2004) model with stochastic private values due ex ante uncertainties about the value of the auctioned object. Identification then becomes more involved. Lastly, given that little is known on bidders' utility function, Guerre, Perrigne and Vuong (2005) exploit some exlcusion restrictions to achieve nonparametric identification of the utility function.

Appendix A

Appendix A gathers proofs of Lemma 1 and Propositions 1, 2, 4–7.

Proof of Lemma 1: First, we prove that (i) and (ii) are necessary. Because $b_i = s(v_i, U, F, I)$ and the v_is are i.i.d., the b_is are also i.i.d. The fact that $G(\cdot) \in \mathcal{G}_R$ follows from applying Lemma 2 to the case with no conditioning variables (Z, I). To prove that (ii) is also necessary, consider (4), where $\lambda(\cdot) \equiv U(\cdot)/U'(\cdot)$. Thus $\lambda(\cdot)$ is defined from $I\!R_+$ to $I\!R_+$ because $\lambda(0) = \lim_{x \downarrow 0} \lambda(x) = 0$, as noted after Definition 1. As $U(\cdot)$ admits $R+2$ continuous derivatives on $(0, +\infty)$ with $U'(\cdot) > 0$, and $\lim_{x \downarrow 0} \lambda^{(r)}$ is finite for $r = 1, \ldots, R+1$, then $\lambda(\cdot)$ has $R+1$ continuous derivatives on $[0, +\infty)$. As $\lambda'(\cdot) = 1 - \lambda(\cdot)U''(\cdot)/U'(\cdot)$, we have $\lambda'(\cdot) \geq 1$ because $\lambda(\cdot) \geq 0$, $U'(\cdot) > 0$ and $U''(\cdot) \leq 0$. It remains to show that $\xi'(\cdot) > 0$. The equilibrium strategy must solve the differential equation (2). As (3) follows from (2), $s(\cdot)$ must satisfy $\xi[s(v), U, G, I] = v$ for all $v \in [\underline{v}, \overline{v}]$. We then obtain $\xi(b, U, G, I) = s^{-1}(b, U, F, I)$. This implies $\xi'(\cdot) = [s^{-1}(\cdot)]' > 0$ using Theorem 1.

Second, we show that (i) and (ii) are together sufficient. First, we construct a pair $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$. Let $U(\cdot)$ be such that $\lambda(\cdot) = U(\cdot)/U'(\cdot)$ or $U'(\cdot)/U(\cdot) = 1/\lambda(\cdot)$. Integrating with the normalization $U(1) = 1$ gives $U(x) = \exp \int_1^x 1/\lambda(t)dt$. We verify that $U(\cdot) \in \mathcal{U}_R$. Because $\lambda(\cdot)$ admits $R+1$ continuous derivatives on $[0, +\infty)$, then Definition 1-(iii) is clearly satisfied. Moreover, in the neighborhood of zero, $\lambda(t) \sim \lambda'(0)t$ with $1 \leq \lambda'(0) < \infty$. Thus $\int_x^1 1/\lambda(t)dt$ diverges to infinity, which implies that $U(x)$ tends to zero as $x \downarrow 0$. Define $U(0) = 0$ so that $U(\cdot)$ is continuous on $[0, +\infty)$. Because $U'(x) = \exp \int_1^x 1/\lambda(t)dt/\lambda(x)$, where $\lambda(\cdot) > 0$ on $(0, +\infty)$, we have $U'(\cdot) > 0$ on $(0, +\infty)$. The second-order derivative gives $U''(x) = [-\lambda'(x) + 1] \exp \int_1^x 1/\lambda(t)dt/\lambda^2(x)$. Since $\lambda'(x) \geq 1$, $U''(\cdot) \leq 0$ on $(0, +\infty)$. It remains to show that $U(\cdot)$ admits $R+2$ continuous derivatives on $(0, +\infty)$. By assumption, $\lambda(\cdot)$ has $R+1$ continuous derivatives on $[0, +\infty)$. It follows that $U(\cdot)$ admits $R+2$ continuous derivatives on $(0, +\infty)$.

Let $F(\cdot)$ be the distribution of $X = b + \lambda^{-1}[G(b)/(I-1)g(b)]$, where $b \sim G(\cdot)$. We verify that $F(\cdot) \in \mathcal{F}_R$. We have $F(x) = \Pr(X \leq x) = \Pr[\xi(b) \leq x] = \Pr[b \leq \xi^{-1}(x)] = G[\xi^{-1}(x)]$, because $\xi'(\cdot) > 0$ by assumption. This implies $F(\cdot) = G[\xi^{-1}(\cdot)]$ on $[\underline{v}, \overline{v}]$, where $\underline{v} \equiv \xi(\underline{b}) = \underline{b}$ and $\overline{v} \equiv \xi(\overline{b}) < \infty$ by continuity of $\xi(\cdot)$. Because $\xi(\cdot)$ and $G(\cdot)$ are strictly increasing, $F(\cdot)$ is strictly increasing on its support $[\underline{v}, \overline{v}]$. Moreover, $\xi(\cdot)$ is $R+1$ continuously differentiable on $[\underline{b}, \overline{b}]$. This follows from the definition of $\xi(\cdot)$, the $R+1$ continuous differentiability of $\lambda^{-1}(\cdot)$ on $[0, +\infty)$, and the $R+1$ continuous differentiability of $G(\cdot)/g(\cdot)$ on $[\underline{b}, \overline{b}]$, which follows from Definition 3-(iv,v). Thus $F(\cdot) = G[\xi^{-1}(\cdot)]$ admits $R+1$ continuous derivatives on $[\underline{v}, \overline{v}]$ because $G(\cdot)$ has $R+1$ continuous derivatives on $[\underline{b}, \overline{b}]$. It remains to show that $f(\cdot) > 0$ on $[\underline{v}, \overline{v}]$. This follows from $f(\cdot) = g[\xi^{-1}(\cdot)]/\xi'[\xi^{-1}(\cdot)]$, where $g(\cdot) > 0$ from Definition 3 and $\xi'(\cdot)$ is finite on $[\underline{b}, \overline{b}]$.

Lastly, we show that the pair $[U, F]$ rationalizes $G(\cdot)$, i.e. that $G(\cdot) = F[s^{-1}(\cdot, U, F, I)]$ on $[\underline{b}, \overline{b}]$, where $s(\cdot, U, F, I)$ solves (2) with the boundary condition $s(\underline{v}, U, F, I) = \underline{v}$. By construction of $F(\cdot)$, $G(\cdot) = F[\xi(\cdot)]$. Thus, it suffices to show that $\xi^{-1}(\cdot)$ solves (2) with the boundary condition $\xi^{-1}(\underline{v}) = \underline{v}$. The boundary condition is straightforward as $\xi(\underline{b}) = \underline{b} = \underline{v}$. By construction of $F(\cdot)$, $f(\cdot)/F(\cdot) = [\xi^{-1}(\cdot)]' g[\xi^{-1}(\cdot)]/G[\xi^{-1}(\cdot)]$. Thus $\xi^{-1}(\cdot)$ solves (2) if $1 = \{(I-1)g[\xi^{-1}(v)]\lambda[v - \xi^{-1}(v)]\}/G[\xi^{-1}(v)]$ for all $v \in [\underline{v}, \overline{v}]$. Making the change of variable $v = \xi(b)$ and noting that $\xi(b) - b = \lambda^{-1}[G(b)/(I-1)g(b)]$ from the definition of $\xi(\cdot)$, it follows that $\xi^{-1}(\cdot)$ solves (2) with boundary condition $\xi^{-1}(\underline{v}) = \underline{v}$.

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

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

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

(ii) The proof for the CARA case is similar. Consider $U \in \mathcal{U}_R^{CARA}$. 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 $\xi(b) = b + (1/a)\log\{1 + [aG(b)/(I-1)g(b)]\}$. We show that there exists $a > 0$ such that $\xi'(b) > 0$ on $[\underline{b}, \overline{b}]$. Differentiating gives

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

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

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

This is satisfied for an infinity of values for $a > 0$ provided the supremum is not $+\infty$. We know that $-[g(b)/G(b)]\{I - 1 + [G(b)/g(b)]'\}$ is R continuously differentiable on $(\underline{b}, \overline{b}]$ and hence continuous on $(\underline{b}, \overline{b}]$ because $R \ge 1$. Moreover, $\lim_{b \downarrow \underline{b}} -[g(b)/G(b)]\{I - 1 + [G(b)/g(b)]'\} = -\infty$

because $g(b)/G(b)$ tends to $+\infty$ and $[G(b)/g(b)]'$ tends to 1. Thus, we can always find an $a > 0$ satisfying (A.2) and hence a CARA model that rationalizes $G(\cdot)$.

Proof of Proposition 2: Let $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$ with bid distribution $G(\cdot) \in \mathcal{G}_R$ by Lemma 1. Let $[\tilde{U}, \tilde{F}]$ be such that $\tilde{U}(\cdot) = [U(\cdot/\delta)/U(1/\delta)]^\delta$, with $\delta \in (0, 1)$ and $\tilde{F}(\cdot)$ be the distribution of

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

where $b \sim G(\cdot)$. It is easy to check that $[\tilde{U}, \tilde{F}] \in \mathcal{U}_R \times \mathcal{F}_R$. Because $\tilde{\xi}(\cdot)$ is the weighted sum of two strictly increasing functions in b, then $\tilde{\xi}(\cdot)$ is strictly increasing. Hence, from Lemma 1 the structures $[U, F]$ and $[\tilde{U}, \tilde{F}]$ are observationally equivalent, and $[U, F]$ is not identified.

