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

Sandra Campo
University of North Carolina

Emmanuel Guerre
LSTA-University of Paris VI and CREST-LS

Isabelle Perrigne
University of Southern California

Quang Vuong
University of Southern California

November 2002

* We are grateful 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.
Correspondence to: I. Perrigne, Department of Economics, University of Southern California, Los Angeles, CA 90089-0253. e-mail address : perrigne@usc.edu.

Abstract

This paper proposes a semiparametric estimation procedure of the first-price auction model with risk averse bidders within the independent private value paradigm. We show that the model is nonidentified in general from observed bids. We then exploit heterogeneity across auctioned objects to establish semiparametric identification under a conditional quantile restriction and parameterization of the bidders' von Neuman Morgenstern utility function. Next we propose a semiparametric method for estimating the corresponding auction model. This method involves several steps and allows to recover the parameters of the utility function as well as the bidders' private values and their density. We show that our semiparametric estimator of the utility function parameters converges at the optimal rate, which is slower than the parametric one. An illustration of the method on U.S. Forest Service timber sales is presented and a test of bidders' risk neutrality is performed.

Key words: Risk Aversion, Independent 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 situations under uncertainty. This is the case in auctions where bidders must deal with various types of uncertainties related to the value of the auctioned object, the strategies used by the other bidders and the private information of their opponents. In particular, a pervasive economic argument for justifying bidder's risk aversion is that the value of the auctioned object is high relative to his assets. Another argument is that a bidder does not have many alternatives for buying the object other than in the auction. On the other hand, many important results in the theory of auctions crucially depend on the assumption of risk neutrality. For instance, within the independent private value (IPV) paradigm, the revenue equivalence theorem established by Vickrey (1961) states that English, descending, first-price sealed-bid and second-price auctions lead to the same expected revenue for the seller provided bidders are risk neutral. Such an important result no longer holds when bidders are risk averse. See Harris and Raviv (1981) and Riley and Samuelson (1981).[1]

Despite the importance of risk aversion in auction modeling, very few empirical studies have attempted to assess the extent of bidders' risk aversion on field data. See Baldwin

[1]Likewise, the optimal auction mechanism is quite involved under bidders' risk aversion as it requires complex transfers among bidders. See Maskin and Riley (1984) and Matthews (1987). For a recent survey of auction theory, see Klemperer (1999).

(1995) and Athey and Levin (2001) using US Forest Service timber auctions. A reason is that economic theory provides few reduced-form implications of risk aversion, which are moreover difficult to test on field data.[2] An alternative approach is to consider that the observed bids are precisely the outcomes of the Bayesian Nash equilibrium of the underlying auction game. This is known as the structural approach, which has been developed by Paarsch (1992) and Laffont, Ossard and Vuong (1995). For a recent survey, see Perrigne and Vuong (1999). 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).

Throughout, we consider first-price sealed-bid auctions with risk averse bidders within the IPV paradigm. From Maskin and Riley (1984) it is known that a first-price auction generates a larger revenue than an ascending auction when bidders are risk averse. Thus, bidders' risk aversion can provide a rationale for the use of first-price auctions.[3] A first part of our paper briefly presents the model and reviews the existence, uniqueness and smoothness of the Bayesian Nash equilibrium strategy. Existence and uniqueness of such a strategy follow from Maskin and Riley (1996, 2000) among others. In particular, these properties are difficult to establish when the reserve price is nonbinding because of a well-known singularity of the differential equation characterizing the equilibrium strategy. In addition to providing an alternative proof of such properties, we derive the smoothness of the equilibrium strategy with respect to the bidder's private value as well as potential exogenous variables characterizing the auctioned object.

A second part is devoted to the identification of the auction model, i.e. whether its structural elements can be uniquely recovered from observed bids. The structural elements are the bidders' utility function and the bidders' private value distribution. Unlike Donald and Paarsch (1996) who consider a constant relative risk aversion (CRRA), we consider a

[2]In contrast, the experimental literature has paid much attention to bidders' risk aversion as it can frequently explain observed overbidding (above the risk neutral Bayesian Nash equilibrium). See Cox, Smith and Walker (1988) and Goere, Holt and Palfrey (2002) among others.

[3]Hence, the development of econometric methods assessing the extent of risk aversion is especially important in first-price auctions. In many situations, empirical researchers have observed the exclusive use of a particular mechanism for a large variety of goods. For instance, ascending auctions are used for art and memorabilia, while first-price auctions are used for procurements and natural resources except for timber, which is sold through both mechanisms.

general von Neuman Morgenstern (vNM) utility function exhibiting possibly risk aversion. First, we show that this general model is nonidentified from observed bids even when the utility function is restricted to belong to well known families of risk aversion. In particular, restricting bidders to have a constant relative risk aversion is not sufficient to achieve identification. 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 a CRRA model and a CARA model cannot be discriminated against each other. It also implies that the game theoretical model does not impose testable restrictions on bids. Furthermore, one can consider either CRRA or CARA utility functions without loss of power for explaining observed bids, despite either model not being identified.

In view of the preceding results, a third part of our paper seeks weak and palatable restrictions that can be used to achieve identification of the auction model with risk averse bidders. Specifically, we exploit heterogeneity across auctioned objects under the assumption that the private value distribution conditional upon the characteristics of the auctioned objects satisfies a parametric quantile restriction. Unlike previous work such as Guerre, Perrigne and Vuong (2000), Li, Perrigne and Vuong (2000, 2002), Campo, Perrigne and Vuong (2002), such an additional restriction was not necessary to identify the auction models considered there. Of course, there are other possible restrictions such as requiring that some quantile be known, but the latter assumption is unattractive as the postulated value of some quantile directly affects the estimated degree of risk aversion. We then restrict the bidders' vNM utility function to be parametric. Under these conditions, we show that the utility function parameters and the conditional private value distribution of the model with risk adverse bidders are semiparametrically identified as no parametric assumption on the latent private value distribution is made. As a matter of fact, we show that these two identifying conditions are necessary as dropping either one of them looses identification. In this sense, our semiparametric modeling is natural, while constituting a new direction for the literature on 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 parameters of the utility function. Given the semi-parametric nature of our model, it is important to study the best (optimal) rate that an estimator of the risk aversion parameters can achieve. To do so, we rely on the minimax theory developed by e.g. Ibragimov and Has'minskii (1981). As is well-known, estimation of the upper boundary of a distribution can be achieved at a faster rate than any other quantile. For this reason, we focus on a parametric restriction on the upper boundary of the private value distribution to achieve a faster convergence rate and a greater precision for the estimator. Specifically, 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 parameters is $N^{(R+1)/(2R+3)}$. Such a rate is independent of the dimension d of heterogeneity and is slower than $\sqrt{N}$, which is unattainable given the assumed smoothness R.

A fifth part of the paper addresses the estimation of the structural elements, i.e. the parameters of the vNM utility function and the underlying conditional density of bidders' private values. We then develop a multistep semiparametric estimation procedure. A first step consists in estimating the conditional density of bids at its upper boundary. This involves nonparametric estimation of the upper boundary using Korostelev and Tsybakov (1993) theory of image reconstruction as well as the corresponding conditional densities at these upper bounds. A second step focuses on the estimation of the utility function parameters exploiting the fundamental equation of the auction model. This leads to (possibly weighted) nonlinear least squares (NLLS) using the nonparametric estimates obtained in the first step. A third step allows us to recover the bidders' private values and their underlying conditional density following Guerre, Perrigne and Vuong (2000) once the utility function parameters are estimated.

We establish the asymptotic properties of this estimator. In particular, we show that our estimator of the utility function parameters attains the optimal rate $N^{(R+1)/(2R+3)}$. This contrasts with most $\sqrt{N}$-consistent semiparametric estimators developed in the econometric literature, see Powell (1994) for a recent survey.[4] As a matter of fact, standard results on $\sqrt{N}$-consistent semiparametric estimators as given in Newey and McFadden (1994) do not apply. Another notable feature of our estimation problem is that it involves a nonlinear regression model with a bias and a variance that decreases and increases with the smoothing parameter, respectively. This diverging variance of the error term in the equation defining the utility function parameters is the main reason why our

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

semiparametric estimator does not achieve the standard parametric rate.

Lastly, in addition to providing an estimator converging at the optimal rate as well as not requiring a parametric specification of the bidders' private value distribution, our method is computationally simple 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. This is the case when risk aversion is not of the simple CRRA form. 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 fairly risk averse.

The paper is organized as follows. Section 2 briefly presents the model and discusses the properties of the Bayesian Nash equilibrium strategy of the corresponding auction game. Section 3 investigates the identification of the first-price auction model with risk averse bidders and provides general nonidentification results. Understanding of such results leads to additional restrictions used to achieve semiparametric identification of the model, which is the purpose of section 4. Section 5 provides an upper bound for the optimal convergence rate, which can be attained by a semiparametric estimator of the utility function parameters. Section 6 presents our semiparametric estimation procedure with its various steps and statistical properties. Section 7 is devoted to an illustration of our method to timber auction data. Section 8 concludes. Five appendices collect the proofs of our theoretical results.

2 Model and Equilibrium Strategy

This section presents the first-price sealed-bid auction model with risk averse bidders within the IPV paradigm. It also establishes the existence, uniqueness and smoothness of the equilibrium bidding strategy.

The Model

A single and indivisible object is sold through a first-price sealed-bid auction. All sealed bids are collected simultaneously. The object is sold to the highest bidder who pays his bid to the seller. Within the IPV paradigm, each bidder is assumed to know his own private value v_i for the auctioned object but not other bidders' private values. The bidders' private values are drawn independently from a common 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. Moreover, each bidder is potentially risk averse.

Let $U_{vNM}(\cdot)$ be a von Neuman Morgensten utility function common to all bidders with $U'_{vNM}(\cdot) > 0$. Because of potential risk aversion, the utility function is assumed to be weakly concave, i.e. $U''_{vNM}(\cdot) \leq 0$. All bidders have the same initial wealth $w \geq 0$. This gives a utility function of the form $U_{vNM}(w+\cdot)$, where the argument refers to the monetary gain from the auction. All bidders are thus identical *ex ante* and the game is said to be symmetric.[5] Bidder i then maximizes his expected utility $\mathrm{E}\Pi_i = U_{vNM}(w+v_i-b_i)\Pr(b_i \geq b_j, j \neq i) + U_{vNM}(w)[1 - \Pr(b_i \geq b_j, j \neq i)]$ with respect to his bid b_i, where b_j is the jth player's bid.

Bidder i's problem is equivalent to maximizing $[U_{vNM}(w+v_i-b_i) - U_{vNM}(w)]\Pr(b_i \geq b_j, j \neq i)$. Let $U(\cdot) = U_{vNM}(w+\cdot) - U_{vNM}(w)$. Note that $U(\cdot)$ is strictly increasing, weakly concave and satisfies $U(0) = 0$. Hereafter, we consider the maximization of

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

where $U(\cdot)$ satisfies the preceding properties. 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).[6] 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.

It is useful to recall some basic properties of utility functions with risk aversion. Let α and β be arbitrary constants with $\alpha > 0$. A common measure of absolute risk aversion is the ratio $-U''_{vNM}(\cdot)/U'_{vNM}(\cdot)$, which can be constant or nonincreasing. This gives the set

[5] Relaxing the assumption of bidders' common wealth, i.e. letting w_i be bidder i's wealth, will lead to an asymmetric game if the w_is are common knowledge and to a multisignal game if the w_is are private information. Both cases are beyond the scope of this paper. For asymmetric extensions, see Campo (2002). For multisignals, see Che and Gale (1998) for a model with budget constraints.

[6] 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 $w(\cdot)$. Because we use a vNM utility function, $u(-b_i, v_i) = U_{vNM}(w + v_i - b_i)$ and $w(0) = U_{vNM}(w)$. Hence, the utility of loosing with no payment is equal to the utility of winning with payment equal to the bidder's value.

$\mathcal{U}^{CARA}$ of constant absolute risk aversion utility functions and the set $\mathcal{U}^{DARA}$ of decreasing absolute risk aversion utility functions. A well-known form for the former is given by $U_{vNM}(x) = \alpha(1-\exp(-ax))+\beta$, with an absolute measure of risk aversion $a > 0$. Another common measure is the relative risk aversion, which is defined as $-xU''_{vNM}(x)/U'_{vNM}(x)$. This quantity can be either constant or nonincreasing, which gives the set $\mathcal{U}^{CRRA}$ of constant relative risk aversion utility functions and the set $\mathcal{U}^{DRRA}$ of decreasing relative risk aversion utility functions. A well-known characterization for the former is given by $U_{vNM}(x) = \alpha x^{1-c}/(1-c) + \beta$ for $c \geq 0$ and $c \neq 1$ and $U_{vNM}(x) = \alpha \log x + \beta$ for $c = 1$. Relative risk aversion is then measured by c. Note that if initial wealth $w = 0$, then $0 \leq c < 1$ because the utility of loosing the auction would be unbounded otherwise. There exist other families of vNM utility functions exhibiting risk aversion, which have been considered in the literature. See e.g. Gollier (2001). Below we consider mostly the above four families, though our results can apply to other families.