Proof of Proposition 4: We first consider the CRRA case. Let $[U, F] \in \mathcal{U}^{CRRA} \times \mathcal{F}_R$ with parameter $c \in [0, 1)$ generating a bid distribution $G(\cdot) \in \mathcal{G}_R$. 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 $\tilde{F}(\cdot) \in \mathcal{F}_R$ leading to the same $G(\cdot)$. 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 5: Let $[U, F]$ satisfy Assumption A1 with parameters (θ, γ) and $G(\cdot|\cdot, \cdot)$ 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 $G(\cdot|\cdot, \cdot)$. We first show that (θ, γ) is identified, i.e $(\theta, \gamma) = (\tilde{\theta}, \tilde{\gamma})$. Writing (5) for each structure gives

$$\frac{1}{I-1}\frac{\alpha}{g[b_\alpha(z, I)|z, I]} = \lambda[v_\alpha(z, I; \gamma) - b_\alpha(z, I); \theta] = \lambda[v_\alpha(z, I; \tilde{\gamma}) - b_\alpha(z, I); \tilde{\theta}],$$

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 (4), we have $v = b + \lambda^{-1}[G(b|z, I)/((I-1)g(b|z, I)); \theta] = \tilde{v}$, for every $b \in [\underline{b}(z, I), \overline{b}(z, I)]$ and $(z, I) \in \mathcal{Z} \times \mathcal{I}$. This shows that $\tilde{F}(\cdot|\cdot, \cdot) = F(\cdot|\cdot, \cdot)$, i.e. that the latter is identified.

Proof of Proposition 6: Let $[U_1, \ldots, U_I, F] \in \mathcal{U}_R^I \times \mathcal{F}_R$, which thus generates $[G_1, \ldots, G_I] \in \mathcal{G}_R^I$ that satisfies the compatibility condition (27). We show that there exists another structure $[\tilde{U}_1, \ldots, \tilde{U}_I, \tilde{F}] \in \mathcal{U}_R^I \times \mathcal{F}_R$ rationalizing $[G_1, \ldots, G_I]$. The proof is in four steps.
Step 1: *Construction of* $[\tilde{U}_1, \ldots, \tilde{U}_I, \tilde{F}]$. Let $\tilde{U}_1(\cdot) = [U_1(\cdot/\delta)/U_1(1/\delta)]^\delta$ with $\delta \in (0, 1)$. Thus, $\tilde{\lambda}_1(\cdot) = \lambda_1(\cdot/\delta)$ and $\tilde{\lambda}_1^{-1}(\cdot) = \delta\lambda_1^{-1}(\cdot)$. For $i = 2, \ldots, I$, let $\tilde{U}_i(x) = \exp\left[\int_1^x 1/\tilde{\lambda}_i(t)dt\right]$ so that $\tilde{\lambda}_i(\cdot) = \tilde{U}_i(\cdot)/\tilde{U}_i'(\cdot)$, where $\tilde{\lambda}_i(\cdot)$ is such that $\tilde{\lambda}_i^{-1}[1/H_i(b_{i\alpha})] = \tilde{\lambda}_1^{-1}[1/H_1(b_{1\alpha})] + b_{1\alpha} - b_{i\alpha}$, for all $\alpha \in [0, 1]$. The latter well-defines $\tilde{\lambda}_i^{-1}(\cdot)$ because $1/H_i(b_{i\alpha})$ strictly increases as α increases given $H_i'(\cdot) < 0$. Moreover, $\tilde{\lambda}_i(\cdot)$ is strictly increasing as shown in Step 3. Note that the

compatibility condition (27) is satisfied by construction. We then let $\tilde{F}(\cdot)$ be the distribution of $\tilde{v}_i \equiv b_i + \tilde{\lambda}_i^{-1}[1/H_i(b_i)] \equiv \tilde{\xi}_i(b_i)$ for an arbitrary i, where $b_i \sim G_i(\cdot)$. Using $\tilde{\lambda}_1^{-1}(\cdot) = \delta\lambda_1^{-1}(\cdot)$, we obtain $\tilde{\lambda}_i^{-1}[1/H_i(b_{i\alpha})] = \delta\lambda_1^{-1}[1/H_1(b_{1\alpha})] + b_{1\alpha} - b_{i\alpha}$. Thus, (27) with $j = 1$ gives

$$\tilde{\lambda}_i^{-1}\left(\frac{1}{H_i(b_{i\alpha})}\right) = \delta\lambda_i^{-1}\left(\frac{1}{H_i(b_{i\alpha})}\right) + (1-\delta)(b_{1\alpha} - b_{i\alpha}). \tag{A.3}$$

Equivalently, $\tilde{\lambda}_i^{-1}[1/H_i(b_{i\alpha})] = \lambda_i^{-1}(1/H_i(b_{i\alpha})) - (1-\delta)\lambda_1^{-1}[1/H_1(b_{1\alpha})]$. In particular, since $\lambda_i^{-1}(\cdot)$ is bidder's i shading, the shading under $[\tilde{U}_1, \ldots, \tilde{U}_I, \tilde{F}]$ is smaller than under $[U_1, \ldots, U_I, F]$, i.e. bidders tend to bid more aggressively under the former than under the latter.

Step 2: *$\tilde{\lambda}_i(0) = 0$ and $\tilde{\xi}_i'(\cdot) > 0$ on $[\underline{b}, \overline{b}]$.* Because $[U_1, \ldots, U_I, F] \in \mathcal{U}_R^I \times \mathcal{F}_R$ so that $[G_1, \ldots, G_I] \in \mathcal{G}_R^I$, we have $\lambda_i^{-1}(0) = 0$ and $\lim_{b \downarrow \underline{b}} 1/H_i(b) = 0$. Thus, (A.3) with the boundary conditions $\underline{b}_1 = \ldots = \underline{b}_I \equiv \underline{b} = \underline{v}$ imply $\tilde{\lambda}_i^{-1}(0) = 0$ and hence $\tilde{\lambda}_i(0) = 0$. Regarding $\tilde{\xi}_i'(\cdot) > 0$, we note that $\tilde{\xi}_i(b_{i\alpha}) = (1-\delta)b_{1\alpha} + \delta\xi_i(b_{i\alpha})$ from (A.3) and (26). Noting that $b_{1\alpha} = G_1^{-1}[G_i(b_{i\alpha})] \equiv B_i(b_{i\alpha})$ and letting $b_{i\alpha} = b$, we obtain $\tilde{\xi}_i'(b) = (1-\delta)B_i'(b) + \delta\xi_i'(b)$, where $B_i'(b) = g_i(b)/g_1[B(b)]$. Hence, $\tilde{\xi}_i'(b) > 0$ since $B_i'(b) > 0$ and $\xi_i'(b) > 0$.

Step 3: *$\tilde{\lambda}_i'(\cdot) \geq 1$.* From (A.3) and (26), $\tilde{\lambda}_i^{-1}[1/H_i(b_{i\alpha})] = \delta\xi_i(b_{i\alpha}) + (1-\delta)b_{1\alpha} - b_{i\alpha}$, i.e. $1/H_i(b_{i\alpha}) = \tilde{\lambda}_i[\delta\xi_i(b_{i\alpha}) + (1-\delta)b_{1\alpha} - b_{i\alpha}]$. From the structure $[U_1, \ldots, U_I, F]$, we have $1/H_i(b_{i\alpha}) = \lambda_i[\xi_i(b_{i\alpha}) - b_{i\alpha}]$. Thus, $\lambda_i[\xi_i(b_{i\alpha}) - b_{i\alpha}] = \tilde{\lambda}_i[\delta\xi_i(b_{i\alpha}) + (1-\delta)b_{1\alpha} - b_{i\alpha}]$. Differentiating with respect to $b = b_{i\alpha}$ and noting that $b_{1\alpha} = G_1^{-1}[G_i(b_{i\alpha})] \equiv B_i(b_{i\alpha})$ gives

$$\tilde{\lambda}_i'(**) = \frac{\xi_i'(b) - 1}{\delta\xi_i'(b) + (1-\delta)B_i'(b) - 1}\lambda_i'(*) \equiv R_i(b)\lambda_i'(*), \tag{A.4}$$

where the different arguments of $\lambda_i'(\cdot)$ and $\tilde{\lambda}_i'(\cdot)$ are indicated by * and **, respectively. Thus, it suffices to show that $R_i(\cdot) \geq 1$ since $\lambda_i'(\cdot) \geq 1$. We note that $\xi_1(b_{1\alpha}) = \xi_i(b_{i\alpha}) = v_\alpha$ for all $\alpha \in [0,1]$ from the compatibility conditions. Using $b_{1\alpha} = B_i(b_{i\alpha})$, this gives $\xi_1[B_i(b)] = \xi_i(b)$ for all $b \in [\underline{b}, \overline{b}]$. Differentiating gives $\xi_1'[B_i(b)]B_i'(b) = \xi_i'(b)$, i.e. $B_i'(b) = \xi_i'(b)/\xi_1'[B_i(b)]$. Hence,

$$R_i(b) = \frac{\xi_i'(b) - 1}{\delta\xi_i'(b) - 1 + (1-\delta)\frac{\xi_i'(b)}{\xi_1'[B_i(b)]}} = 1 + \frac{(1-\delta)\xi_i'(b)\{\xi_1'[B_i(b)] - 1\}}{\delta\xi_i'(b)\{\xi_1'[B_i(b)] - 1\} - \{\xi_1'[B_i(b)] - \xi_i'(b)\}}, \tag{A.5}$$

for $b \in [\underline{b}, \overline{b}]$. Note that $\xi_i'(\cdot) > 1$ on $(\underline{b}, \overline{b}]$ for every $i = 1, \ldots, I$ since differentiating (26) gives $\xi_i'(b) = 1 - \lambda_i^{-1'}[1/H_i(b)][H_i'(b)/H_i^2(b)]$, where $\lambda_i^{-1'}(\cdot) > 0$ and $H_i'(\cdot) < 0$ by assumption. Hence, $\xi_i'(\cdot) \geq 1$ on $[\underline{b}, \overline{b}]$ by continuity. Since $1 - \delta > 0$ and $\xi_i'(\cdot) > 0$, it suffices to show that the denominator $D_i(b)$ (say) in the RHS is strictly positive for all $b \in [\underline{b}, \overline{b}]$ and some $\delta \in [\delta^*, 1]$.

To study the sign of $D_i(\cdot)$ on $[\underline{b}, \overline{b}]$, we note that

$$\frac{g_j(b)}{G_j(b)} = \frac{g_j(\underline{b}) + o(1)}{g_j(\underline{b})(b - \underline{b}) + o(b - \underline{b})} = \frac{1}{b - \underline{b}}\frac{g_j(\underline{b}) + o(1)}{g_j(\underline{b}) + o(1)} = \frac{1}{b - \underline{b}}(1 + o(1)).$$

Thus, a Taylor expansion of $1 = \lambda_i[\xi_i(b) - b]\sum_{j\neq i}[g_j(b)/G_j(b)]$ from (26) gives

$$1 = \{\lambda_i'(0)[\xi_i'(\underline{b}) - 1](b - \underline{b}) + o(b - \underline{b})\}\frac{I-1}{b-\underline{b}}(1 + o(1)) = \{\lambda_i'(0)[\xi_i'(\underline{b}) - 1](I-1)\} + o(1).$$

Hence, $\xi_i'(\underline{b}) = 1 + \{1/[(I-1)\lambda_i'(0)]\} > 1$ as $\lambda_i'(\cdot) \geq 1$. Thus, because $\xi_i'(\cdot) > 1$ on $[\underline{b}, \overline{b}]$, then $D_i(\cdot) > 0$ on $[\underline{b}, \overline{b}]$ if and only if $\delta > \delta^* \equiv \max_{b\in[\underline{b},\overline{b}]}\underline{\delta}(b)$, where $\underline{\delta}(\cdot)$ is continuous on $[\underline{b}, \overline{b}]$ with

$$\underline{\delta}(b) \equiv \frac{\xi_1'[B_i(b)] - \xi_i'(b)}{\xi_i'(b)\{\xi_1'[B_i(b)] - 1\}} = \frac{1}{\xi_i'(b)}\left[1 - \frac{\xi_i'(b) - 1}{\xi_1'[B_i(b)] - 1}\right].$$

It remains to show that $\underline{\delta}(\cdot) < 1$ on $[\underline{b}, \overline{b}]$ so that $\delta^* < 1$. Clearly, $\underline{\delta}(\cdot) < 1$ on $(\underline{b}, \overline{b}]$ as $\xi_i'(\cdot) > 1$ on $(\underline{b}, \overline{b}]$. Moreover, at $b = \underline{b}$, we have $\underline{\delta}(\underline{b}) = [1/\xi_i'(\underline{b})]\{1 - [\lambda_i'(0)/\lambda_1'(0)]\} < 1$.

Step 4: $[\tilde{U}_1, \ldots, \tilde{U}_I, \tilde{F}] \in \mathcal{U}_R^I \times \mathcal{F}_R$. From the previous steps and the rationalization result given after (27), it follows that $[\tilde{U}_1, \ldots, \tilde{U}_I, F]$ rationalizes $[G_1, \ldots, G_I]$. It remains to show that $[\tilde{U}_1, \ldots, \tilde{U}_I, \tilde{F}] \in \mathcal{U}_R^I \times \mathcal{F}_R$. From the proof of Lemma 1, it sufffices to show that $\tilde{\lambda}_i(\cdot)$ is $R+1$ continuously differentiable on $[0, \infty)$ for $i = 1, \ldots, I$. This follows from (A.4)–(A.5) and the $R+1$ continuous differentiablity of $\lambda_i(\cdot)$ and $\xi_i(\cdot)$ as $[G_1, \ldots, G_I] \in \mathcal{G}_R^I$.

Proof of Proposition 7: Consider any pair (i, j) of individuals such that $c_i \neq c_j$. The compatibility condition (27) for a CRRA model is

$$b_{j\alpha} - b_{i\alpha} = (1 - c_i)/H_i(b_{i\alpha}) - (1 - c_j)/H_j(b_{j\alpha}) \quad \text{for all } \alpha \in [0, 1]. \tag{A.6}$$

We first show that there exists $(\alpha, \tilde{\alpha}) \in [0, 1]^2$ such that $H_i(b_{i\alpha})/H_j(b_{j\alpha}) \neq H_i(b_{i\tilde{\alpha}})/H_j(b_{j\tilde{\alpha}})$. Suppose not, then $H_i(b_{i\alpha})/H_j(b_{j\alpha})$ is a constant. Because (A.6) at $\overline{b}$ gives $H_i(\overline{b})/H_j(\overline{b}) = (1 - c_i)/(1 - c_j)$, then $H_i(b_{i\alpha})/H_j(b_{j\alpha}) = (1 - c_i)/(1 - c_j)$ for all $\alpha \in [0, 1]$. Using this in (A.6) gives $b_{i\alpha} = b_{j\alpha}$, which implies $G_i(\cdot) = G_j(\cdot)$ and hence $H_i(\cdot) = H_j(\cdot)$ because $H_i(\cdot) = [g_j(\cdot)/G_j(\cdot)] + \sum_{k\neq i,j}[g_k(\cdot)/G_k(\cdot)]$ and $H_j(\cdot) = [g_i(\cdot)/G_i(\cdot)] + \sum_{k\neq i,j}[g_k(\cdot)/G_k(\cdot)]$. From (A.6) 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 $H_i(b_{i\alpha})/H_j(b_{j\alpha}) \neq H_i(b_{i\tilde{\alpha}})/H_j(b_{j\tilde{\alpha}})$. This guarantees that (A.6) 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 $F(\cdot)$ follows from (26). A similar result can be found in Campo (2005) who assumes that $H_i(b_{i\alpha})/H_j(b_{j\alpha})$ is not constant.

Appendix B

Appendix B proves Theorem 2 and Theorem 3. The proofs use some lemmas, which are proved in Appendix C. 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 $u = (u_{i\ell}) \in I\!\!R^N$, 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 2: The proof is in three steps.

STEP 1: *Smoothness of* $m(z, I; \beta)$. We have

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 (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 $\psi(\cdot) : I\!\!R_- \to I\!\!R$ be an infinitely differentiable function on $I\!\!R_-$ with support $[-1, 0]$, such that $\psi(0) = 1$, $\int \psi(x)dx = 0$. Let $1\!\!I_p = (1, \ldots, 1)' \in I\!\!R^p$. For a fixed constant $\kappa > 0$, consider the following perturbations of θ_0 and $g_0(b|z, I)$, where $I \in \mathcal{I}$,

$$
\begin{aligned}
\beta_N &= (\theta_N', \overline{v}_0)' = (\theta_0' + 2t1\!\!I_p'/\rho_N, \overline{v}_0)' = \beta_0 + (2t1\!\!I_p'/\rho_N, 0)', \\
g_N(b|z, I) &= g_0(b|z, I) + \pi_N(z, I)\psi\left[\kappa \rho_N^{1/(R+1)}\left(b - \overline{b}_0(z, I)\right)\right], \\
\pi_N(z, I) &= m(z, I; \beta_N) - m(z, I; \beta_0) = \frac{\partial m(z, I; \beta_0)}{\partial \beta}(\beta_N - \beta_0) + o(\|\beta_N - \beta_0\|) = O(1/\rho_N).
\end{aligned}
$$

Without loss, we can assume that $\{\beta_N; N = 1, 2, \ldots\}$ is in a compact subset $\mathcal{B}_c \subset \mathcal{B}$ 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 2-(iii) it follows that $g_N(\cdot|z, I)$ is a conditional density with support $[\underline{b}_0, \overline{b}_0(z, I)]$ 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 $(z, I) \in \mathcal{Z} \times \mathcal{I}$, $G_N(\cdot|z, I)$ *is rationalized by the IPV auction structure with risk aversion* $[U(\cdot; \theta_N), F_N(\cdot|z, I)]$, *where* $F_N(\cdot|\cdot, I) \in \mathcal{F}_R^*$ *with support* $[\underline{v}_0, \overline{v}_0]$,

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

STEP 3: *Lower Bound.* Using the triangular inequality we have for any estimator $\tilde{\beta}$

$$
\begin{aligned}
\Pr_{\beta_N, f_N}\left(\|\rho_N(\tilde{\beta} - \beta_N)\|_\infty \geq t\right) &\geq \Pr_{\beta_N, f_N}\left(\|\rho_N(\beta_N - \beta_0)\|_\infty - \|\rho_N(\tilde{\beta} - \beta_0\|_\infty \geq t\right) \\
&\geq \Pr_{\beta_N, f_N}\left(\|\rho_N(\tilde{\beta} - \beta_0)\|_\infty < t\right)
\end{aligned}
$$

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

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

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

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

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 $g_j(\cdot|\cdot,\cdot)$, $j=0,N$. 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

$$\begin{aligned}
\Pr_e(\mathcal{F}_L) &= 1-\Big(\Pr_0(\|\rho_N(\tilde\beta-\beta_0)\|_\infty<t)-\Pr_N(\|\rho_N(\tilde\beta-\beta_0)\|_\infty<t)\Big)\\
&\geq 1-\sup_A|\Pr_0(A)-\Pr_N(A)| = 1-\frac{1}{2}\int|d\Pr_0-d\Pr_N|\\
&\geq 1-\left[\int\left(\sqrt{d\Pr_0}-\sqrt{d\Pr_N}\right)^2\right]^{1/2} = 1-\sqrt{2}\left(1-\int\sqrt{d\Pr_0 d\Pr_N}\right)^{1/2}\\
&= 1-\sqrt{2}\left(1-\prod_{\ell=1}^{L}\prod_{i=1}^{I_\ell}\int_{\underline{b}_0}^{\overline{b}_0(Z_\ell,I_\ell)}\sqrt{g_0(b_{i\ell}|Z_\ell,I_\ell)g_N(b_{i\ell}|Z_\ell,I_\ell)}db_{i\ell}\right)^{1/2}. \quad \text{(B.2)}
\end{aligned}$$