Existence, Uniqueness and Smoothness of the Equilibrium Strategy

¿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.[7] 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.

Maskin and Riley (1984, Theorem 2) prove the existence and uniqueness of $s(\cdot)$ by noting that the differential equation (2) with boundary condition has a unique solution when there is a binding reserve price, i.e. $p_0 > \underline{v}$. In our case, the reserve price is not binding. Consequently, there is a well-known singularity of (2) at the boundary $\underline{v}$, which prevents the use of such an argument for establishing existence and uniqueness of $s(\cdot)$. This is the purpose of the next result, which provides in addition the regularity properties of $s(\cdot)$ used in the following section.

[7] 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.

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 $\lambda^{(r)}(\cdot)$ denotes the rth derivative.

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. For instance, if $U(x) = U_{vNM}(w + x) - U_{vNM}(w)$ with $U_{vNM}(\cdot)$ a suitably normalized CRRA utility function, these conditions are satisfied for $c \geq 1$ when $w > 0$ and for $0 \leq c < 1$ when $w \geq 0$. In either case, $R = \infty$. Similarly, with $U_{vNM}(\cdot)$ a suitable normalized CARA utility function, these conditions are satisfied with $R = \infty$.

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 $0 \leq \underline{v} < \overline{v} < +\infty$,
(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 $\overline{v}$ in (i). Relaxing (i) is possible but would involve more technical aspects in addition to allowing the possibility of 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 $[\underline{v}, \overline{v}]$.

When the reserve price is nonbinding, existence of a pure equilibrium strategy follows from Maskin and Riley (2000) and Athey (2001), while its uniqueness has been established by Maskin and Riley (1996) using an argument similar to Lebrun (1999). The main contribution of Theorem 1 is to derive the smoothness of the equilibrium strategy.

Theorem 1 is an immediate consequence of the more general Theorem A1 in Appendix A, which allows for exogenous variables. Moreover, the proof of Theorem A1 differs significantly from previous work (e.g. Lebrun (1999), Lizzeri and Persico (2000)) and is based on a functional approach viewing $s(\cdot)$ as a zero of a nonlinear operator. A Functional Implicit Function Theorem and a continuation argument then allow us to establish existence, uniqueness and smoothness of $s(\cdot)$, especially with respect to exogenous variables as needed in estimation.

Except for some particular choices for $U(\cdot)$ and $F(\cdot)$, the equilibrium strategy does not have an explicit form. In practice, the integral form of the differential equation (2) can be useful. It is

$$s(v)F^{I-1}(v) = \int_{\underline{v}}^{v} \{s(x) + \lambda(x - s(x))\} \, dF^{I-1}(x).$$

This form can be also written as $s(v) = s_N(v) + \int_{\underline{v}}^{v}\{\lambda(x - s(x)) - x + s(x)\} \, dF^{I-1}(x)$, where $s_N(\cdot)$ is the well-known equilibrium strategy in the risk neutral case derived by e.g. Riley and Samuelson (1981). Because $\lambda(u) \geq u$ for $u \geq 0$ (see below), it follows that the equilibrium bid under risk aversion is strictly larger than under risk neutrality for $v > \underline{v}$ as noted by Milgrom and Weber (1982).

3 General Nonidentification Results

In this section we address the problem of identification of the structure $[U, F]$ from observables. Specifically, we assume that the number I of bidders is observed, as is typically the case 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 distribution $G(\cdot)$ for an observed bid can be rationalized by a structure $[U, F]$ given I. Such a question is important in itself as it is connected to the validity of the auction model under consideration.

Following Guerre, Perrigne and Vuong (2000) for the risk neutral case, we can express the differential equation (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 the differential equation (2) can be written equivalently 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 its 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, which are implied by the smoothness of $s(\cdot)$ stated in Theorem 1 and the regularity assumptions on $[U, F]$. It is convenient to introduce the following definition.

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 a necessary and sufficient condition for rationalizing a distribution of observed bids by an IPV auction model with risk aversion. Hereafter, we say that a distribution is *rationalized* by a risk averse auction model 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) $\mathbf{G}(b_1, \ldots, b_I) = \prod_{i=1}^{I} G(b_i)$,
(ii) $G(\cdot) \in \mathcal{G}_R$,
(iii) $\exists \lambda : I\!\!R_+ \to I\!\!R_+$ *with* $R+1$ *continuous derivatives on* $[0, +\infty)$, $\lambda(0) = 0$ *and* $\lambda'(\cdot) \geq 1$ *such that* $\xi'(\cdot) > 0$ *on* $[\underline{b}, \bar{b}]$, *where*

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

Lemma 1 provides a necessary and sufficient condition for rationalizing an observed bid distribution. The first condition is related to the use of the IPV paradigm and requires that bids be independently and identically distributed in agreement with private values. The second condition requires that the marginal observed bid distribution $G(\cdot)$ satisfies the regularity assumptions embodied in the set $\mathcal{G}_R$ of Definition 3. The third condition arises from the fact that $\xi(\cdot, U, G, I)$ is the inverse of the equilibrium strategy, which is strictly increasing. As shown in the proof of Lemma 1, if such a condition 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 $b \sim G(\cdot)$.[8]

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.

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 with $c \in [0, 1)$ as well as a CARA structure with both zero wealth and a private value distribution in $\mathcal{F}_R$.*

Proposition 1 is striking. First, it implies that the restriction (iii) in Lemma 1 for rationalizing a bid distribution by an IPV auction model with risk averse bidders is redundant unlike Condition C2 in Guerre, Perrigne and Vuong (2000, Theorem 1) for the risk neutral case. Specifically, our proof indicates that we can always find a function $\lambda(\cdot)$ corresponding to either a CRRA or CARA utility function so that (iii) is satisfied whenever $G(\cdot) \in \mathcal{G}_R$.

[8]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 by Definition 1-(i).

Alternatively, the IPV auction model with general risk aversion does not impose any restrictions on observed bids beyond their independence and the weak regularity conditions embodied in $\mathcal{G}_R$.

Second, because a general risk aversion 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 with zero wealth that are observationally equivalent to $[U, F]$. In other words, irrespective of initial wealth, 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 with zero wealth. Hence, without loss of power for explaining bids, an analyst could consider either a CRRA or a CARA model with zero wealth, provided these models are identified and can be estimated.

Because a risk neutral model is a special case of a risk averse model, it follows from Proposition 1 that any risk neutral model is observationally equivalent to a risk averse model such as a CRRA or a CARA model. An interesting question is whether the reverse holds, i.e. whether any risk averse model is observationally equivalent to a risk neutral model. This is not true.[9] Thus, by allowing for risk aversion, one does enlarge the set of rationalizable bid distributions relative to risk neutrality. As a matter of fact, Proposition 1 says that allowing for very simple forms of risk aversion such as CRRA or CARA rationalizes any bid distribution in $\mathcal{G}_R$.

A model is a set of structures $[U, F]$. For instance, the CARA model (with regularity R) is the set of structures $[U, F]$ where $U(\cdot) \in \mathcal{U}_R^{CARA} \equiv \{U(\cdot) = U_{vNM}(w + \cdot) - U_{vNM}(w); w \in$

[9]The following is an example with $I = 2$ of a CRRA structure that is not observationally equivalent to any risk neutral structure. Consider the distribution $G(b) = kb$ 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 the equation $-8x^2 + 11x - 2 = 0$ and k such that $G(\cdot)$ is continuous at $b = 1/2$. The corresponding density $g(\cdot)$ is flat on $[0, 1/2]$ and sharply increasing on $[1/2, 1]$. This distribution satisfies the regularity conditions of Definition 3 with $R = 1$. Letting $U(x) = x^{1-c}$ gives $\lambda(x) = x/(1 - c)$. The bid distribution $G(\cdot)$ can be rationalized by a CRRA structure where the inverse bidding strategy is $\xi(b, c, G) = b + (1 - c)G(b)/g(b)$ as soon as $2/5 < c < 1$ by Lemma 1-(iii). On the other hand, from Guerre, Perrigne and Vuong (2000) the distribution $G(\cdot)$ 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 $0 \le b \le 1/2$ and $\xi(b, G) = -\frac{8}{3}(b - \frac{1}{2})(b - \frac{5}{4}) + 1$ for $1/2 \le b \le 1$. It can be easily checked that this function is not strictly increasing. Hence there does not exist a risk neutral structure that is observationally equivalent to the preceding CRRA structure.

$I\!R_+, U_{vNM}(\cdot) \in \mathcal{U}^{CARA}\} \cap \mathcal{U}_R$ and $F(\cdot) \in \mathcal{F}_R$. The sets $\mathcal{U}_R^{DARA}$, $\mathcal{U}_R^{CRRA}$ and $\mathcal{U}_R^{DRRA}$ are similarly defined. As suggested by Proposition 1, auction models with risk averse bidders are nonidentified, in general. Hereafter, we say that a structure $[U, F]$ is *nonidentified* if there exists another structure $[\tilde{U}, \tilde{F}]$ within the model under consideration that leads to the same equilibrium bid distribution. If no such $[\tilde{U}, \tilde{F}]$ exists for any $[U, F]$, we say that the model is (globally) identified.

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. Similarly, any structure $[U, F]$ in $\mathcal{U}_R^{DARA} \times \mathcal{F}_R$, $\mathcal{U}_R^{DRRA} \times \mathcal{F}_R$, $\mathcal{U}_R^{CARA} \times \mathcal{F}_R$ or $\mathcal{U}_R^{CRRA} \times \mathcal{F}_R$ is not identified.*

As shown by Guerre, Perrigne and Vuong (2000), the IPV auction model with risk neutral bidders is identified from observed bids. 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. The second part of Proposition 2 indicates that restricting $U(\cdot)$ to be derived from the four well known families of utility functions is still insufficient for achieving identification even if $U(\cdot)$ is restricted to a simple parametric specification such as the CRRA model.

It is useful to understand the source of nonidentification by considering the CRRA model with zero wealth.[10] Hence, $U(x) = x^{1-c}$ with $0 \leq c < 1$ and $F(\cdot) \in \mathcal{F}_R$. Let $G(\cdot)$ be the corresponding equilibrium bid distribution. Consider the alternative CRRA structure $[\tilde{U}, \tilde{F}]$, where $\tilde{c}$ satisfies $c < \tilde{c} < 1$, while $\tilde{F}(\cdot)$ is the distribution of

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

[10]For the general risk aversion model, where $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$ with arbitrary wealth, let $[\tilde{U}, \tilde{F}]$ be such that $\tilde{U}(\cdot) = [U(\cdot/\alpha)/U(1/\alpha)]^{\alpha}$, with $\alpha \in (0, 1)$ and $\tilde{F}(\cdot)$ be the distribution of

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

with $b \sim G(\cdot)$, the equilibrium bid distribution under $[U, F]$. The second equality follows from $\tilde{\lambda}(\cdot) \equiv \tilde{U}(\cdot)/\tilde{U}'(\cdot) = U(\cdot/\alpha)/U'(\cdot/\alpha) = \lambda(\cdot/\alpha)$. It is easy to check that $[\tilde{U}, \tilde{F}] \in \mathcal{U}_R \times \mathcal{F}_R$. Note that $\tilde{\xi}(\cdot)$ can be decomposed as the sum of $(1-\alpha)b$ and $\alpha\xi(b) = \alpha[b + \lambda^{-1}(G(b)/(I-1)g(b))]$, which are two strictly increasing functions in b. Hence, $\tilde{\xi}(\cdot)$ is strictly increasing. Thus, from Lemma 1 the structures $[U, F]$ and $[\tilde{U}, \tilde{F}]$ are observationally equivalent, i.e. lead to the same bid distribution $G(\cdot)$.

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 state that the CRRA model is identified. In fact, because of their assumption 1, these authors restrict the distribution $\tilde{F}(\cdot)$ to have the same support as $F(\cdot)$. In contrast, our result shows that the CRRA model is not identified as $F(\cdot)$ and $\tilde{F}(\cdot)$ may have different supports. In particular, at $b = \bar{b}$ the above equation indicates that the support of $\tilde{F}(\cdot)$ must shrink, i.e. $\tilde{\bar{v}} < \bar{v}$, 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 $F(\cdot)$ by the same argument.