But, because $g_j(\cdot|\cdot,\cdot)$, $j=0,N$, 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 $g_N(\cdot|\cdot,\cdot)$ and a Taylor expansion

$$\begin{aligned}
&\int_{\underline{b}_0}^{\overline{b}_0(Z_\ell,I_\ell)}\sqrt{g_0(b|Z_\ell,I_\ell)g_N(b|Z_\ell,I_\ell)}db\\
&= \int_{\underline{b}_0}^{\overline{b}_0(Z_\ell,I_\ell)} g_0(b|Z_\ell,I_\ell)\sqrt{1+\frac{\pi_N(Z_\ell,I_\ell)}{g_0(b|Z_\ell,I_\ell)}\psi\left(\kappa\rho_N^{\frac{1}{R+1}}\left(b-\overline{b}_0(Z_\ell,I_\ell)\right)\right)}db\\
&= \int_{\underline{b}_0}^{\overline{b}_0(Z_\ell,I_\ell)} g_0(b|Z_\ell,I_\ell)\Bigg[1+\frac{\pi_N(Z_\ell,I_\ell)}{2g_0(b|Z_\ell,I_\ell)}\psi\left(\kappa\rho_N^{\frac{1}{R+1}}\left(b-\overline{b}_0(Z_\ell,I_\ell)\right)\right)\\
&\quad -\frac{\pi_N^2(Z_\ell,I_\ell)}{8g_0^2(b|Z_\ell,I_\ell)}\psi^2\left(\kappa\rho_N^{\frac{1}{R+1}}\left(b-\overline{b}_0(Z_\ell,I_\ell)\right)\right)\Bigg]db+O\left(\frac{1}{\rho_N^3}\right)\\
&= 1+\frac{\pi_N(Z_\ell,I_\ell)}{2}\int_{\underline{b}_0}^{\overline{b}_0(Z_\ell,I_\ell)}\psi\left(\kappa\rho_N^{\frac{1}{R+1}}\left(b-\overline{b}_0(Z_\ell,I_\ell)\right)\right)db\\
&\quad -\frac{\pi_N^2(Z_\ell,I_\ell)}{8\kappa\rho_N^{\frac{1}{R+1}}}\int_{-1}^{0}\frac{\psi^2(x)}{g_0\left(\overline{b}_0(Z_\ell,I_\ell)+\rho_N^{-\frac{1}{R+1}}x/\kappa\right)}dx+O\left(\rho_N^{-3}\right)\\
&= 1+0+O\left(\rho_N^{-\frac{1}{R+1}-2}\right)=1+O\left(\rho_N^{-\frac{2R+3}{R+1}}\right),
\end{aligned}$$

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

$$\begin{aligned}
&\prod_{\ell=1}^{L}\prod_{i=1}^{I_\ell}\int_{\underline{b}_0}^{\overline{b}_0(Z_\ell,I_\ell)}\sqrt{g_0(b_{i\ell}|Z_\ell,I_\ell)g_N(b_{i\ell}|Z_\ell,I_\ell)}db_{i\ell}\\
&= \left[1+O\left(\rho_N^{-\frac{2R+3}{R+1}}\right)\right]^N = \exp\left[N\log\left(1+O\left(\rho_N^{-(2R+3)/(R+1)}\right)\right)\right]=1+o(1).
\end{aligned}$$

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

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

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

Proof of Theorem 3: The proof is in three steps. Hereafter, the compactness of $\mathcal{Z} \times \mathcal{I}$ is heavily exploited, while all limits are taken as $N \to \infty$.

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

Lemma B3: *Let A2–A3 and A4-(iii,iv) hold.*
(i) The variables $Y_{i\ell}$ (or $\epsilon_{i\ell}$), $1 \leq i \leq I_\ell$, $1 \leq \ell \leq L$ are independent given $\mathcal{F}_L$,
(ii) Uniformly in (i,ℓ),

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

(iii) Uniformly in (i,ℓ)

$$\begin{aligned}
\mathrm{Var}[\epsilon_{i\ell}|\mathcal{F}_L] &= \frac{g_0(\overline{b}_0(Z_\ell,I_\ell)|Z_\ell,I_\ell)+o(1)}{h_N}\int\Phi^2(x)dx = \frac{m(Z_\ell,I_\ell;\beta_0)+o(1)}{h_N}\int\Phi^2(x)dx, \\
\max_{1\leq\ell\leq L, 1\leq i\leq I_\ell}|\epsilon_{i\ell}| &\leq \frac{2\sup_{x\in\mathbb{R}}|\Phi(x)|}{h_N}.
\end{aligned}$$

The second lemma is a Central Limit Theorem, which is useful for weighted averages of $\epsilon_{i\ell}$.

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

$$\begin{aligned}
\left|\sum_{\ell=1}^L\sum_{i=1}^{I_\ell} u_{i\ell}\left(\hat{Y}_{i\ell}-Y_{i\ell}\right)\right| &\leq \sum_{\ell=1}^L\sum_{i=1}^{I_\ell}|u_{i\ell}|\left|\hat{Y}_{i\ell}-Y_{i\ell}\right| = O_P\left[\max\left(\|u\|_1\frac{a_N}{h_N}, \|u\|_2\frac{\sqrt{a_N}}{h_N}\right)\right], \\
\sum_{\ell=1}^L\sum_{i=1}^{I_\ell}|u_{i\ell}|\left(\hat{Y}_{i\ell}-Y_{i\ell}\right)^2 &= \|u\|_1 O_P\left(\frac{a_N}{h_N^2}\right).
\end{aligned}$$

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 (19).

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

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

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

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

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

The last lemma deals with the following processes

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

Lemma B9: *Let A2–A3 and A4-(i,ii,iii,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 $|\max_{1\le\ell\le L}\hat{\overline{b}}_N(Z_\ell,I_\ell)-\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\overline{b}_0(z,I)|<\delta/4$ 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, $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\overline{b}_0(z,I)+\delta/4<\max_{1\le\ell\le L}\hat{\overline{b}}_N(Z_\ell,I_\ell)+\delta/2<\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\overline{b}_0(z,I)+3\delta/4<\overline{v}_0<\overline{v}_{\sup}$ 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, (12), (13) and the triangular inequality give

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

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

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

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

$$\begin{aligned}
Q_N(\beta) &\leq 3 \sum_{\ell=1}^{L} \sum_{i=1}^{I_\ell} \omega(Z_\ell, I_\ell) \left[\left(m(Z_\ell, I_\ell; \beta) - m(Z_\ell, I_\ell; \beta_0) \right)^2 + e_{i\ell}^2 + \epsilon_{i\ell}^2 \right] \\
&= O_P(N) + O_P\left(N h_N^{2(R+1)} \right) + O_P(N/h_N) = O_P(N/h_N), \quad \text{(B.4)}
\end{aligned}$$

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

$$\begin{aligned}
\sup_{\beta \in \mathcal{B}_{\delta/4}} |\hat{Q}_N(\beta) - Q_N(\beta)| &= \sup_{\beta \in \mathcal{B}_{\delta/4}} \left| \left(\hat{Q}_N^{1/2}(\beta) - Q_N^{1/2}(\beta) \right) \left(2\hat{Q}_N^{1/2}(\beta) + Q_N^{1/2}(\beta) - \hat{Q}_N^{1/2}(\beta) \right) \right| \\
&\leq 2 \sup_{\beta \in \mathcal{B}_{\delta/4}} Q_N^{1/2}(\beta) \sup_{\beta \in \mathcal{B}_{\delta/4}} |\hat{Q}_N^{1/2}(\beta) - Q_N^{1/2}(\beta)| + \sup_{\beta \in \mathcal{B}_{\delta/4}} |\hat{Q}_N^{1/2}(\beta) - Q_N^{1/2}(\beta)|^2 \\
&= N O_P\left(\sqrt{\frac{a_N}{h_N^3}} \right) + N O_P\left(\frac{a_N}{h_N^2} \right) = o_P(N), \quad \text{(B.5)}
\end{aligned}$$

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 (19). We have

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

using (10) and (12). Hence,

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

using Lemmas 2-(i,iv), B1, B3-(ii), B9 and A4-(iv). Thus,

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

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 $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}} \overline{b}_0(z,I) + \delta < \overline{v}_0 < \overline{v}_{\text{sup}}$. 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

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

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

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

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

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

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

$$\hat{S}_N = \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell} \omega(Z_\ell, I_\ell)\left(Y_{i\ell} - m(Z_\ell, I_\ell; \beta_0)\right)\frac{\partial m(Z_\ell, I_\ell; \hat{\beta}_N)}{\partial\beta}$$

$$+\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)\left(m(Z_\ell,I_\ell;\beta_0)\frac{\partial m(Z_\ell,I_\ell;\hat\beta_N)}{\partial\beta}-\hat m(Z_\ell,I_\ell;\beta_0)\frac{\partial\hat m(Z_\ell,I_\ell;\hat\beta_N)}{\partial\beta}\right)$$
$$+\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)\left[\left(\hat Y_{i\ell}-Y_{i\ell}\right)\frac{\partial\hat m(Z_\ell,I_\ell;\hat\beta_N)}{\partial\beta}+Y_{i\ell}\left(\frac{\partial\hat m(Z_\ell,I_\ell;\hat\beta_N)}{\partial\beta}-\frac{\partial m(Z_\ell,I_\ell;\hat\beta_N)}{\partial\beta}\right)\right].$$

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

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

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

$$\begin{aligned}\hat S_N &= \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)\left(e_{i\ell}+\epsilon_{i\ell}\right)\frac{\partial m(Z_\ell,I_\ell;\beta_0)}{\partial\beta}+o_P\left(Nh_N^{R+1}+\sqrt{\frac{N}{h_N}}\right)\\ &= Nh_N^{R+1}\mathbf{b}(\beta_0,g_0)+\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell,I_\ell)\epsilon_{i\ell}\frac{\partial m(Z_\ell,I_\ell;\beta_0)}{\partial\beta}+o_P\left(Nh_N^{R+1}+\sqrt{\frac{N}{h_N}}\right).\end{aligned}\quad\text{(B.9)}$$

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

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

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

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

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

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

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 (B.8).

Appendix C

Appendix C gathers proofs of Lemmas B1–B9 stated in Appendix B.

Proof of Lemma B1: By Lemma 2-(i), $\overline{b}_0(\cdot, I)$ has $R+1$ continuous derivatives on $\mathcal{Z}$. The desired result then follows from (7), $\lambda(\cdot;\cdot) = U(\cdot;\cdot)/U'(\cdot;\cdot)$, which is strictly positive and $R+1$ continuous differentiable on $(0,\infty)\times\Theta$, and the definition of $\mathcal{B}$. That $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}} \overline{b}_0(z, I) < \overline{v}_0$ follows from the compactness of $\mathcal{Z}$, the finiteness of $\mathcal{I}$, Lemma 2-(i) and Theorem 1-(i).

Proof of Lemma B2: The proof is in four steps. For $j = 0, N$, define

$$\xi_j(b; z, I) = b + \lambda^{-1}\left(\frac{1}{I-1}\frac{G_j(b|z, I)}{g_j(b|z, I)}; \theta_j\right),\ b \in [\underline{v}_0, \overline{b}_0(z, I)],\ (z, I) \in \mathcal{Z}\times\mathcal{I}.$$

Step 1: *$G_N(\cdot|\cdot,\cdot)$ satisfies the properties of Lemma 2.* Let $\Psi(b) = \int_{-\infty}^{b}\psi(x)dx$. We have

$$G_N(b|z, I) \quad = \quad G_0(b|z, I) + \pi_N(z, I)\kappa^{-1}\rho_N^{-\frac{1}{R+1}}\Psi\left(\kappa\rho_N^{\frac{1}{R+1}}\left(b - \overline{b}_0(z, I)\right)\right).$$

In particular, $G_N(\cdot|z, I)$ and $g_N(\cdot|z, I)$ are equal to $G_0(\cdot|z, I)$ and $g_0(\cdot|z, I)$ on $[\underline{b}_0, \overline{b}_0(z, I) - \rho_N^{-1/(R+1)}]$, while differing from the latter on $[\overline{b}_0(z, I) - \kappa^{-1}\rho_N^{-1/(R+1)}, \overline{b}_0(z, I)]$. Now, $G_0(\cdot|\cdot,\cdot)$ satisfies Lemma 2 under Assumption A2. Moreover, $\Psi(\cdot)$ is infinitely differentiable on $I\!R_-$, while $\pi_N(\cdot, I)$ is $R+1$ continuously differentiable on $\mathcal{Z}$ in view of Lemma B1. Therefore, for N large enough, $G_N(\cdot|\cdot,\cdot)$ satisfies the properties of Lemma 2 with $\underline{b}_N(z, I) = \underline{b}_0(z, I) = \underline{v}_0$ and $\overline{b}_N(z, I) = \overline{b}_0(z, I)$ for all $(z, I) \in \mathcal{Z}\times\mathcal{I}$. In particular, as for $G_0(\cdot|\cdot, I)/g_0(\cdot|\cdot, I)$, the ratio $G_N(\cdot|\cdot, I)/g_N(\cdot|\cdot, I)$ is $R+1$ continuously differentiable on $S_I(G_0)$ for every $I \in \mathcal{I}$.

Step 2: *Properties of $\xi_N(\cdot|z, I) - \xi_0(\cdot|z, I)$ and its derivatives.* Step 1 and Lemma 2 yield

$$\frac{\partial^r G_N(b|z, I)}{\partial b^r} - \frac{\partial^r G_0(b|z, I)}{\partial b^r} = \pi_N(z, I)\kappa^{r-1}\rho_N^{\frac{r-1}{R+1}}\Psi^{(r)}\left(\kappa\rho_N^{\frac{1}{R+1}}\left(b - \overline{b}_0(z, I)\right)\right),$$

$$\frac{\partial^r g_N(b|z, I)}{\partial b^r} - \frac{\partial^r g_0(b|z, I)}{\partial b^r} = \pi_N(z, I)\kappa^{r}\rho_N^{\frac{r}{R+1}}\psi^{(r)}\left(\kappa\rho_N^{\frac{1}{R+1}}\left(b - \overline{b}_0(z, I)\right)\right),$$

where $b \in [\underline{v}_0, \overline{b}_0(z,I)]$ for $0 \leq r \leq R+1$ with the exception that $b \in (\underline{v}_0, \overline{b}_0(z,I)]$ when $r = R+1$ in the second equality. By Lemma 2-(i) there is a b_* with $\underline{v}_0 < b_* < \overline{b}_0(z,I)$ for all (z,I). Because $\sup_{z,I} |\pi_N(z,I)| = O(1/\rho_N)$ and $|\psi^{(r)}(\cdot)|$ and $|\Psi^{(r)}(\cdot)|$ are bounded, this gives

$$\sup_{(b,z,I)\in \cup_{\mathcal{Z}\times\mathcal{I}}[b_*,\overline{b}_0(z,I)]\times\{z,I\}} \left| \frac{\partial^r G_N(b|z,I)}{\partial b^r} - \frac{\partial^r G_0(b|z,I)}{\partial b^r} \right| = \kappa^{r-1} O\left(\rho_N^{\frac{r-1-(R+1)}{R+1}} \right), \quad \text{(C.1)}$$

$$\sup_{(b,z,I)\in \cup_{\mathcal{Z}\times\mathcal{I}}[b_*,\overline{b}_0(z,I)]\times\{z,I\}} \left| \frac{\partial^r g_N(b|z,I)}{\partial b^r} - \frac{\partial^r g_0(b|z,I)}{\partial b^r} \right| = \kappa^{r} O\left(\rho_N^{\frac{r-(R+1)}{R+1}} \right), \quad \text{(C.2)}$$

for $r = 0, \ldots, R+1$, where the remainder terms are independent of κ.

Now, for L large enough, $\partial^r \xi_N(b;z,I)/\partial b^r = \partial^r \xi_0(b;z,I)/\partial b^r$ for $r \geq 0$ and $(b,z,I) \in [\underline{v}_0, b_*] \times \mathcal{Z} \times \mathcal{I}$, while for $r = 0, \ldots, R+1$ and $(b,z,I) \in [b^*, \overline{b}_0(z,I)] \times \mathcal{Z} \times \mathcal{I}$,

$$\begin{aligned}
&\left| \frac{\partial^r \xi_N(b;z,I)}{\partial b^r} - \frac{\partial^r \xi_0(b;z,I)}{\partial b^r} \right| \\
&\leq \left| \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_N(b|z,I)}{(I-1)g_1(b|z,I)}; \theta_N \right) \right] - \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)}; \theta_N \right) \right] \right| \\
&\quad + \left| \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)}; \theta_N \right) \right] - \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)}; \theta_0 \right) \right] \right| \\
&\leq \sup_{\theta\in\Theta} \left| \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_1(b|z,I)}{(I-1)g_1(b|z,I)}; \theta \right) \right] - \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)}; \theta \right) \right] \right| \\
&\quad + \left| \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)}; \theta_N \right) \right] - \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)}; \theta_0 \right) \right] \right|. \quad \text{(C.3)}
\end{aligned}$$

The difference in the sup term is a difference of polynomials in the variables $1/g_j(b|z,I)$, $g_j^{(k)}(b|z,I)$, $G_j^{(k)}(b|z,I)$ and $\partial^k \lambda^{-1}(\cdot;\theta_j)/\partial x^k$ evaluated at $G_j(b|z,I)/[(I-1)g_j(b|z,I)]$, for $k = 0, \ldots, r$. Therefore, for $r = 0, \ldots, R+1$ the first term in (C.3) is of order $\kappa^r O\left(\rho_N^{(r-(R+1))/(R+1)}\right)$ uniformly on $\cup_{\mathcal{Z}\times\mathcal{I}}[b_*, \overline{b}_0(z,I)] \times \{z,I\}$ by (C.1)-(C.2). Regarding the second term in (C.3), note that $\lambda^{-1}(\cdot;\cdot)$ is $R+1$ continuously differentiable on $[0,+\infty) \times \Theta$ because $\lambda(\cdot;\cdot)$ is $R+1$ continuously differentiable on $[0,+\infty) \times \Theta$ by Assumption A2-(i) and $\lambda'(\cdot;\theta) \geq 1$. Thus, because $G_0(\cdot|\cdot,I)/g_0(\cdot|\cdot;I)$ is $R+1$ continuously differentiable on $S_I(G_0)$, the function $\partial^r \lambda^{-1}\{G_0(b|z,I)/[(I-1)g_0(b|z,I)];\theta\}/\partial b^r$ is continuous on $S_I(G_0) \times \Theta$. Hence, the second term is of order $\|\theta_N - \theta_0\|_\infty = O(1/\rho_N) = o(1)$ uniformly on $S_I(G_0)$, for $r = 0, \ldots, R+1$, because $\rho_N \to \infty$. Collecting results and using the finiteness of $\mathcal{I}$, (C.3) yields

$$\sup_{I\in\mathcal{I}} \sup_{(b,z)\in S_I(G_0)} \left| \frac{\partial^r \xi_N(b;z,I)}{\partial b^r} - \frac{\partial^r \xi_0(b;z,I)}{\partial b^r} \right| = \kappa^r O\left(\rho_N^{\frac{r-(R+1)}{R+1}} \right) + o(1), \ r = 0, \ldots, R+1. \quad \text{(C.4)}$$

Step 3: *Proof of (i).* Because $R \geq 1$, applying (C.4) for $r = 1$ yields that $\xi_N'(\cdot;z,I) > 0$ on $[\underline{v}_0, \overline{b}_0(z,I)]$ for every $(z,I) \in \mathcal{Z} \times \mathcal{I}$. Because $G_N(\cdot|z,I)$ satisfies Definition 3, Lemma 1 shows that $G_N(\cdot|z,I)$ is rationalized by $[U(\cdot;\theta_N), F_N(\cdot|z,I)]$ for every $(z,I) \in \mathcal{Z}\times\mathcal{I}$, where $F_N(\cdot|z,I)$ is the distribution of $\xi_N(b;z,I)$ with $b \sim G_N(\cdot|z,I)$. It remains to verify that $F_N(\cdot|\cdot,\cdot) \in \mathcal{F}_R^*$. The support of $F_N(\cdot|z,I)$ is the same as $F_0(\cdot|z,I)$, namely $[\underline{v}_0, \overline{v}_0]$ since $\underline{v}_N(z,I) = \xi_1(\underline{v}_0; z, I) = \underline{v}_0$, while $\overline{v}_N(z,I) = \xi_N(\overline{b}_0(z,I); z, I)$ together with $\psi(0) = 1$ give

$$\begin{aligned}
\overline{v}_N(z,I) &= \overline{b}_0(z,I) + \lambda^{-1}\left(\frac{1}{I-1}\frac{1}{g_0(\overline{b}_0(z,I)|z,I) + m(z,I;\beta_N) - m(z,I;\beta_0)}; \theta_N\right) \\
&= \overline{b}_0(z,I) + \lambda^{-1}\left(\frac{1}{I-1}\frac{1}{m(z,I;\beta_N)}; \theta_N\right) \\
&= \overline{b}_0(z,I) + \lambda^{-1}\left(\lambda\left(\bar{v}_0 - \overline{b}_0(z,I);\theta_N\right); \theta_N\right) = \overline{v}_0,
\end{aligned}$$

because $g_0(\overline{b}_0(z,I)|z,I) = m(z,I;\beta_0)$ by (6) and (7). Also, it can be seen that $F_N(\cdot|\cdot,I)$ is $R+1$ continuously differentiable on $[\underline{v}_0, \overline{v}_0] \times \mathcal{Z}$ with a density satisfying Definition 4-(iii) in view of the properties of $\xi_N(\cdot;\cdot,I)$ and $G_N(\cdot|\cdot,I)$ as $F_N(v|z,I) = G_N(\xi_N^{-1}(v;z,I)|z,I)$.