To summarize, considering risk aversion even under its simplest form such as CRRA much increases the explanatory power of the model relative to the risk neutral case since all bid distributions in $\mathcal{G}_R$ can now be rationalized. On the other hand, the validity of such a model can no longer be tested as the theory does not provide restrictions beyond the independence of bids and the regularity conditions of $\mathcal{G}_R$. Moreover, risk averse models as simple as CRRA and CARA models are nonidentified from observed bids. In particular, parameterizing the utility function is not sufficient for achieving identification. Additional identifying restrictions are thus needed.

4 Semiparametric Identification

The purpose of this section is to exploit heterogeneity across auctioned objects combined with palatable identifying restrictions to achieve semiparametric identification of first-price auction models with risk averse bidders.[11] Heterogeneity across objects is characterized by a vector of observed variables Z, which can be discrete or continuous with values z in $\mathcal{Z} \subset I\!R^d$. For instance, Z can include a dummy variable for the quality of the auctioned object or a continuous variable indicating the object's appraisal value. As before, we assume that the number of bidders $I \in \mathcal{I}$ is observed, which can be either constant or varying across auctions. Hereafter, we thus consider that private values are

[11]For instance, if $\bar{v}$ was known, Donald and Paarsch (1996) result would apply and the CRRA model would be identified. Assuming that $\bar{v}$ is known is, however, very strong as $\bar{v}$ directly affects the risk aversion parameter c because $\bar{v} = \bar{b} + (1-c)/[(I-1)g(\bar{b})]$.

drawn from the conditional distribution $F(\cdot|Z,I)$.[12] Our preceding nonidentification results of risk averse models then hold when the whole structure depends on (Z,I) namely $[U,F]=\{[U(\cdot|z,I),F(\cdot|z,I)],z\in\mathcal{Z},I\in\mathcal{I}\}$.

A first natural restriction is to require that the utility function $U(\cdot)$ be independent of (Z,I). Hence risk aversion is independent of the characteristics of the auctioned objects and the number of bidders. This is justified in the case studied here as bidders do not face uncertainty about the quality and equivalent monetary value of the auctioned object. Restricting $U(\cdot)$ to be independent of (Z,I) is, however, insufficient for identifying the model as noted later. Thus we need to consider additional restrictions on both $U(\cdot)$ and $F(\cdot|\cdot,\cdot)$ to achieve identification. We impose the following ones.

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) $F(\cdot|\cdot,\cdot)\in\mathcal{F}_R(\mathcal{Z}\times\mathcal{I})\equiv\{F(\cdot|\cdot,\cdot):F(\cdot|z,I)\in\mathcal{F}_R,\forall(z,I)\in\mathcal{Z}\times\mathcal{I}\}$. The support of $F(\cdot|z,I)$ is denoted $[\underline{v}(z,I),\overline{v}(z,I)]$,
(iii) For some $\alpha\in(0,1]$, the α-quantile $v_\alpha(z,I)$ of $F(\cdot|z,I)$ satisfies $v_\alpha(z,I)=v_\alpha(z,I;\gamma)$ for all $(z,I)\in\mathcal{Z}\times\mathcal{I}$ and some $\gamma\in\Gamma\subset I\!R^q$,
(iv) The function $\psi_\alpha(z,I;\theta,\gamma)\equiv\lambda(v_\alpha(z,I;\gamma)-b_\alpha(z,I);\theta)$ for $(z,I)\in\mathcal{Z}\times\mathcal{I}$ determines uniquely $(\theta,\gamma)\in\Theta\times\Gamma$, where $b_\alpha(z,I)$ is the α-quantile of the equilibrium bid distribution $G(\cdot|z,I)$ generated by the structure $[U,F]$.

Condition (i) requires that $U(\cdot)$ belongs to a parametric family of utility functions that are smooth. Utility functions derived from CRRA and CARA vNM utility functions satisfy such a condition with $R=\infty$. It is also satisfied by many parametric families that allow for flexible patterns of risk aversion. Note that if initial wealth w is unknown, then w must be included in the parameter vector θ. Condition (ii) requires that the conditional distribution $F(\cdot|z,I)$ satisfies the regularity conditions of Definition 2 for every $(z,I)\in\mathcal{Z}\times\mathcal{I}$.

[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 (2002). Note that unobserved heterogeneity across bidders is allowed through differences in bidders' private values. On the other hand, observed heterogeneity across bidders is ruled out as it leads to an asymmetric auction model. See Campo, Perrigne, and Vuong (2002) where bidders are *ex ante* different under risk neutrality.

Condition (iii) is a parametric conditional quantile restriction on $F(\cdot|z,I)$, as frequently used in the semiparametric literature. See Powell (1994). For instance, $v_\alpha(\cdot,\cdot;\gamma)$ can be chosen to be a constant or more generally a polynomial, where γ is the vector of unknown coefficients. Note that $\alpha = 1$ is allowed, in which case a parametric specification of the upper bound $\overline{v}(z,I)$ is considered. On the other hand, no assumption is made on the lower bound $\underline{v}(z,I)$ corresponding to $\alpha = 0$ as $\underline{v}(z,I)$ is nonparametrically identified from the boundary condition $\underline{v}(z,I) = \underline{b}(z,I)$. An alternative identifying assumption to (iii) would be to require that the difference $v_\alpha(z,I) - \underline{v}(z,I)$ is a parametric function of (z,I). This is equivalent to imposing a restriction on the α-quantile as the lower boundary $\underline{v}(z,I)$ can be recovered from $\underline{b}(z,I)$. In particular when $\alpha = 1$, $v_\alpha(z,I) = \overline{v}(z,I)$, in which case this alternative identifying assumption would correspond 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 a standard identifying condition of the parameter vector (θ,γ) from the knowledge of the function $\psi_\alpha(\cdot,\cdot;\theta,\gamma)$ on $\mathcal{Z}\times\mathcal{I}$. It implies the "order" condition Card $\mathcal{Z}\times\mathcal{I} \geq p+q$. Condition (iv) bears on $[U,G]$, where G implicitly depends on the structure $[U,F]$. In particular, it can be easily verified. For instance, consider a CRRA model with zero wealth and a constant (unknown) α-quantile of $F(\cdot|\cdot,\cdot)$, i.e. $v_\alpha(z,I) = \gamma$, in which case $p = 1$ and $q = 1$. Condition (iv) is then satisfied as soon as there are two α-quantiles $b_\alpha(z_1,I_1)$ and $b_\alpha(z_2,I_2)$ that differ as shown in Proposition 3.

The next proposition establishes the semiparametric identification of the first-price auction model with risk averse bidders. It relies upon the key equation (4) giving the inverse of the equilibrium strategy, taking into account the conditioning variables (Z,I), the parameterization of the utility function $U(\cdot;\theta)$ and the α-quantile $v_\alpha(z,I)$ of $F(\cdot|z,I)$. Specifically, because the equilibrium strategy $s(\cdot,U,F,I)$ is strictly increasing, then $b_\alpha(z,I) = s(v_\alpha(z,I),U,F,I)$. Hence, (4) evaluated at the α-quantile $b_\alpha(z,I)$ gives

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

for any $(z,I)\in\mathcal{Z}\times\mathcal{I}$, where $\lambda(\cdot;\theta) = U(\cdot;\theta)/U'(\cdot;\theta)$. This equation combined with (iv) suggests how the parameter vector (θ,γ) can be identified given the knowledge of $g(b_\alpha(z,I)|z,I)$, and the specified parametric forms for $\lambda(\cdot;\theta)$ and $v_\alpha(\cdot;\gamma)$.

Proposition 3: *The semiparametric model defined as the set of structures $[U,F]$ satisfying Assumption A1 is identified. In particular, if there exists $(z_j,I_j)\in\mathcal{Z}\times\mathcal{I}, j = 1,2,$*

such that $b_\alpha(z_1, I_1) \neq b_\alpha(z_2, I_2)$, then the CRRA model and the CARA model with zero wealth, a constant conditional quantile restriction $v_\alpha = \gamma$ and $F(\cdot|\cdot,\cdot) \in \mathcal{F}_R(\mathcal{Z} \times \mathcal{I})$ are semiparametrically 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 its parametric conditional quantile restriction. Moreover, the proof shows that the parameter γ is identified. Note that in the CRRA and CARA models with zero wealth and a constant quantile restriction, the additional requirement that the α-quantile $b_\alpha(z, I)$ varies with (z, I) is readily verifiable. For instance, suppose that I does not vary, while Z is reduced to one dichotomous variable indicating e.g. the quality of the auctioned object. The CRRA model is then identified if the α-quantiles of the conditional bid distributions corresponding to the two values of Z differ.

It is worthnoting that parameterizing the utility function and the α-quantile of the distribution of private values arises naturally. In particular, dropping either one of these parameterizations would lead to a nonidentified model as the following examples indicate. For instance, assume that the conditional quantile $v_\alpha(z, I)$ is left unspecified, but a parametric specification for the utility function is retained. Specifically, consider the semiparametric model composed of the structures $[U, F]$ satisfying A1-(i,ii). Such a model would not be necessarily identified. An example is the CRRA model with $U(x) = x^{1-c}$ for $c \in [0, 1)$ and $F(\cdot|\cdot,\cdot)$ belonging to $\mathcal{F}_R(\mathcal{Z} \times \mathcal{I})$. 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 by itself identification of the semiparametric model, despite that $U(\cdot)$ does not vary with (Z, I).

Likewise, suppose that the restriction to a parametric specification of the utility function is relaxed while the parametric conditional quantile restriction is retained. That is, consider 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

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

where $0 < \delta < 1$, $c_1 = c_2[U(\delta)]^{1-\delta}$, and $c_2 = 1/U(1 + \delta(1-\delta))$.[13] Let $\tilde{F}(\cdot|z, I)$ be the distribution of

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

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).[14] Hence, the parameterization of the conditional quantile of $F(\cdot|z, I)$ is not sufficient by itself for identification.

5 Optimal Convergence Rate

The previous section has shown that the auction model with risk averse bidders is semiparametrically identified through a parameterization of the utility function and a parametric quantile restriction on the distribution of private values. This naturally leads to the search for estimators of $[U, F]$, and in particular for semiparametric estimators of θ as θ parameterizes $U(\cdot)$. ¿From the semiparametric literature, it is known that many semiparametric estimators can attain the parametric rate of convergence, while others converge at a slower rate. For the former, see Robinson (1988) and Newey and McFadden (1994) and Powell (1994) for surveys. For the latter, see Manski (1985), Horowitz (1992),

[13] Note that $\tilde{U}(0) = 0$, and $\tilde{U}(\cdot)$ has $R + 2$ continuous derivatives on $(0, \delta) \cup (\delta, +\infty)$. Thus $\tilde{U}(\cdot)$ would belong to U_R if $\tilde{U}(\cdot)$ has $R + 2$ continuous derivatives at $x = \delta^2$. In fact, $\tilde{U}(\cdot)$ has only one continuous derivative at $x = \delta^2$. Hence $\tilde{U}(\cdot)$ should be smoothed out in the neighborhood of $x = \delta^2$ to be $R + 2$ continuously differentiable on $(0, +\infty)$. We omit this smoothing requirement and use $\tilde{U}(\cdot)$ directly.

[14] From Lemma 1, we need to show that $\tilde{\xi}'(\cdot|z, I) > 0$ for any $(z, I) \in \mathcal{Z} \times \mathcal{I}$. We have

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

Because $\xi'(\cdot|z, I)$ is strictly positive, $\tilde{\xi}'(\cdot|z, I)$ is strictly positive as required. Hence $[\tilde{U}, \tilde{F}]$ rationalizes the bid distribution $G(\cdot|\cdot, \cdot)$. It remains to show that $\tilde{F}(\cdot|\cdot, \cdot)$ satisfies A1-(iii). The α-quantile $\tilde{v}_\alpha(z, I)$ of $\tilde{F}(\cdot|z, I)$ satisfies $\tilde{v}_\alpha(z, I) = \tilde{\xi}(b_\alpha(z, I)|z, I)$. Consider $G(b_\alpha(z, I)|z, I)/[(I-1)g(b_\alpha(z, I)|z, I)] = \alpha/[(I-1)g(b_\alpha(z, I)|z, I)]$. Because $\lambda(\cdot)$ is strictly increasing with $\lambda(0) = 0$, there exists δ sufficiently small so that $0 < \lambda(\delta) < \alpha/\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}[(I-1)g(b_\alpha(z, I)|z, I)]$, where the latter is assumed to be finite. Thus $\tilde{v}_\alpha(z, I) = \xi(b_\alpha(z, I)|z, I) - \delta(1-\delta) = v_\alpha(z, I) - \delta(1-\delta) > 0$ for δ sufficiently small. Hence $\tilde{v}_\alpha(z, I)$ satisfies A1-(iii) whenever $v_\alpha(z, I; \gamma)$ contains a constant term.

Kyriazidou (1997) and Honoré and Kyriazidou (2000). 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 θ. This is the primary purpose of this section. The optimal convergence rate 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 their counterparts in Definition 2, condition (i) needs further discussion. Because of the singularity of the differential equation (2) at the lower boundary of the support of the private value distribution, assuming a constant lower boundary $\underline{v}(z, I) = \underline{v}$ simplifies the proof of Theorem A1 establishing the smoothness of the equilibrium strategy $s(\cdot)$ with respect to (v, z), which is needed to obtain the smoothness of the equilibrium bid distribution. On the other hand, such a restriction is not used in estimation as $\underline{v}(z, I)$ can be recovered from $\underline{b}(z, I)$. Regarding the upper boundary restriction, Section 4 indicates that a parameterization of a quantile of $F(\cdot|\cdot,\cdot)$ is necessary for achieving identification. The upper boundary is a particular quantile corresponding to $\alpha = 1$. Our estimation procedure will rely on (5), which requires an estimate for $b_\alpha(z, I)$. There is then an important difference between estimating a quantile corresponding to $\alpha \in (0, 1)$ and estimating the upper boundary. In particular, the convergence rate for estimating the latter is much 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 $(z, I) \in \mathcal{Z} \times \mathcal{I}$.[15]

It will be useful to derive the smoothness properties of the equilibrium bid distribution $G(\cdot|\cdot,\cdot)$ corresponding to a structure $[U, F]$ in $\mathcal{U}_R \times \mathcal{F}_R^*$. Such properties are important as they relate to the implied statistical model for the observables, which are the bids, the number of bidders and the exogenous variables. They follow from Theorem A1, which establish the existence, uniqueness and smoothness of the equilibrium strategy with respect to (v, z) in $[\underline{v}, \overline{v}] \times \mathcal{Z}$ for every $i \in \mathcal{I}$.[16]

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

Lemma 2 extends Lemma 1-(ii) to the case with exogenous variables (Z, I). It parallels Proposition 1 in Guerre, Perrigne and Vuong (2000) for the risk neutral case.

In view of the above, we then consider the semiparametric model composed of structures $[U, F]$ satisfying the following assumption.

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 $\psi_1(z, I; \theta, \overline{v}) \equiv \lambda(\overline{v} - \overline{b}(z, I); \theta)$ for $(z, I) \in \mathcal{Z} \times \mathcal{I}$ determines uniquely

[15] When there are no exogenous variables Z, assuming that $F(\cdot|I)$s have a common support agrees with the theoretical assumption that the private value distribution is independent of the number of bidders.

[16] To simplify the presentation, we exclude discrete exogenous variables by requiring $\mathcal{Z}$ to have a nonempty interior. Our next results continue to hold with suitable modifications.

$(\theta, \overline{v}) \in \Theta \times (0, +\infty)$, *where* $\overline{b}(z, I)$ *is the upper boundary of the equilibrium bid distribution* $G(\cdot|z, I)$ *generated by the structure* $[U, F]$ *with* I *bidders.*

Conditions (i) and (ii) strengthen A1-(i,ii,iii) of A1. Condition (iii) simply expresses A1-(iv) at the upper boundary under a constant restriction. Then (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 the heterogeneity of the auctioned objects. This gives a total number $N = \sum_{\ell=1}^{L} I_\ell$ of bids, where L is the number of auctions. Thus $f(\cdot|Z_\ell, I_\ell)$ is the density of private values conditional upon (Z_ℓ, I_ℓ) in auction ℓ. Following the game theoretical model of Section 2, we make the following assumption on the data generating process and the specification of the semiparametric model composed of structures $[U, F]$ satisfying A2.

Assumption A3:
(i) The variables (Z_ℓ, I_ℓ), $\ell = 1, 2 \ldots$ *are independently and identically distributed with support* $\mathcal{Z} \times \mathcal{I}$, *where* $\mathcal{I}$ *is finite and* $0 < \inf_{(z,I) \in \mathcal{Z} \times \mathcal{I}} g(z, I) \leq \sup_{(z,I) \in \mathcal{Z} \times \mathcal{I}} g(z, I) < +\infty$,
(ii) For every ℓ, *the private values* $V_{i\ell}, i = 1, \ldots, I_\ell$ *are independently and identically distributed 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 Assumption A2 for some* $\theta_0 \in \Theta$ *and* $0 \leq \underline{v}_0 < \overline{v}_0 < +\infty$.

In particular, private values and hence bids are independent across auctions.[17]

We are now in a position to establish the optimal rate at which $\beta = (\theta, \overline{v})$ can be estimated. As in Horowitz (1993), we invoke the minimax theory developed by e.g. Ibragimov and Has'minskii (1981). We consider the following norms

$$||\beta||_\infty = \max(\max_{1 \leq k \leq p} |\theta_k|, |\overline{v}|), \quad ||f(\cdot|\cdot, \cdot)||_\infty = \sup_{(z,I) \in \mathcal{Z} \times \mathcal{I}} \sup_{v \in [\underline{v}, \overline{v}]} |f(v|z, I)|$$

[17]Not all of A3 is used to prove Theorem 2. In particular, (Z_ℓ, I_ℓ) need not be independently and identically distributed. Furthermore, A3-(i) can be weakened allowing Z_ℓs not to be independently and identically distributed as Theorem 3 is derived conditionally upon $(Z_1, I_1, \ldots, Z_\ell, I_\ell)$.

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.

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

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 $(\theta, f, f_{Z,I})$, where $f_{Z,I}$ is the joint density of the (Z_ℓ, I_ℓ)s. The next theorem establishes an upper bound for the optimal rate when estimating β_0. It crucially relies on Lemmas 1 and 2 and Proposition 3. Let Θ^o denote the interior of Θ.

Theorem 2: *Under Assumptions 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*

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

for any $t \geq 0$, 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 $N^{(R+1)/(2R+3)}$, and hence at the parametric rate $\rho_N = \sqrt{N}$. On the other hand, Theorem 3 in the next section will show that there exists an estimator $\hat{\beta}_N$ that converges 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 introduce some 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 $\psi : I\!R^- \to I\!R$ 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 (i) the parameter θ characterizing the bidders' utility function $U(\cdot; \theta)$ and hence bidders' risk aversion, and (ii) the conditional latent density $f(\cdot|\cdot, \cdot)$ of bidders' private values. Because we do not restrict $f(\cdot|\cdot, \cdot)$ to belong to a parametric family, the estimation problem is semiparametric. A first subsection presents our semiparametric procedure and its different steps, while a second subsection establishes the asymptotic properties of our estimator of θ.

6.1 A Semiparametric Procedure

Our semiparametric procedure follows closely the semiparametric identification result. By (6) and (7), it relies on the identifying relation

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

where the subscript 0 indicates quantities at the truth. If one knew the upper boundary $\overline{b}_0(\cdot, \cdot)$ and the density $g_0(\cdot|\cdot, \cdot)$, one could recover the utility function parameters θ_0 from (8) given the chosen parametric form for $\lambda(\cdot; \cdot)$. From the knowledge of $G_0(\cdot|\cdot, \cdot)$ and θ_0, one could then recover every bidder's private value v_i from (4) to estimate $f_0(\cdot|\cdot, \cdot)$. Unfortunately, $\overline{b}_0(\cdot, \cdot)$, $G_0(\cdot|\cdot, \cdot)$ and $g_0(\cdot|\cdot, \cdot)$ are unknown, but they can be estimated from observed bids. 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 obtained in the first step, estimate $\beta_0 \equiv (\theta_0, \overline{v}_0)$ using NLLS,
- 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 $f_0(\cdot|\cdot,\cdot)$.

The next subsections detail each of these steps.

Nonparametric Boundary Estimation

This step consists in estimating the upper boundary $\overline{b}_0(\cdot,\cdot)$ of the bid distribution 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), the upper boundary $\overline{b}_0(\cdot, I)$ is $R+1$ continuously differentiable on $\mathcal{Z}$. Following Korostelev and Tsybakov (1993), one introduces a partition of $\mathcal{Z}$ into bins increasing with the number of observations. 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

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

This estimator converges at the uniform rate $(N/\log N)^{2/3}$, which is strictly faster than $\sqrt{N}$ and sufficient for our purpose.

Turning to the estimation of the density $g(\cdot|\cdot,\cdot)$ at the upper boundary, 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. Specifically, let

$\Phi(\cdot)$ be a one-sided kernel with support $[-1, 0]$ satisfying some conditions, which include $\int \Phi(x)dx = 1$ and $\int x^r \Phi(x)dx = 0$ for $r = 1, \ldots, R$ (see Assumption 4-(iii)). For every $\ell = 1, \ldots, L$ and $i = 1, \ldots, I_\ell$, 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 some bandwidth. Lemma D1 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] In practice, one does not know $\overline{b}_0(\cdot, \cdot)$. We thus define $\hat{Y}_{i\ell}$ similarly to $Y_{i\ell}$, where $\overline{b}_0(\cdot, \cdot)$ is replaced by its estimator $\hat{\overline{b}}_N(\cdot, \cdot)$ obtained previously.

Semiparametric Estimation of θ_0

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

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

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 D1 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 a variance of the error term diverging to infinity as h_N vanishes. These features raise some technical difficulties when deriving the asymptotic properties of our estimator of θ. In particular, the diverging variance of the error term is the main reason why our estimator does not achieve the parametric rate $\sqrt{N}$ of convergence. Specifically, its rate $N^{(R+1)/(2R+3)}$ is smaller than $\sqrt{N}$ but is optimal in the minimax sense in view of Section 5.

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

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

[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.

with respect to $\beta = (\theta, \overline{v}) \in \mathcal{B}_\delta$, where the $\omega(Z_\ell, I_\ell)$s are strictly positive weights. The set $\mathcal{B}_\delta$ is defined as $\mathcal{B}_\delta = \{(\theta, \overline{v}); \theta \in \Theta, \sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}} \overline{b}_0(z, I) + \delta \leq \overline{v} \leq \overline{v}_{\sup}\}$ 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 (see Lemma C1). Because $\overline{b}_0(\cdot, \cdot)$ and hence $m(\cdot, \cdot; \beta)$ are unknown, the preceding estimator is infeasible. We thus replace $\overline{b}_0(\cdot, \cdot)$ by its estimator obtained in Step 1. Thus, our estimator of β is defined as $\hat{\beta}_N = \mathrm{Argmin}_{\beta\in\mathcal{B}_N} \hat{Q}_N(\beta)$, where

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

and

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

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 now estimated using the estimator $\hat{\theta}_N$ of θ_0 in Step 2, while $\lambda(\cdot)$ was known and equal to the identity in that paper.

Specifically, 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

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

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

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

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 $\hat{\overline{b}}_N(\cdot, I)$ is the previous boundary estimator and $\hat{\underline{b}}_N(\cdot, I)$ is defined similarly.

The N pseudo private values $\hat{V}_{i\ell}$ hence obtained are used in a standard kernel estimation of $f_0(\cdot|\cdot)$.[19] Namely, for an arbitrary pair (v,z), $f(v|z)$ is estimated by

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

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 our semiparametric estimator $\hat{\theta}_N$ converges at a faster rate, as shown in the next section, 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 $(N/\log N)^{R/(2R+d+3)}$. Moreover, the latter is the optimal rate for estimating $f_0(\cdot|\cdot)$ from observed bids.

6.2 Asymptotic Properties

In this section, we derive the asymptotic properties of our estimator of θ_0. In particular, we show that our semiparametric estimator $\hat{\theta}_N$ converges at the rate $N^{(R+1)/(2R+3)}$, which is independent of the dimension of Z. In view of Theorem 2, it follows that the optimal rate for estimating θ_0 is $N^{(R+1)/(2R+3)}$ and that $\hat{\theta}_N$ converges at this optimal rate.

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

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

(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 $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}} \omega(z,I) < \infty$,

[19] As in Guerre, Perrigne and Vuong (2000), we focus on the conditional density $f_0(v|z)$, though similar results are obtained for $f_0(v|z,I)$.

(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 Assumption 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 Assumption A4-(i) is standard in parametric estimation. In particular, it strengthens the identification requirement of β_0 from the parametric specification $m(\cdot, \cdot; \beta)$. For instance, it implies that $\overline{b}_0(z, I)$ must have at least $p + 1$ different values. As shown in Lemma D5, combined with Assumption A4-(ii), it also ensures that the usual matrices

$$
\begin{aligned}
A(\beta) &\equiv \frac{1}{\mathrm{E[I]}}\mathrm{E}\left[I\omega(z, I)\frac{\partial m(z, I; \beta)}{\partial \beta} \cdot \frac{\partial m(z, I; \beta)}{\partial \beta'}\right] && (14)\\
B(\beta) &\equiv \frac{1}{\mathrm{E[I]}}\mathrm{E}\left[I\omega^2(z, I)m(z, I; \beta)\frac{\partial m(z, I; \beta)}{\partial \beta} \cdot \frac{\partial m(z, I; \beta)}{\partial \beta'}\right] && (15)
\end{aligned}
$$

are of full rank in a neighborhood of β_0.