Step 4: *Proof of (ii).* Let $s_j(v;z,I) = \xi_j^{-1}(v;z,I)$. Using the same argument as in Step 2 of the proof of Lemma B1 in Guerre, Perrigne, Vuong (2000), it follows that $s_N(v;z,I) - s_0(v;z,I)$ also satisfies (C.4) with $\xi_j(b;z,I)$ replaced by $s_j(v;z,I)$, for $j = 0, 1$. Now, $f_N(v|z,I) = g_N(s_1(v;z,I)|z,I)s'_N(v;z,I)$. Thus, following Step 3 of that lemma we obtain

$$\sup_{I\in\mathcal{I}} \sup_{(v,z)\in[\underline{v}_0,\overline{v}_0]\times\mathcal{Z}} \left|\frac{\partial^r f_N(v|z,I)}{\partial v^r} - \frac{\partial^r f_0(v|z,I)}{\partial v^r}\right| = \kappa^{r+1} O\left(\rho_N^{\frac{r-R}{R+1}}\right) + o(1), \quad r = 0, 1, \ldots, R.$$

Letting $r = 0$ we have $\|f_N(\cdot|\cdot,\cdot) - f_0(\cdot|\cdot,\cdot)\|_\infty < \epsilon$ for N sufficiently large. Moreover, for $r = R$ the triangular inequality gives $\|\partial^R f_N(\cdot|\cdot,\cdot)/\partial v^R\|_\infty < M$ as $\|\partial^R f_0(\cdot|\cdot,\cdot)/\partial v^R\|_\infty < M$, provided κ is sufficiently small. Because $\|\beta_N - \beta_0\|_\infty < \epsilon$ for N large enough, the desired result follows.

Proof of Lemma B3: (i) The variables $Y_{i\ell}$ are independent given $\mathcal{F}_L$ because the $V_{i\ell}$s (and then the $B_{i\ell}$s) are independent given $\mathcal{F}_L$. The same property holds for $\epsilon_{i\ell} = Y_{i\ell} - \mathrm{E}[Y_{i\ell}|\mathcal{F}_L]$.

(ii) The proof is standard. We have $0 < h_N \leq \inf_{(z,I)\in\mathcal{Z}\times\mathcal{I}}(\overline{b}_0(z,I) - \underline{b}_0(z,I))$ by A4-(iv), $\mathcal{I}$ finite and Lemma 2-(i). Thus, by Lemma 2-(iv) a Taylor expansion of order $R+1$ gives

$$\begin{aligned}
&\mathrm{E}[Y_{i\ell}|\mathcal{F}_L] \\
&= \frac{1}{h_N}\int_{\underline{b}_0(Z_\ell,I_\ell)}^{\overline{b}(Z_\ell,I_\ell)} \Phi\left(\frac{b - \overline{b}_0(Z_\ell,I_\ell)}{h_N}\right) g_0(b|Z_\ell,I_\ell)db = \int_{-\infty}^{0} \Phi(x) g_0(\overline{b}_0(Z_\ell,I_\ell) + h_N x|Z_\ell,I_\ell)dx \\
&= \int \Phi(x)\left[\sum_{r=0}^{R+1} g_0^{(r)}(\overline{b}_0(Z_\ell,I_\ell)|Z_\ell,I_\ell)\frac{(h_N x)^r}{r!} + o\left((h_N x)^{R+1}\right)\right]dx \\
&= g_0(\overline{b}_0(Z_\ell,I_\ell)|Z_\ell,I_\ell) + \frac{h_N^{R+1}}{(R+1)!}\left(\frac{\partial^{R+1} g_0(\overline{b}_0(Z_\ell,I_\ell)|Z_\ell,I_\ell)}{\partial b^{R+1}}\int x^{R+1}\Phi(x)dx + o(1)\right)
\end{aligned}$$

using A4-(iii). Then (ii) follows because $g_0(\overline{b}_0(\cdot,I)|\cdot,I) = m(\cdot,I;\beta_0)$ by (8).

(iii) Similarly, using (ii), Lemma 2 and $h_N \to 0$, we have

$$
\begin{aligned}
\mathrm{Var}[Y_{i\ell}|\mathcal{F}_L] &= \frac{1}{h_N^2}\int_{-\infty}^{\overline{b}_0(Z_\ell,I_\ell)} \Phi^2\left(\frac{b-\overline{b}_0(Z_\ell,I_\ell)}{h_N}\right) g_0(b|Z_\ell,I_\ell)db - \mathrm{E}^2[Y_{i\ell}|\mathcal{F}_L] \\
&= \frac{1}{h_N}\int_{-\infty}^{0} \Phi^2(x) g_0(\overline{b}_0(Z_\ell,I_\ell)+h_N x|Z_\ell,I_\ell)dx + o\left(\frac{1}{h_N}\right) \\
&= g_0(\overline{b}_0(Z_\ell,I_\ell)|Z_\ell,I_\ell)\frac{1+o(1)}{h_N}\int \Phi^2(x)dx, \\
\max_{i,\ell}|\epsilon_{i\ell}| &= \frac{1}{h_N}\max_{i,\ell}\left|\Phi\left(\frac{B_{i\ell}-\overline{b}_0(Z_\ell,I_\ell)}{h_N}\right) - \mathrm{E}\left[\Phi\left(\frac{B_{i\ell}-\overline{b}_0(Z_\ell,I_\ell)}{h_N}\right)|\mathcal{F}_L\right]\right| \le \frac{2\sup_{x\in I\!R}|\Phi(x)|}{h_N}.
\end{aligned}
$$

Proof of Lemma B4: It suffices to check the Lyapounov condition of Theorem 7.3 in Billingsley (1968) given $\mathcal{F}_L$. From Lemma B3-(iii), we have

$$
|u_{i\ell}\varepsilon_{i\ell}| \le \frac{2\|u\|_\infty \sup_{x\in I\!R}|\Phi(x)|}{h_N}, \qquad \mathrm{Var}^{1/2}\left[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell} u_{i\ell}\epsilon_{i\ell}|\mathcal{F}_L\right] \asymp \|u\|_2/\sqrt{h_N}.
$$

Thus, $\mathrm{E}[|u_{i\ell}\epsilon_{i\ell}|^3|\mathcal{F}_L] \le \mathrm{E}[|u_{i\ell}\epsilon_{i\ell}|^2|\mathcal{F}_L]O(\|u\|_\infty/h_N)$. Hence, by independence of the $\epsilon_{i\ell}$s given $\mathcal{F}_L$ and $\mathrm{E}[\epsilon_{i\ell}|\mathcal{F}_L]=0$, we obtain

$$
\frac{1}{\mathrm{Var}^{3/2}[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell} u_{i\ell}\epsilon_{i\ell}|\mathcal{F}_L]}\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\mathrm{E}[|u_{i\ell}\epsilon_{i\ell}|^3|\mathcal{F}_L] \le \frac{O(\|u\|_\infty/h_N)}{\mathrm{Var}^{1/2}[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell} u_{i\ell}\epsilon_{i\ell}|\mathcal{F}_L]},
$$

which is an $O\left(\|u\|_\infty/[\sqrt{h_N}\|u\|_2]\right)$, and hence an $o_P(1)$ by the assumptions of the lemma.

Proof of Lemma B5: In view of A4-(v), let $\overline{a}_N \asymp a_N$ be such that the event $\mathcal{E}_N = \{\max_\ell|\hat{\overline{b}}(Z_\ell,I_\ell)-\overline{b}_0(Z_\ell,I_\ell)| \le \overline{a}_N\}$ has a probability larger than $1-\epsilon$, where $\epsilon>0$ can be chosen arbitrary small. By A4-(iii), $\Phi(\cdot)$ is continuously differentiable on $I\!R^-$, with support $[-1,0]$. In particular, $Y_{i\ell}=\hat{Y}_{i\ell}=0$ if $B_{i\ell}\le \overline{b}_0(Z_\ell,I_\ell)-\overline{a}_n-h_n$. Note also that $\overline{a}_N < h_N$ for N sufficiently large by A2-(v). In order to bound $\hat{Y}_{i\ell}-Y_{i\ell}$ on $\mathcal{E}_N$, we use a first-order Taylor expansion of $\Phi(\cdot)$ when $\overline{b}_0(Z_\ell,I_\ell)-\overline{a}_n-h_n \le B_{i\ell} \le \overline{b}_0(Z_\ell,I_\ell)-2\overline{a}_n$, while we use the boundedness of $\Phi(\cdot)$ when $\overline{b}_0(Z_\ell,I_\ell)-2\overline{a}_n \le B_{i\ell}\le \overline{b}_0(Z_\ell,I_\ell)$. This gives

$$
\begin{aligned}
|\hat{Y}_{i\ell}-Y_{i\ell}|\,I\!I(\mathcal{E}_N) &= \frac{I\!I(\mathcal{E}_N)}{h_N}\left|\Phi\left(\frac{B_{i\ell}-\hat{\overline{b}}(Z_\ell,I_\ell)}{h_N}\right) - \Phi\left(\frac{B_{i\ell}-\overline{b}_0(Z_\ell,I_\ell)}{h_N}\right)\right| \\
&\le \frac{\overline{a}_N\sup_x|\Phi'(x)|}{h_N^2} I\!I_{[-\bar{a}_N-h_N,-2\overline{a}_N]}\Big(B_{i\ell}-\overline{b}_0(Z_\ell,I_\ell)\Big) + \frac{2\sup_x|\Phi(x)|}{h_N} I\!I_{[-2\overline{a}_N,0]}\Big(B_{i\ell}-\overline{b}_0(Z_\ell,I_\ell)\Big).
\end{aligned}
$$