Assumptions A4-(iii,iv) are standard in kernel estimation when using higher order kernels though our kernel is one-sided. Assumption A4-(v) requires that the boundary estimator $\hat{\overline{b}}_N(\cdot, \cdot)$ converges faster than the semiparametric estimator $\hat{\theta}_N$ (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) (see previous subsection). 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 Assumption 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 the boundary estimator $\hat{\overline{b}}_N(\cdot, \cdot)$ to satisfy Assumption A4-(v).

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

$$
\begin{aligned}
A_N(\beta) &= \sum_{\ell=1}^{L} I_\ell \omega(Z_\ell, I_\ell)\frac{\partial m(Z_\ell, I_\ell; \beta)}{\partial \beta} \cdot \frac{\partial m(Z_\ell, I_\ell; \beta)}{\partial \beta'}, && (16)\\
B_N(\beta) &= \sum_{\ell=1}^{L} I_\ell \omega^2(Z_\ell, I_\ell)m(Z_\ell, I_\ell; \beta)\frac{\partial m(Z_\ell, I_\ell; \beta)}{\partial \beta} \cdot \frac{\partial m(Z_\ell, I_\ell; \beta)}{\partial \beta'}, && (17)
\end{aligned}
$$

which, when normalized by N, are consistent estimators of $A(\beta)$ and $B(\beta)$ as shown in Lemma D5. 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 $m(\cdot;\beta)$ replaced by $\hat{m}(\cdot;\beta)$. Moreover, let

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

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 Assumptions A2–A4,*
(i) $\hat{\beta}_N$ is a consistent estimator of β_0 with

$$\hat{\beta}_N - \beta_0 = O_P\left(h_N^{R+1} + \frac{1}{\sqrt{Nh_N}}\right) ,$$

so the best rate of convergence of $\hat{\beta}_N$ is $N^{(R+1)/(2R+3)}$, which is achieved when the exact order of the bandwidth h_N is $N^{-1/(2R+3)}$.
(ii) If $\lim_{N\to\infty}\sqrt{Nh_N}h_N^{R+1} = \infty$, then

$$\frac{1}{h_N^{R+1}}\left(\hat{\beta}_N - \beta_0\right) \xrightarrow{P} A(\beta_0)^{-1}\mathbf{b}(\beta_0, g_0) .$$

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

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

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

On technical grounds, the proof of Theorem 3-(i) is complicated by the divergence of the variance (11) of the error term $\epsilon_{i\ell}$ in the nonlinear model (10). In particular, omitting the estimation of the upper boundary $\overline{b}(\cdot,\cdot)$, which has no effect by Assumption 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 the normalization $h_N Q_N(\beta)/N$, but its limit is a constant independent of β. To overcome such a difficulty, we show that $(Q_N(\beta) - Q_N(\beta_0) - \overline{Q}_N(\beta))/N$ vanishes asymptotically, where

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

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

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

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

This expansion corresponds to the usual bias/variance decomposition of 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

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

When $Nh_N^{2R+3} \to \infty$, 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 $\hat{\beta}_N - \beta_0 = O_P(N^{-(R+1)/(2R+3)})$.[20]

The best convergence rate $N^{(R+1)/(2R+3)}$ of $\hat{\beta}_N$ is independent of the dimension d of Z and corresponds to the optimal rate for estimating an univariate density with $R+1$ bounded derivatives. This seems surprising in view of the key relation (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. Note that the previous moment conditions are similar to those considered by Newey and McFadden (1994) though Assumptions (iii)-(iv) of their Theorem 8.1 is not satisfied in our case. In fact, because the variance (11) is diverging, our proof shows that the average gradient $(1/N)\partial\hat{Q}_N(\beta_0)/\partial\beta = O_{\mathrm{P}}(h^{R+1} +$

[20] 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.

$1/\sqrt{Nh_N})$, which is different from the usual $O_{\mathrm{P}}(1/\sqrt{N})$. Hence, our estimator converges at a slower rate than $\sqrt{N}$.

In practice, Theorem 3-(iii) is used to make inference on β_0 as it gives an estimate of the variance of $\hat{\beta}_N$, namely $(\int \Phi^2(x)dx/h_N)\hat{A}_N^{-1}(\hat{\beta}_N)\hat{B}_N(\hat{\beta}_N)\hat{A}_N^{-1}(\hat{\beta}_N)$. Note that $\hat{\beta}_N$ depends on the weights $\omega(\cdot,\cdot)$, which can be chosen optimally to decrease the asymptotic variance of $\hat{\beta}_N$ 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 the optimal weights are 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 $e_{i\ell} = 0$. An interesting question is then the existence of an estimator based upon the $Y_{i\ell}$s with a smaller asymptotic variance in a local minimax sense.

7 Empirical Application

This section illustrates the previous methodology on timber auction sales from the US Forest Service. A first subsection briefly presents the data. A second subsection discusses the implementation of our estimation method for a CRRA utility specification and gives the estimation results. In particular, risk aversion is found to be significant.

7.1 Data

The US Forest Service (USFS) timber auction data have been widely used in empirical studies on auctions. Comparing revenues generated from ascending and sealed-bid auctions, Hansen (1985) tests the revenue equivalence theorem. Adopting an independent private value framework, Baldwin, Marshall and Richard (1997) study collusion, while Baldwin (1995) attempts to test for the presence of risk aversion. More recently, Athey and Levin (2001) study the practice of skewed bidding when bidders bid on species and when payments are based on actual harvested value. Their analysis suggests that bidders are risk averse. Haile (2001) analyzes the bidding behavior when there are resale opportunities after the auctions. Each of these papers focuses on a particular 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. For this reason, many characteristics of these auctions such as collusion, skewed bidding and resale markets are left aside to focus on the issue of risk aversion.

The Western half of the United States has a large part of its forestry publicly owned and is an important provider of timber in the country.[21] The Forest Service uses both oral ascending and first-price sealed-bid auctions for selling its standing timber. We focus here on the first-price sealed-bid auctions for the year 1979. There is a total of 378 auctions involving a total of 1,400 sealed bids from sawmills.

The data contain a set of variables characterizing each timber lot on sale varying from the various species included in the lot, the estimated volume measured in mbf, the logging cost in dollars, the acreage of the lot, the term of the contract measured in months, the month during which the auction was held, the location of the lot, the total reserve price in dollars and the appraisal value in dollars. The latter is an estimated value of the lot provided by the USFS taking into account the quality and quantity of timber. In addition to these variables, the data provide the number of bidders who have submitted a sealed bid as well as their bid in dollars and their identities. Table 1 gives some summary statistics on the bids per mbf, the winning bid per mbf, the reserve price per mbf, the appraisal value per mbf, the volume in mbf, the density computed as the ratio of the volume per acre, the acreage and the number of bidders.

The auctioned lots display important heterogeneity in size and quality. When regressing the logarithm of bids per mbf on a complete set of variables characterizing the auctioned lot including region dummies and seasonality effects, only two variables are strongly significant, namely the number of bidders and the appraisal value. As expected, a larger number of bidders increases competition and therefore the bids, while bids are increasing in the lot value.[22] Thus, the appraisal value seems to be the best candidate to capture the heteregoneity across auctioned objects. Such a feature has been already

[21] The data analyzed here come from Regions 1 to 6, covering the states of Idaho, Montana, North and South Dakota, Nebraska, Kansas, Colorado, Wisconsin, Arizona, New Mexico, Nevada, Utah, California, Oregon and Washington.

[22] A quadratic term has been included as well to capture a decrease after some value for the number of bidders as predicted by the common value model. It is not significant.

observed in previous empirical studies, see e.g. Haile (2001).

Table 1: Some Summary Statistics

Variable	Mean	STD	Min	Max
Bids	97.28	71.51	1.05	665.18
Winning Bid	117.03	88.67	4.37	665.18
Reserve Price	62.95	46.01	1.00	217.36
Appraisal Value	57.07	45.41	1.00	199.58
Volume	1,621.93	3,153,48	11.00	23,500.00
Density	2.05	5.17	0.002	46.43
Acreage	1,348.35	3,590.69	1.00	38,850
Number of Bidders	3.70	1.81	2	12

The auctions are organized with a posted reserve price. It is well known that this reserve price does not act as a screening device to participating to this auction, see e.g. Haile (2001). To assess such a statement, we have estimated the probability of a bid being close to the reserve price. Using a nonparametric estimator, we have estimated the conditional probability $\Pr(p_0 \leq b \leq (1+\delta)p_0|Z)$, where p_0 denotes the reserve price, b the bid variable, δ an arbitrary value larger than 0 and Z the appraisal value. At the average value $Z = 57.07$, we find this probability to be equal to 1.4% for $\delta = 0.05$, 4.5% for $\delta = 0.10$ and 9.5% for $\delta = 0.20$. These results indicate that only a few bids are in the neighborhood of the reserve price and that the possible screening effect of the reserve price is negligible. It is also interesting to note that no strong relationship was found between the number of bidders and the value of each lot using various regression models including Poisson regression models. As a matter of fact, the number of bidders is very slightly decreasing in the appraisal value.

7.2 Estimation Results

The first step consists in estimating nonparametrically the upper boundary $\bar{b}_0(Z_\ell, I_\ell)$ and the bid density at this upper boundary $g_0(\bar{b}_0(Z_\ell, I_\ell)|Z_\ell, I_\ell)$ for $\ell = 1, \ldots, L$. This step needs to be conducted for each value $I = 2, \ldots, 12$ separately to take into account the dependence of the observed bids on the number of bidders. In practice, the data provide enough auctions for two and three bidders. Above four bidders, the number of observed auctions is too small for implementing a nonparametric estimator. As the bids increase in the number of bidders, it is expected that the upper boundary also increases in the number of bidders at z given. The boundary estimator of Section 6.1 was applied separately for the 107 auctions with two bidders and the 109 auctions with three bidders. No significant increase was found. In view of this, we have pooled the data and estimated a unique upper boundary for the 378 auctions.

Specifically, we let $R = 1$ hereafter. First, we choose a partition of the interval $[1, 199.58]$ for Z into 20 equal bins of approximate length 9.93.[23] For each bin, we estimate the coefficients (a_k, b_k) of the optimal straight line of Section 6.1, which then provides the estimated boundary $\hat{\overline{b}}_N(\cdot)$. Next, we need to specify a one-sided kernel $\Phi(\cdot)$ defined on $[-1, 0]$ satisfying $\int_{-1}^{0} \Phi(x)dx = 1$ and being of order one, i.e. $\int_{-1}^{0} x\Phi(x)dx = 0$. The linear kernel $\Phi(x) = (6x + 4)\mathbb{1}(-1 \leq x \leq 0)$ satisfies such requirements. The density at the upper boundary $g_0(\overline{b}_0(Z_\ell, I_\ell)|Z_\ell, I_\ell)$ is estimated by $\hat{Y}_{i\ell}$ as given by (9), where h_N is proportional to $(1,400)^{-1/5}$ following the optimal rate of Theorem 3.

The second step consists in estimating the parameter of risk aversion θ. Following previous experimental studies on auctions, we choose a CRRA specification, namely $U(x) = x^\theta$, where $\theta = 1 - c$. In particular, this choice allows us to test for risk neutrality corresponding to $c = 0$. In this case, $m(z, I; \beta_0)$ takes a simple form with (8) reducing to

$$Y_{i\ell} = \frac{1}{I_\ell - 1} \frac{\theta_0}{\overline{v}_0 - \overline{b}_0(Z_\ell, I_\ell)} + e_{i\ell} + \epsilon_{i\ell}, \tag{20}$$

where $e_{i\ell}$ is a vanishing bias and $\epsilon_{i\ell}$ is an error term. The optimal weighted NLLS estimator of $(\theta_0, \overline{v}_0)$ is obtained as $(\hat{\theta}_N^*, \hat{\overline{v}}_N^*) = \text{Argmin}_{(\theta,\overline{v})\in\mathcal{B}_N} \hat{Q}_N(\theta, \overline{v})$, where

$$\hat{Q}_N(\theta, \overline{v}) = \sum_{\ell=1}^{L} \sum_{i=1}^{I_\ell} \omega^*(Z_\ell, I_\ell) \left(\hat{Y}_{i\ell} - \frac{\theta}{(I_\ell - 1)(\overline{v} - \hat{\overline{b}}_N(Z_\ell, I_\ell))} \right)^2, \tag{21}$$

where the optimal weights $\omega^*(Z_\ell, I_\ell)$ are equal to $(I_\ell - 1)(\overline{v}_0 - \overline{b}_0(Z_\ell, I_\ell))$ (see Section 6.2). This estimator can be implemented by a standard two-step procedure in which the optimal weights are first estimated by ordinary NLLS.