Let $\zeta_{i\ell}$ denote the right-hand side. Because $\overline{a}_N \asymp a_N = o(h_N)$ by A4-(v), and $\max_{I\in\mathcal{I}} \sup_{(b,z)\in S_I(G_0)} g_0(b|z,I) < \infty$, it is easily seen that

$$\mathrm{E}[\zeta_{i\ell}|\mathcal{F}_L] = O\left(\frac{\overline{a}_N}{h_N}\right) = O\left(\frac{a_N}{h_N}\right), \quad \mathrm{Var}[\zeta_{i\ell}|\mathcal{F}_L] \le \mathrm{E}[\zeta_{i\ell}^2|\mathcal{F}_L] = O\left(\frac{\overline{a}_N^2}{h_N^3}\right) + O\left(\frac{\overline{a}_N}{h_N^2}\right) = O\left(\frac{a_N}{h_N^2}\right).$$

By independence of the $\zeta_{i\ell}$s given $\mathcal{F}_L$, this gives

$$\begin{aligned}
\mathrm{E}\left[\left(\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}|u_{i\ell}|\zeta_{i\ell}\right)^2 \mid \mathcal{F}_L\right] &= \mathrm{E}^2\left[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}|u_{i\ell}|\zeta_{i\ell} \mid \mathcal{F}_L\right] + \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell} u_{i\ell}^2 \mathrm{Var}[\zeta_{i\ell} \mid \mathcal{F}_L] \\
&= O\left[\left(\|u\|_1 \frac{a_N}{h_N}\right)^2 + \|u\|_2^2 \frac{a_N}{h_N^2}\right], \\
\mathrm{E}\left[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}|u_{i\ell}|\zeta_{i\ell}^2 \mid \mathcal{F}_L\right] &= \|u\|_1 O\left(\frac{a_N}{h_N^2}\right).
\end{aligned}$$

Using $E[|X|] \le E^{1/2}[X^2]$, Markov inequality given $\mathcal{F}_L$ and $\Pr(\mathcal{E}_N) \ge 1-\epsilon$ completes the proof.

Proof of Lemma B6: By definition of $\mathcal{B}_\delta$, we have $\delta \le \overline{v} - \overline{b}_0(Z_\ell, I_\ell) \le \overline{v}_{\sup}$ for all ℓ. By A4-(v), we have $|\hat{\overline{b}}(Z_\ell, I_\ell) - b_0(Z_\ell, I_\ell)| < \delta/2$ for all ℓ with probability approaching one. Thus, $\delta/2 \le \overline{v} - \hat{\overline{b}}(Z_\ell, I_\ell) \le \overline{v}_{\sup} + \delta/2$ for all ℓ with probability approaching one. Now,

$$\hat{m}(Z_\ell, I_\ell; \beta) - m(Z_\ell, I_\ell; \beta) = \frac{1}{I_\ell - 1}\left(\frac{1}{\lambda\left(\overline{v} - \hat{\overline{b}}(Z_\ell, I_\ell); \theta\right)} - \frac{1}{\lambda\left(\overline{v} - \overline{b}_0(Z_\ell, I_\ell); \theta\right)}\right).$$

Hence, the denominators are uniformly bounded away from 0 with probability approaching one. The desired result follows from A2-(i) since $\lambda(x;\theta)$ is uniformly continuous on the compact $[\delta/2, \overline{v}_{\sup} + \delta/2] \times \Theta$. The study of the derivatives is similar.

Proof of Lemma B7: By A3-(i) and A4-(ii), there exists some $\underline{C} > 0$ such that $\overline{Q}(\beta) \ge \underline{C}\sum_{I\in\mathcal{I}}\int_{\mathcal{Z}}[m(z,I;\beta) - m(z,I;\beta_0)]^2\,dz$. Thus, $\overline{Q}(\beta) = 0$ is equivalent to $m(z,I;\beta) = m(z,I;\beta_0)$ for all $(z,I) \in \mathcal{Z}\times\mathcal{I}$ by continuity of $m(\cdot,I;\beta) - m(\cdot,I;\beta_0)$ whenever $\beta \in \mathcal{B}_\delta$ in view of Lemma B1 and $\mathcal{B}_\delta \subset \mathcal{B}$. Using (7), this is equivalent to $\lambda[\overline{v} - \overline{b}_0(z,I);\theta] = \lambda[\overline{v}_0 - \overline{b}_0(z,I);\theta_0]$ for all $(z,I) \in \mathcal{Z}\times\mathcal{I}$. Hence, $\beta = \beta_0$ by A2-(iii). Therefore, $\overline{Q}(\beta) = 0$ if and only if $\beta = \beta_0$. Moreover, Lemma B1 yields that $\overline{Q}(\cdot)$ is continuous on $\mathcal{B}_\delta$ and hence on $\mathcal{B}_\delta \cap \{\|\beta - \beta_0\| \ge \epsilon\}$ by the Lebesgue Dominated Convergence Theorem. This implies the first claim as $\mathcal{B}_\delta$ is compact.

Next, consider $A(\beta)$ as $B(\beta)$ is treated similarly. The Lebesgue Dominated Convergence Theorem and Lemma B1 yield that $\det A(\beta)$ is continuous in $\beta \in \mathcal{B}_\delta$. Thus, it is sufficient to show that $A(\beta_0)$ is of full rank. Suppose not, then there exists $t \in I\!R^{p+1}\setminus\{0\}$ with $t'A(\beta_0)t = 0$. From

(7) we have $\partial m(z, I; \beta)/\partial\beta = -(I-1)m^2(z, I; \beta)\partial\lambda[\overline{v} - \overline{b}_0(z, I); \theta]/\partial\beta$, with $m(z, I; \beta) \geq \underline{m} > 0$ for all $\beta \in \mathcal{B}_\delta$, $z \in \mathcal{Z}$, $I \in \mathcal{I}$ because $m(\cdot, I; \cdot)$ does not vanish and is continuous on $\mathcal{Z} \times \mathcal{B}_\delta$ by Lemma B1. Thus, arguing as for $\overline{Q}(\cdot)$ and letting $\beta = \beta_0$ give

$$0 = t'A(\beta_0)t \geq \underline{C} \sum_{I \in \mathcal{I}} \int_{\mathcal{Z}} \left(t' \cdot \frac{\partial\lambda\left(\overline{v}_0 - \overline{b}_0(z, I); \theta_0\right)}{\partial\beta} \right)^2 dz,$$

which implies $t' \cdot \partial\lambda[\overline{v}_0 - \overline{b}_0(z, I); \theta_0]/\partial\beta = 0$ for all $(z, I) \in \mathcal{Z} \times \mathcal{I}$, thereby contradicting A4-(i).

Proof of Lemma B8: Observe that the function $I_\ell \omega(Z_\ell, I_\ell)\left[m(Z_\ell, I_\ell; \beta) - m(Z_\ell, I_\ell; \beta_0)\right]^2$ is a Lipschitz function with respect to $\beta \in \mathcal{B}_\delta$ with a Lipschitz constant that can be chosen independently of $(\beta, \beta_0, Z_\ell, I_\ell)$ by Lemma B1 and the compactness of $\mathcal{B}_\delta \times \mathcal{Z} \times \mathcal{I}$. The first statement of the lemma with the order $O_P(1/\sqrt{L})$ follows from the maximal inequality (19.36) in van der Vaart (1998) upon computing the bracketing number of the class of functions $\{[m(\cdot, \cdot; \beta) - m(\cdot, \cdot; \beta_0)]^2; \beta \in \mathcal{B}_\delta\}$ on $\mathcal{Z} \times \mathcal{I}$. See Example 19.7 in van der Vaart (1998). The other statements of the lemma are direct consequences of the Lindeberg-Levy Central Limit Theorem since $L \asymp N$ and $N/L = \mathrm{E}[I] + o_P(1)$, writing for instance

$$\frac{A_N(\beta)}{N} = \left(\frac{N}{L}\right)^{-1} \frac{1}{L} \sum_{\ell=1}^{L} I_\ell \omega(Z_\ell, I_\ell) \frac{\partial m(Z_\ell, I_\ell; \beta)}{\partial\beta} \cdot \frac{\partial m(Z_\ell, I_\ell; \beta)}{\partial\beta'}.$$

Proof of Lemma B9: Define $w_{\ell L}(\beta) = \sqrt{h_N} \sum_{i=1}^{I_\ell} \omega(Z_\ell, I_\ell; \beta) m(Z_\ell, I_\ell; \beta)\epsilon_{i\ell}$ so that $W_N(\beta) = (L/N)^{1/2} L^{-1/2} \sum_{\ell=1}^{L} w_{\ell L}(\beta)$ with $w_{\ell L}$ i.i.d. within rows. Because L/N is bounded, it suffices to show that $\sup_{\beta \in \mathcal{B}_\delta} \left| L^{-1/2} \sum_{\ell=1}^{L} w_{\ell L}(\beta) \right| = O_P(1)$. Using A4-(ii), Lemmas B1, B3-(iii) and the compactness of $\mathcal{Z} \times \mathcal{I} \times \mathcal{B}_\delta$, there exists some constant $C > 0$ such that $\max_{1 \leq \ell \leq L} \sup_{\beta \in \mathcal{B}_\delta} |w_{\ell L}(\beta)| \leq C/\sqrt{h_N}$, $\max_{1 \leq \ell \leq L} \sup_{\beta \in \mathcal{B}_\delta} \mathrm{Var}[w_{\ell L}(\beta)] \leq C$, and $\max_{1 \leq \ell \leq L} \mathrm{E}^{1/2}[w_{\ell L}(\beta) - w_{\ell L}(\beta')]^2 \leq C\|\beta - \beta'\|$ for all L. This is sufficient to apply the maximal inequality of Lemma 19.36 in van der Vaart (1998) with bracketing number as in Example 19.7. This gives

$$\mathrm{E}\left[\sup_{\beta \in \mathcal{B}_\delta} \left| L^{-1/2} \sum_{\ell=1}^{L} w_{\ell L}(\beta) \right|\right] \leq C'\left(1 + \frac{C''}{\sqrt{L h_N}}\right) = O(1),$$

where C' and C'' are positive constants independent of L. Applying Markov inequality, this gives $\sup_{\beta \in \mathcal{B}_\delta} |W_N(\beta)| = O_P(1)$. A similar proof establishes that $\sup_{\beta \in \mathcal{B}_\delta} |W_N^{(1)}(\beta)| = O_P(1)$. Because $\tilde{\beta}_N \in \mathcal{B}_\delta$ with probability approaching one as $\tilde{\beta}_N \xrightarrow{P} \beta_0 \in \mathcal{B}_\delta$, we obtain $W_N^{(1)}(\tilde{\beta}_N) = o_P(1)$.