[23] We have tried larger numbers such as 30 and 40. The estimated upper boundaries are different but did not much affect the estimation results for θ in the second step.

Another possibility for implementing this estimator is to solve the first-order conditions associated with the maximization of $\hat{Q}_N$ with respect to $(\theta, \overline{v})$. The resulting estimator has the same asymptotic properties as given by Theorem 3. The first-order conditions show that the estimator is similar to an IV estimator of a linear model whose error term is $\hat{\eta}_{i\ell} = \hat{\bar{b}}_N(Z_\ell, I_\ell)(I_\ell - 1)\hat{Y}_{i\ell} - \overline{v}_0(I_\ell - 1)\hat{Y}_{i\ell} + \theta_0$ and instruments $1/(I_\ell - 1)[\overline{v}_0 - \bar{b}_0(Z_\ell, I_\ell)]$ and $1/(I_\ell - 1)[\overline{v}_0 - \bar{b}_0(Z_\ell, I_\ell)]^2$. Alternatively, this IV interpretation can be obtained by writing (20) as

$$(I_\ell - 1)\bar{b}_0(Z_\ell, I_\ell)Y_{i\ell} = -\theta_0 + \overline{v}_0(I_\ell - 1)Y_{i\ell} + \eta_{i\ell},$$

where $\eta_{i\ell} = -(I_\ell - 1)[\overline{v}_0 - \bar{b}_0(Z_\ell, I_\ell)](e_{i\ell} + \epsilon_{i\ell})$. This equation is linear in $(\theta_0, \overline{v}_0)$. The error term $\eta_{i\ell}$ is correlated with the regressor requiring the use of an IV estimator. Following Chamberlain (1987), the optimal instrumental variables are as above.[24] This estimator is implemented through a two-step procedure, where the first step involves a standard IV estimator with a vector of instruments $(1, Z, Z^2, \ldots)$ and $[Y_{i\ell}, \bar{b}_0(Z_\ell, I_\ell)]$ replaced by $[\hat{Y}_{i\ell}, \hat{\bar{b}}_N(Z_\ell, I_\ell)]$. This provides an estimate for $\overline{v}_0$, which can be used for the optimal instruments in the second step.

This estimator has one main advantage. It avoids to maximize the objective function $\hat{Q}_N(\theta, \overline{v})$ while imposing the constraint on the set of parameters embodied in $\mathcal{B}_N$. In particular, constraining $\overline{v}$ to be larger than $\hat{\bar{b}}_N(Z_\ell, I_\ell) + \delta/2$ for all ℓ can raise some problems when the data set contains some outliers. The highest bid in the data, taking a value at 665, is clearly an outlier as the second and third highest bids are in the upper 300 range. Consequently, this constrains $\overline{v}$ to be too large. The above IV estimator circumvents the need for constraining β to be in $\mathcal{B}_N$ and of choosing δ.

Theorem 3 derives the asymptotic distribution of the estimator for θ_0 and $\overline{v}_0$. In practice, it suffices to compute the matrix $\hat{A}_N(\hat{\theta}_N, \hat{\overline{v}}_N)$ and its inverse to obtain the variance of the estimator (see Section 6.2). Using $[1, Z]$ as instruments in the first step, in the second step we find $\hat{\theta}_N = 0.394$ with a standard error equal to 0.286, while $\hat{\overline{v}}_N = 242.20$ with a standard error equal to 2.51. We are interested in testing whether bidders are risk averse, namely whether $\theta < 1$. If $\theta = 1$ (or $c = 0$), bidders are risk neutral. The one-sided

[24]The optimal instruments are defined as $\mathrm{E}[\partial\rho(Y_{i\ell}, Z_\ell, I_\ell, \theta_0, \overline{v}_0)/\partial\beta \mid Z_\ell, I_\ell]/\mathrm{E}[\eta_{i\ell}^2 \mid Z_\ell, I_\ell]$, where $\eta_{i\ell} = \rho(Y_{i\ell}, Z_\ell, I_\ell, \theta_0, \overline{v}_0) = \bar{b}_0(Z_\ell, I_\ell)(I_\ell - 1)Y_{i\ell} + \theta_0 - \overline{v}_0(I_\ell - 1)Y_{i\ell}$. If another specification for the utility function is chosen such as CARA, it would lead to a nonlinear IV estimator.

test of $\theta = 1$ leads us to reject the null hypothesis at 5% with a t-value equal to -2.12. Thus bidders are risk averse with a constant relative risk aversion coefficient of 0.606. This coefficient is close to the one found in the experimental literature at about 0.5. As assuming an upper bound $\overline{v}$ independent of Z can be restrictive, we have tried a linear parameterization of the upper boundary, namely $\overline{v}(Z_\ell) = \gamma_0 + \gamma_1 Z_\ell$, $\ell = 1, \ldots, 378$. The same method as described above applies with some adjustments. In particular there are three parameters to be estimated instead of two. The estimate for γ_1 appears to be non significant and the estimate for θ_0 is similar.

Such 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. Namely, for an auction with 4 bidders (the average number of bidders), a risk aversion parameter at 0.606 is roughly equivalent of having 6 bidders in an auction with risk neutrality.

The third step can be then implemented. Applying the rule of thumb for the constants with triweight kernels and using the appropriate vanishing rates give $h_g = 253.53$ and $h_G = 322.77$. The estimated inverse equilibrium strategy is increasing in b satisfying the restriction imposed by the model on observables as required by Lemma 1. We observe, however, some boundary effects. As a result, some observations need to be trimmed out for the estimation of the underlying conditional density of private values. Specifically, 5 auctions are trimmed out of the original 378 auctions. The estimated conditional density is displayed in Figure 1 and has been obtained with a bandwidth h_f equal to 279.48. The shape roughly follows a log-normal density with some irregularities.

To assess further the impact of risk aversion, we can compute the winners' gain in value $\hat{v}_w - b_w$ and in percentage $(\hat{v}_w - b_w)/v_w$ for the auctions for which a good estimate of the private value is obtained. The results are given in the following table. As expected, the presence of risk aversion, which renders bidding more aggressive, tends to reduce the gain or informational rent for the winners. As a matter of fact, the USFS captures on average about 80% of the bidders' willingness to pay.

Table 2: Winners' Informational Rents

	Mean	STD	Min	Max
Informational Rent in \$ per mbf	84.92	84.99	0.56	992.77
Informational Rents in %	19.83	12.18	3.11	75.73

It is interesting to find an economic rationale for such risk aversion. Bidders' risk aversion in timber auctions has been suspected by many authors. See Athey and Levin (2001) and Baldwin (1995). They observe that the bidders face uncertainty about the exact volume of each species in a lot leading bidders to split their bids across different species. The split of bids is then an indicator of bidder's risk aversion. Risk aversion has also been found in a different data set of timber auctions in which there is no bidding on species and bidders pay for their bids and not for harvested timber. See Perrigne (2001). A reason could arise from the uncertainty of the supply of timber and the limited outside opportunities besides the timber auctions organized by public institutions. In the western regions of the US, the USFS is a large supplier of timber. It is likely that given the uncertainty of the supply outside these auctions, it is important for bidders to win these auctions. Though more empirical studies need to be performed for other sectors, the experimental literature shows that overbidding is frequent though the financial stakes are almost inexistent in the experiments. It seems that risk aversion is a natural component of the agent's behavior when facing uncertainty. See the recent work by Goere, Holt and Palfrey (2002), where the deviations from the risk neutral Nash equilibrium are mainly explained by bidders' risk aversion.

Measuring risk aversion is important for the seller when implementing the auction design. Though the optimal mechanism with risk averse bidders is especially difficult to implement as it involves some complex transfers (see Maskin and Riley (1984) and Matthews (1987)), an optimal posted reserve price can be set to generate more revenue for the seller. For $c \neq 1/I$, the optimal reserve price p_0^* is solution of

$$p_0^* = v_0 + \frac{\frac{1-c}{1-cI}[F^{(I-1)c/(1-c)}(p_0^* \mid z) - F(p_0^* \mid z)]}{f(p_0^* \mid z)},$$

where v_0 is the auctioned object value for the seller. Assuming that v_0 is equal to the USFS appraisal value, we find $\hat{p}_0^*$ equal to approximatively \$93 for a lot with average characteristics in terms of value (\$57) and number of bidders (4). The same estimate conducted for an auction with risk neutral bidders ($c = 0$) would give an optimal reserve price at \$132, which is significantly larger. The idea is that because the bidders tend to bid more aggressively with risk aversion, the precommitment effect does not need to be as important thereby reducing the level of the reserve price that generates the maximum profit for the seller. These figures show that assessing risk neutrality when implementing an optimal reserve price policy when risk aversion prevails may have dramatic effects on seller's profit and revenue.

8 Conclusion

This paper extends the structural analysis of auction data to the case where bidders are risk averse. In particular, the methods developed in this paper 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 parameters 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 problem raised by the theoretical auction model.

Specifically, up to some general smoothness conditions, we show that any bid distribution can be rationalized by some auction model with risk averse bidders. On the other hand, 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 and zero wealth can be considered without loss of explanatory power. Because of these nonidentification results, we propose minimal restrictions to establish identification. Parameterization of the utility function and a conditional quantile of the latent distribution of private values is shown to achieve semiparametric identification of the model. This naturally leads to a semiparametric estimation method involving nonparametric boundary estimation, kernel estimators and weighted nonlinear least squares. We show that our estimator cannot achieve $\sqrt{N}$ consistency unlike many other semiparametric estimators, though it attains the best (optimal) rate for estimating the risk aversion parameters. An illustration of the method is proposed on US Forest Service auction data. It is found that bidders are risk averse.

Many extensions can be entertained based on our methodology. A first interesting extension relates to the practice of random reserve prices in auctions such as in timber, wine, art and web auctions. Within a private value paradigm, Li and Tan (2000) show that the overbidding effect due to risk aversion accentuated by a secret reserve price may dominate the precommitment effect of a posted reserve price. Perrigne (2001) extends the method of the present paper to a model with random reserve prices to assess empirically the gain for the seller of keeping the reserve price secret instead of posting it. Relying on the results developed here, Campo (2002) considers an auction model with heterogeneous bidders, where bidders' characteristics such as capacity and experience may affect their attitude towards risk leading to an asymmetric game. Using construction procurement auction data, she shows that risk aversion is decreasing in bidders' experience. A third extension is conducted by Lu (2002) relying on Eso and White (2001) model in which bidders' private values are stochastic because of the many ex ante uncertainties about the value of the auctioned object. Bidders' risk aversion then affects bidding behavior by introducing a risk premium. The results obtained by Lu (2002) show that identification is more involved though some of his restrictions are similar to those in this paper.

Considering risk aversion in auctions represents a great potential for empirical work. The use of first-price auctions and random reserve prices can be justified by bidders' risk aversion. This concerns a large number of applications as these mechanisms are commonly used. Lastly, from an economic point of view, bidders' risk aversion is reminiscent of having financially constrained bidders as financial constraints represent an extreme form of risk aversion. Che and Gale (1998) study a model in which bidders have two private signals, one for the value of the auctioned object and the other for their financing ability. This situation seems realistic for analyzing business-to-business auctions and represents a promising line of research to explore.

Appendix A

Appendix A establishes Theorem 1 as an immediate application to the case with no conditioning variables of the more general Theorem A1 stated below. Proofs of Lemmas used to prove Theorem A1 are given in Appendix E.

For $R \geq 1$, we consider the (nonparametric) model defined by structures $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R^*$, where $\mathcal{U}_R$ and $\mathcal{F}_R^*$ are as given in Definitions 1 and 4. For any such structure, the next result establishes the existence, uniqueness, and smoothness of the equilibrium strategy $s(\cdot; z, I)$. In addition to obtaining the smoothness of $s(\cdot; \cdot, I)$ with respect to (v, z), which is nontrivial because $s(\cdot; \cdot, \cdot)$ does not have an explicit form in general, its proof is interesting in its own right as it tackles directly the singularity at $\underline{v}$ of the differential equation characterizing $s(\cdot; \cdot, \cdot)$ in contrast to previous work (e.g. Maskin and Riley (1996), Lebrun (1999), Lizzeri and Persico (2000)).

Theorem A1: *Let $\mathcal{I} \subset \{2, 3, \ldots\}$, $R \geq 1$ and $\mathcal{Z}$ be a rectangular compact of $I\!R^d$ with nonempty interior. Suppose that $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R^*$, then there exists a unique (symmetric) equilibrium and its equilibrium strategy $s(\cdot; \cdot, \cdot)$ satisfies:*
(i) $\forall (v, z, I) \in (\underline{v}, \overline{v}] \times \mathcal{Z} \times \mathcal{I}$, $s(v; z, I) < v$, while $s(\underline{v}; z, I) = \underline{v}$,
(ii) $\forall (v, z, I) \in [\underline{v}, \overline{v}] \times \mathcal{Z} \times \mathcal{I}$, $s'(v; z, I) > 0$ with $s'(\underline{v}; z, I) = (I-1)\lambda'(0)/[(I-1)\lambda'(0) + 1] < 1$,
(iii) $\forall I \in \mathcal{I}$, $s(\cdot; \cdot, I)$ admits $R+1$ continuous derivatives on $[\underline{v}, \overline{v}] \times \mathcal{Z}$.

Proof of Theorem A1: For any $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R^*$, it follows from the proof of Theorem 2 in Maskin and Riley (1984) that a (symmetric) Bayesian Nash equilibrium strategy $s(\cdot; z, I)$, when it exists, must be a strictly increasing and continuous function on $[\underline{v}, \overline{v}]$, differentiable on $(\underline{v}, \overline{v}]$, and satisfying the differential equation

$$s'(v; z, I) = (I-1)\frac{f(v|z, I)}{F(v|z, I)}\lambda\left(v - s(v; z, I)\right), \ v \in (\underline{v}, \overline{v}], \tag{A.1}$$

with initial condition $s(\underline{v}; z, I) = \underline{v}$ for every $(z, I) \in \mathcal{Z} \times \mathcal{I}$. Furthermore, they show that such functions are Bayesian Nash equilibria. As seen from the proof of their Theorem 2, existence and uniqueness of the equilibrium strategy crucially depends on the existence and uniqueness of a solution to (A.1). When the reserve price is binding, the latter properties are straightforward from standard existence and uniqueness results for first-order differential equations. However, when the reserve price is nonbinding, the explosive behavior of the ratio $f(v|z, I)/F(v|z, I)$ around $\underline{v}$ prevents the application of these standard results.

The main idea of the proof is to introduce a suitable transformation of (A.1) that is sufficiently regular to establish the existence and uniqueness of a solution to (A.1). To do so, we first show that a Bayesian Nash equilibrium strategy must be continuously differentiable on $[\underline{v},\overline{v}]$ (including at $\underline{v}$) and that it must satisfy properties (i)-(ii) of Theorem A1.

Lemma A1: *Let $\mathcal{I}\subset\{2,3,\dots\}$ and $R\geq 1$. Suppose that $[U,F]\in\mathcal{U}_R\times\mathcal{F}_R^*$. Then, for every $(z,I)\in\mathcal{Z}\times\mathcal{I}$, an equilibrium strategy $s(\cdot;z,I)$, if it exists, is continuously differentiable on $[\underline{v},\overline{v}]$, satisfies properties (i)-(ii) of Theorem A1, and solves (A.1) for $v\in[\underline{v},\overline{v}]$.*

In view of Lemma A1, it is convenient to introduce the following set of functions

$$S_1(\underline{v})=\left\{\begin{array}{c} s\colon [\underline{v},\overline{v}]\to I\!\!R \text{ continuously differentiable with} \\ s(\underline{v})=\underline{v},\ 0<s'(\underline{v})<1, \text{ and } s(v)<v,\ s'(v)>0 \text{ for all } v\in(\underline{v},\overline{v}] \end{array}\right\}, \tag{A.2}$$

since a function in $S_1(\underline{v})$ that solves (A.1) is a Bayesian Nash equilibrium. We then introduce some appropriate changes of variables and operator notations. Assume temporarily that (A.1) has a solution $s(\cdot;z,I)$ in $S_1(\underline{v})$ and, for $u\in[0,1]$, define[25]

$$\begin{aligned} \sigma_I(u;v,z) &= \frac{s\left(\underline{v}+u(v-\underline{v});z,I\right)-\underline{v}}{v-\underline{v}} \text{ for } v>\underline{v}, \qquad \sigma_I(u;\underline{v},z)=s'(\underline{v};z,I)u, \\ \Lambda\left(x;v\right) &= \frac{\lambda\left(\left(v-\underline{v}\right)x\right)}{v-\underline{v}} \text{ for } v>\underline{v}, \text{ and } x\in I\!\!R_+, \qquad \Lambda(x;v)=\lambda'(0)x \text{ for } x\in I\!\!R_- \text{ or } v=\underline{v}, \\ \Phi_I(u;v,z) &= (I-1)\frac{(v-\underline{v})uf\left(\underline{v}+u(v-\underline{v})|z,I\right)}{F\left(\underline{v}+u(v-\underline{v})|z,I\right)} \text{ for } v>\underline{v}, \qquad \Phi_I(u;\underline{v},z)=(I-1). \end{aligned} \tag{A.3}$$

Given $[U,F]\in\mathcal{U}_R\times\mathcal{F}_R^*$, the above functions at $v=\underline{v}$ are obtained by taking their limits as $v\downarrow\underline{v}$. Note also that $\Lambda(x;v)$ is continuously differentiable with respect to $x\in I\!\!R$. Moreover,

$$s(v;z,I)=\underline{v}+(v-\underline{v})\sigma_I(1;v,z)=\underline{v}+(\overline{v}-\underline{v})\sigma_I\left(\frac{v-\underline{v}}{\overline{v}-\underline{v}};\overline{v},z\right),\ v\in[\underline{v},\overline{v}], \tag{A.4}$$

which shows with (A.3) that $\sigma_I(\cdot;\cdot,z)$ and $s(\cdot;z,I)$ are in a one-to-one relationship.[26] In what follows $\sigma_I^{(k)}(u;v,z)$, $\Phi_I^{(k)}(u;v,z)$ and $\Lambda^{(k)}(x;v)$ denote the kth derivatives of $\sigma_I(u;v,z)$, $\Phi_I(u;v,z)$ and $\Lambda(x;v)$ with respect to u or x.

[25]The introduction of the additional variable u is standard when studying the smoothness of the solutions of a first-order differential equation via the Functional Implicit Function Theorem, see Theorem 4.D in Zeidler (1985). In particular, $\sigma_I(\cdot;v,z)$ allows us to study (A.1) in the subinterval $[\underline{v},v]$ of $[\underline{v},\overline{v}]$. Dividing by $v-\underline{v}$ in the definitions of σ_I and Λ regularizes (A.1) as seen later. This technique can also be applied to the more general class of utility functions $U(b,v)$ with $U(\underline{v},\underline{v})=0$ considered in Maskin and Riley (1984).

[26]In particular, the function $s(\cdot;z,I)$ on $[\underline{v},v]$ is given by the function $\sigma_I(u;v,z)$ for $u\in[0,1]$. Note also that $\sigma_I(\cdot;v,z)$ can be computed from $\sigma_I(\cdot;v',z)$ whenever $v'>v$.

We now derive a differential equation for $\sigma_I(\cdot; v, z)$. For $v > \underline{v}$, we have

$$\begin{aligned}
\sigma_I^{(1)}(u; v, z) &= s'\left(\underline{v} + u(v-\underline{v}); z, I\right) \\
&= (I-1)\frac{f\left(\underline{v} + u(v-\underline{v})|z, I\right)}{F\left(\underline{v} + u(v-\underline{v})|z, I\right)}\lambda\left(\underline{v} + u(v-\underline{v}) - s\left(\underline{v} + u(v-\underline{v}); z, I\right)\right) \\
&= (I-1)\frac{u(v-\underline{v})f\left(\underline{v} + u(v-\underline{v})|z, I\right)}{F\left(\underline{v} + u(v-\underline{v})|z, I\right)}\frac{\lambda\left((v-\underline{v})u - (v-\underline{v})\sigma_I(u; v, z)\right)}{u(v-\underline{v})} \\
&= \Phi_I(u; v, z)\frac{\Lambda\left(u - \sigma_I(u; v, z); v\right)}{u}, \qquad \text{(A.5)}
\end{aligned}$$

since $u - \sigma_I(u; v, z) \in I\!R_+$ for all $u \in [0,1]$ by (A.3) when $s(\cdot; z, I) \in S_1(\underline{v})$. That is, $\sigma_I(\cdot; v, z)$ solves the flow of differential equations with "parameters" $(v, z, I) \in (\underline{v}, \overline{v}] \times \mathcal{Z} \times \mathcal{I}$

$$E_I(v, z) \;\; : \;\; \sigma^{(1)}(u) = \Phi_I(u; v, z)\frac{\Lambda\left(u - \sigma(u); v\right)}{u}, \; u \in [0,1], \qquad \text{(A.6)}$$

and initial condition $\sigma_I(0; v, z) = 0$ by (A.3) and $s(\underline{v}; z, I) = \underline{v}$. Moreover, $E_I(\underline{v}, z)$ can be defined by taking the limit as $v \downarrow \underline{v}$, giving

$$E_I(\underline{v}, z) \;\; : \;\; \sigma^{(1)}(u) = \frac{(I-1)}{u}\lambda'(0)\left(u - \sigma(u)\right), \; u \in [0,1].$$

A comparison of $E_I(\underline{v}, z)$ with (A.1) shows that the first-order differential equation $E_I(\underline{v}, z)$ corresponds to a CRRA model with constant relative risk aversion parameter $c = 1 - 1/\lambda'(0) \in [0,1)$ since $\lambda'(0) \geq 1$, with uniform distribution on $[0,1]$ of the private value. Hence, $E_I(\underline{v}, z)$ has a *unique* solution satisfying $\sigma_I(0; \underline{v}, z) = 0$, namely

$$\sigma_I(u; \underline{v}, z) = \frac{(I-1)\lambda'(0)}{(I-1)\lambda'(0) + 1}u, \; u \in [0,1]. \qquad \text{(A.7)}$$

Now, solving (A.6) is actually solving (A.1) in the neighborhood of $\underline{v}$. In particular, for each $(v, z, I) \in [\underline{v}, \overline{v}] \times \mathcal{Z} \times \mathcal{I}$, solving (A.6) with initial condition $\sigma_I(0; v, z) = 0$ can be viewed as finding zero(s) of the operator

$$\mathbf{E}_I(\cdot; v, z) \, : \, \sigma \in C_1^0 \to \mathbf{E}_I(\sigma; v, z) = \sigma^{(1)}(u) - \Phi_I(u; v, z)\frac{\Lambda\left(u - \sigma(u); v\right)}{u}, \; u \in [0,1], \qquad \text{(A.8)}$$

where

$$C_1^0 = \left\{\sigma : [0,1] \to I\!R \, ; \, \sigma(0) = 0 \, \text{and} \, \sigma \, \text{continuously differentiable on} \, [0,1]\right\},$$

and $\mathbf{E}_I(\cdot; \underline{v}, z)$ is defined by taking the limit as for $E_I(\underline{v}, z)$ above. Note that $\mathbf{E}_I(\sigma; v, z)$ is a function defined on $[0,1]$. Moreover, for any $v \in [\underline{v}, \overline{v}]$ and $\sigma \in C_1^0$, the definition of $\Lambda(x; v)$ implies

$\Lambda\left(u-\sigma(u);v\right)/u \sim \lambda'(0)(u-\sigma(u))/u \to \lambda'(0)(1-\sigma'(0))$ when $u \downarrow 0$, since $\lim_{u\downarrow 0}\sigma(u)/u = \sigma^{(1)}(0)$. Hence, $\mathbf{E}_I(\cdot;v,z)$ is an operator, parameterized by (v,z,I), from C_1^0 to

$$C_0 = \{\zeta : [0,1] \to I\!R\,;\ \zeta \text{ continuous on } [0,1]\}\,.$$

Define the norms

$$\|\sigma\|_{1,\infty} = \max_{r=0,1}\left\|\sigma^{(r)}\right\|_\infty = \left\|\sigma^{(1)}\right\|_\infty, \quad \|\zeta\|_\infty = \sup_{u\in[0,1]}|\zeta(u)|\,,$$

using $|\sigma(u)| = \left|\int_0^u \sigma^{(1)}(y)dy\right| \le \|\sigma^{(1)}\|_\infty$ for all $u \in [0,1]$ to obtain $\|\sigma\|_{1,\infty} = \|\sigma^{(1)}\|_\infty$. Hence, C_1^0 equipped with $\|\cdot\|_{1,\infty}$ and C_0 equipped with $\|\cdot\|_\infty$ are Banach spaces.