References

Athey, S. (2001): "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information," *Econometrica*, 69, 861-889.

Athey, S. and **P. Haile** (2006): "Nonparametric Approaches to Auctions," in J. Heckman and E. Leamer, eds., *Handbook of Econometrics, Volume VI*, Amsterdam: North Holland.

Athey, S. and **J. Levin** (2001): "Information and Competition in U.S. Forest Service Timber Auctions," *Journal of Political Economy*, 109, 375-417.

Athey, S., **J. Levin** and **E. Seira** (2004): "Comparing Open and Sealed Bid Auctions: Theory and Evidence from Timber Auctions," Working Paper, Stanford University.

Bajari, P. and **A. Hortcsu** (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., **R. Marshall** and **J.F. Richard** (1997): "Bidder Collusion at Forest Service Timber Sales," *Journal of Political Economy*, 105, 657-699.

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

Billingsley, P. (1968): *Convergence of Probability Measures*, Wiley.

Bickel, P.J., **C.A.J. Klaassen**, **Y. Ritov** and **J.A. Wellner** (1993): *Efficient and Adaptive Estimation for Semiparametric Models*, Johns Hopkins University Press.

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., **I. Perrigne** and **Q. Vuong** (2003): "Asymmetry in First-Price Auctions with Affiliated Private Values," *Journal of Applied Econometrics,* 18, 197-207.

Che, Y.K. and **I. Gale** (1998): "Standard Auctions with Financially Constrained Bidders," *Review of Economic Studies*, 65, 1-21.

Donald, S. and **H. Paarsch** (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 **L. White** (2004): "Precautionary Bidding in Auctions," *Econometrica*, 72, 77-92.

Goeree, J., **C. Holt** and **T. Palfrey** (2002): "Quantal Response Equilibrium and Overbidding in Private Value Auctions," *Journal of Economic Theory*, 104, 247-272.

Gollier, C. (2001): *The Economics of Risk and Time*, MIT Press.

Guerre, E., **I. Perrigne**, and **Q. Vuong** (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68, 525-574.

Guerre, E., **I. Perrigne**, and **Q. Vuong** (2005): "Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions," Working Paper, Pennsylvania State University.

Guerre, E. and **Q. Vuong** (2006): "Existence, Uniqueness and Smoothness of the Bayesian Nash Equilibirum Strategy in First-Price Sealed-Bid Auctions with Risk Averse Bidders," Working Paper, Pennsylvania State University.

Haile, P. (1996): "Auctions with Resale Markets: An Application to U.S. Forest Service Timber Sales," Working Paper, University of Wisconsin.

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

Hansen, R. (1985): "Empirical Testing of Auction Theory," *American Economic Review*, 75, 156-159.

Härdle, W. and **O. Linton** (1994): "Applied Nonparametric Methods," in R.F. Engle and D.L. McFadden, eds., *Handbook of Econometrics, Volume IV*, Amsterdam: North Holland.

Hengartner, N.W. (1997): "Asymptotic Unbiased Density Estimators," Working Paper, Yale University.

Honoré, B.E. and **E. Kyriazidou** (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.

Horowitz, J.L. (1993): "Optimal Rates of Convergence of Parameter Estimators in the Binary Response Model with Weak Distributional Assumptions," *Econometric Theory*, 9, 1-18.

Ibragimov, I.A. and **R.Z. Has'minskii** (1981): *Statistical Estimation. Asymptotic Theory*, New York: Springer Verlag.

Korostelev, A.P. and **A.B. Tsybakov** (1993): *Minimax Theory of Image Reconstruction, Lecture Notes in Statisics*, 82, New York: Springer Verlag.

Krishna, V. (2002): *Auction Theory*, Academic Press.

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

Laffont, J.J., **H. Ossard** and **Q. Vuong** (1995): "Econometrics of First-Price Auctions," *Econometrica*, 63, 953-980.

Laffont, J.J. and **Q. Vuong** (1996): "Structural Analysis of Auction Data," *American Economic Review, Papers and Proceedings,* 86, 414-420.

Lebrun, B. (1999): "First-Price Auctions and the Asymmetric N Bidders Case," *International Economic Review*, 40, 135-142.

Li, T., **I. Perrigne** and **Q. Vuong** (2000): "Conditionally Independent Private Information in OCS Wildcat Auctions," *Journal of Econometrics*, 98(1), 129-161.

Li, T., **I. Perrigne** and **Q. Vuong** (2002): "Structural Estimation of the Affiliated Private Value Auction Model," *Rand Journal of Economics*, 33, 171- 193.

Liu, R.C. and **L.D. Brown** (1993): "Nonexistence of Informative Unbiased Estimators in Singular Problems," *Annals of Statistics*, 21, 1-13.

Lu, J. (2002): "Stochastic Private Values in Auctions: Identification and Estimation," Working Paper, University of Southern California.

Lu, J. and **I. Perrigne** (2005): "Estimating Risk Aversion from Ascending and Sealed-Bid Auctions: The Case of Timber Auction Data," Working Paper, Pennsylvania State University.

Manski, C.F. (1985): "Semiparametric Analysis of Discrete Response: Asymptotic Properties of the Maximum Score Estimator," *Journal of Econometrics*, 27, 205-228.

Maskin, E. and **J. Riley** (1984): "Optimal Auctions with Risk Averse Buyers," *Econometrica*, 52, 1473-1518.

Maskin, E. and **J. Riley** (2003): "Uniqueness in Sealed High Bid Auctions," *Games and Economic Behavior*, 45, 395-409.

Maskin, E. and **J. Riley** (2000a): "Asymmetric Auctions," *Review of Economic Studies*, 67, 413-438.

Maskin, E. and **J. Riley** (2000b): "Equilibrium in Sealed High Bid Auctions," *Review of Economic Studies*, 67, 439-454.

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

Milgrom, P. and **R. Weber** (1982): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50, 1089-1122.

Newey, W.K. and **D.L. McFadden** (1994): "Large Sample Estimation and Hypothesis Testing," in R.F. Engle and D.L. McFadden, eds., *Handbook of Econometrics, Volume IV*, Amsterdam: North Holland.

Paarsch, H. (1992): "Deciding between the Common and Private Value Paradigms in Empirical Models of Auctions," *Journal of Econometrics*, 51, 191-215.

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

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

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

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.

van der Vaart, A. (1998): *Asymptotic Statistics*, Cambridge University Press.

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

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

Table 1: Summary Statistics

Variable	Mean	STD
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

Table 2: Estimation Results

	Constant	Linear	Quadratic
CRRA	$\hat{\theta}$ =0.0199	$\hat{\theta}$ =0.6813	$\hat{\theta}$ =0.6797
	(0.0015)	*(0.0317)*	*(0.0104)*
	$\hat{\gamma}_0$ =78.6576	$\hat{\gamma}_0$ =10.6115	$\hat{\gamma}_0$ =10.4648
	(0.0267)	*(0.8498)*	*(0.6340)*
		$\hat{\gamma}_1$ =5.2266	$\hat{\gamma}_1$ =5.2750
		(0.1157)	*(0.1983)*
			$\hat{\gamma}_2$ =-0.0031
			(0.0115)
	SSE=2.1950	SSE= 1.6138	SSE= 1.6107
CRRA with wealth	$\hat{\theta}$ =0.1565	$\hat{\theta}$ =0.7331	$\hat{\theta}$ =0.7305
	(0.1081)	*(0.1098)*	*(0.0998)*
	$\hat{\gamma}_0$ =82.5494	$\hat{\gamma}_0$ =10.8534	$\hat{\gamma}_0$ =10.7060
	(0.2772)	*(3.0051)*	*(2.3924)*
		$\hat{\gamma}_1$ =5.2164	$\hat{\gamma}_1$ =5.2650
		(0.0472)	*(0.1642)*
			$\hat{\gamma}_2$ =-0.0031
			(0.0097)
	$\hat{w} = 10^{-10}$	$\hat{w} = 10^{-10}$	$\hat{w} = 10^{-10}$
	(1.8×10^{-8})	*(0.0075)*	*(0.0060)*
	SSE= 2.1504	SSE= 1.5487	SSE=1.5431
CARA	No Convergence	$\hat{\theta}$ =0.0001	$\hat{\theta}$ =0.00002
		(0.0024)	*(0.0034)*
		$\hat{\gamma}_0$ =9.8972	$\hat{\gamma}_0$ =11.9472
		(0.7086)	*(1.4472)*
		$\hat{\gamma}_1$ =7.4882	$\hat{\gamma}_1$ =2.2882
		(0.2110)	*(2.2140)*
			$\hat{\gamma}_2$ =1.1219
			(0.4602)
		SSE=1.3616	SSE=1.3431