The idea of our proof is as follows. In a first step, using that (A.7) is the unique zero of $\mathbf{E}_I(\cdot;\underline{v},z)$, we show that (A.8) has a unique zero for v in a small interval $[\underline{v},v_0]$ using a continuation argument given by Proposition 6.10 in Zeidler (1985). Such an argument requires so-called a priori conditions on the zeros of (A.8) as summarized by the set of functions

$$\Sigma_1^0 = \left\{\sigma \in C_1^0;\ 0 < \sigma'(0) < 1 \text{ and } \sigma'(u) > 0,\, \sigma(u) < u \text{ for } u \in (0,1]\ \right\},$$

which is the counterpart of $S_1(\underline{v})$ by (A.3) and (A.5). Note that (A.7) is in Σ_1^0, which is an open subset of C_1^0 since the open ball $V(\sigma;\epsilon) = \{\zeta \in C_1^0; \|\zeta-\sigma\|_{1,\infty} < \epsilon\} \subset \Sigma_1^0$ for any $\sigma \in \Sigma_1^0$, provided that $\epsilon = \epsilon_\sigma > 0$ is sufficiently small so that such ζ's satisfy the constraints defining Σ_1^0. Because of the relationship between $\sigma_I(\cdot;v_0,z)$ and $s(\cdot;z,I)$, this step gives us a unique solution of (A.1) on $[\underline{v},v_0]$. In a second step, we apply the standard Picard-Lindelöf Theorem to extend the unique solution of (A.1) for $v \in [\underline{v},v_0]$ to $v \in [v_0,\overline{v}]$, where the ratio $f(v|z,I)/F(v|z,I)$ remains bounded, using the value of the local solution at v_0 as an initial condition. These two steps are done in Lemma A4. The smoothness of $s(v;z,I)$ as a function of the "parameters" (v,z) will be given by a Functional Implicit Function Theorem 4.B in Zeidler (1985). This is done in Lemmas A5 and A6.

To apply the aforementioned theorems, we need to study the partial derivatives of the operator $\mathbf{E}_I(\sigma;v,z)$ with respect to (σ,v,z). This is the purpose of Lemma A3, which relies on some smoothness properties of $\Lambda(x;v)$ and $\Phi_I(u;v,z)$ summarized in the next lemma.

Lemma A2: *Under the conditions of Theorem A1, let $[U,F] \in \mathcal{U}_R \times \mathcal{F}_R^*$. Then*
(i) $\Lambda(x;v)$ is R continuously differentiable with respect to $(x,v) \in I\!R_+ \times [\underline{v},\overline{v}]$, Moreover, the function $(1/x)\partial^{r_2}\Lambda(x;v)/\partial v^{r_2}$ is continuous on $I\!R_+ \times [\underline{v},\overline{v}]$, for $r_2 = 0,\ldots,R$,
(ii) For every $I \in \mathcal{I}$, $\Phi_I(u;v,z)$ is R continuously differentiable with respect to $(v,z) \in [\underline{v},\overline{v}] \times \mathcal{Z}$ and the partial derivative $\partial^{r_2+r_3}\Phi_I(u;v,z)/\partial v^{r_2}\partial z^{r_3}$, $r_2+r_3 = 0,\ldots,R$, is continuous with respect to $(u,v,z) \in [0,1] \times [\underline{v},\overline{v}] \times \mathcal{Z}$.

For every $I \in \mathcal{I}$, we now compute the partial derivatives of the operator $\mathbf{E}_I(\sigma; v, z)$ when $(\sigma, v, z) \in \Sigma_1^0 \times [\underline{v}, \overline{v}] \times \mathcal{Z}$. Because σ is a function, it is necessary to use the notions of Fréchet and Gâteaux derivatives (see e.g. Zeidler (1985)). In what follows

$$\mathbf{E}_I^{r_1 r_2 r_3}(\sigma; v, z) = \frac{\partial^{r_1+r_2+r_3}\mathbf{E}_I(\sigma; v, z)}{\partial\sigma^{r_1}\partial v^{r_2}\partial z^{r_3}},\ 0 \le r_1 + r_2 + r_3 \le R,$$

denotes the partial derivatives of $\mathbf{E}_I(\sigma; v, z)$, which are linear operators from $(C_1^0)^{r_1} \times \mathbb{R}^{r_2} \times (\mathbb{R}^d)^{r_3}$ taking values in C_0 as $\mathbf{E}_I(\sigma; v, z)$ does. Expressions for the operators $\mathbf{E}_I^{r_1 r_2 r_3}(\sigma; v, z)$, $r_1 \ge 1$ are easier found using Gâteaux differentiation, namely

$$\mathbf{E}_I^{r_1 r_2 r_3}(\sigma; v, z)(\eta, \ldots, \eta)(u) = \frac{d^{r_1}}{dt^{r_1}}\left(\frac{\partial^{r_2+r_3}\mathbf{E}_I(\sigma + t\eta; v, z)}{\partial v^{r_2}\partial z^{r_3}}\right)_{t=0}(u),\ \eta \in C_1^0,$$

and then changing the term η^{r_1} appearing in the above partial derivative into $\eta_1 \times \cdots \times \eta_{r_1}$ to obtain a multilinear form of order r_1.[27] Specifically, using (A.8) and Lemma A2, we obtain

$$\mathbf{E}_I^1(\sigma; v, z)(\eta_1)(u) = \mathbf{E}_I^{1,0,0}(\sigma; v, z)(\eta_1)(u) = \eta_1^{(1)}(u) + \Phi_I(u; v, z)\Lambda^{(1)}\left(u - \sigma(u); v\right)\frac{\eta_1(u)}{u}. \quad \text{(A.9)}$$

Because $\eta_1(y)/y$ is continuous over $[0,1]$ by definition of C_1^0, Lemma A2 shows that $\mathbf{E}_I^1(\sigma; v, z)(\eta_1)$ belongs to C_0 as required. More generally, for $r_1 \ge 1$ and $(r_2, r_3) \neq (0,0)$, (A.8) gives

$$\begin{aligned}&\mathbf{E}_I^{r_1 r_2 r_3}(\sigma; v, z)(\eta_1, \ldots, \eta_{r_1})(u)\\ &\quad = (-1)^{r_1+1}\frac{\partial^{r_2+r_3}}{\partial v^{r_2}\partial z^{r_3}}\left(\Phi_I(u; v, z)\Lambda^{(r_1)}\left(u - \sigma(u); v\right)\right)\frac{\eta_1(u) \times \cdots \times \eta_{r_1}(u)}{u},\end{aligned}$$

which again is in C_0. On the other hand, if $r_1 = 0$, we have by the Liebnitz-Newton formula

$$\begin{aligned}\mathbf{E}_I^{0 r_2 r_3}(\sigma; v, z)(u) &= -\sum_{j=0}^{r_2}\frac{r_2!}{j!(r_2-j)!}\frac{1}{u}\frac{\partial^j\Lambda\left(u - \sigma(u); v\right)}{\partial v^j}\frac{\partial^{r_2+r_3-j}\Phi_I(u; v, z)}{\partial v^{r_2-j}\partial z^{r_3}}\\ &= \frac{\sigma(u) - u}{u}\sum_{j=0}^{r_2}\frac{r_2!}{j!(r_2-j)!}\frac{1}{u - \sigma(u)}\frac{\partial^j\Lambda\left(u - \sigma(u); v\right)}{\partial v^j}\frac{\partial^{r_2+r_3-j}\Phi_I(u; v, z)}{\partial v^{r_2-j}\partial z^{r_3}},\end{aligned}$$

which belongs to C_0 because $(\sigma(u) - u)/u \in C_0$, $u - \sigma(u) \ge 0$ since $\sigma \in \Sigma_1^0$, and Lemma A2.

[27] Note that the operator $\mathbf{E}_I(\sigma; v, z)$ depends upon $\Lambda(u - \sigma(u); v)$ and that $\Lambda(x; v)$ has derivatives with respect to x for $x \in \mathbb{R}_+$ by Lemma A2 but not necessarily for $x \in \mathbb{R}_-$. However, $\sigma + t\eta \in \Sigma_1^0$ for t sufficiently small because Σ_1^0 is open. The definition of Σ_1^0 then yields that $u - \sigma(u) - t\eta(u) \ge 0$ for all $u \in [0,1]$ and the Gâteaux differentiation above is correct because we restrict σ to belong to Σ_1^0, which is sufficient for our purpose. Derivation over C_1^0 can be achieved by defining $\Lambda(x; v)$ for $x < 0$ using the $R + 1$ Taylor expansion of $\lambda(\cdot)$ at 0 instead of its first-order Taylor expansion as done here.

The next lemma establishes that the above Gâteaux derivatives are the Fréchet derivatives of $\mathbf{E}_I(\sigma; v, z)$ on $\Sigma_1^0 \times [\underline{v}, \overline{v}] \times \mathcal{Z}$. Moreover, it checks the conditions of the Functional Implicit Function Theorem, which will be a key tool to study the existence, uniqueness and smoothness with respect to (v, z) of the solution to (A.1).

Lemma A3: *Under the conditions of Theorem A1, let* $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R^*$. *Then, for every* $I \in \mathcal{I}$,
(i) $\mathbf{E}_I(\sigma; v, z)$ *takes its values in* C_0 *for every* $(\sigma, v, z) \in C_1^0 \times [\underline{v}, \overline{v}] \times \mathcal{Z}$,
(ii) $\mathbf{E}_I(\sigma; v, z)$ *is* R *continuously differentiable on* $\Sigma_1^0 \times [\underline{v}, \overline{v}] \times \mathcal{Z}$, *with Fréchet partial derivatives* $\mathbf{E}_I^{r_1 r_2 r_3}(\sigma; v, z)$ *as given above that are uniformly continuous over* $\Sigma_1^0 \times [\underline{v}, \overline{v}] \times \mathcal{Z}$, *provided* $0 \leq r_1 + r_2 + r_3 \leq R$,
(iii) $\mathbf{E}_I^1(\sigma; v, z)$ *is a one-to-one mapping from* C_1^0 *to* C_0 *for every* $(\sigma, v, z) \in \Sigma_1^0 \times [\underline{v}, \overline{v}] \times \mathcal{Z}$.

We are now ready to prove Theorem A1. Lemma A1 combined with the next lemma establishes that there exists a unique Bayesian Nash equilibrium strategy, which moreover satisfies properties (i)-(ii) of Theorem A1.[28]

Lemma A4: *Under the conditions of Theorem A1, let* $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R^*$. *Then, for every* $(z, I) \in \mathcal{Z} \times \mathcal{I}$, *the first-order differential equation (A.1) has a unique continuously differentiable solution on* $[\underline{v}, \overline{v}]$ *with initial condition* $s(\underline{v}; z, I) = \underline{v}$. *Moreover, this solution satisfies properties (i)-(ii) of Theorem A1.*

It remains to establish property (iii) of Theorem A1, i.e. the smoothness of $s(v; z, I)$ with respect to (v, z). This is done in the following two lemmas using a Functional Implicit Function Theorem.

Lemma A5: *Under the conditions of Theorem A1, let* $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R^*$. *For every* $(z, I) \in \mathcal{Z} \times \mathcal{I}$, *let* $s(\cdot; z, I)$ *be the unique solution of (A.1) with* $s(\underline{v}; z, I) = \underline{v}$. *Then, for every* $I \in \mathcal{I}$, *the functions* $s(v; z, I)$ *and* $s'(v; z, I)$ *are* R *continuously differentiable with respect to* $(v, z) \in [\underline{v}, \overline{v}] \times \mathcal{Z}$.

Lemma A5 yields the existence and continuity of all partial derivatives up to order $(R+1)$ of $s(\cdot; \cdot, I)$, with the exception of $\partial^{R+1} s(\cdot; \cdot, I)/\partial z^{R+1}$. Thus, it remains to consider the latter to complete the proof of Theorem A1-(iii). Let $S_I(G)$ be as in Lemma 2.

Lemma A6: *Under the conditions of Theorem A1, let* $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R^*$. *Then, for every* $I \in \mathcal{I}$,

[28] Lemma A4 can be established under weaker conditions as seen from its proof. First, it is possible to assume $\overline{v} = +\infty$. Second, the density $f(\cdot|z, I)$ can vanish at $\underline{v}$ provided $\lim_{v \downarrow \underline{v}} (v - \underline{v}) f(v|z, I)/F(v|z, I) \in (0, +\infty)$. These conditions weaken the ones used in Maskin and Riley (1996) and Lebrun (1999).

(i) The conditional bid distribution $G(b|z,I) = F(s^{-1}(b;z,I)|z,I)$ admits up to $R+1$ continuous partial derivatives on $S_I(G)$, with $\inf_{(b,z)\in S_I(G)} g(b|z,I) > 0$ and a support $[\underline{b}(z,I), \overline{b}(z,I)]$ satisfying $\inf_{z\in\mathcal{Z}} \left(\overline{b}(z,I) - \underline{b}(z,I)\right) > 0$ and $\underline{b}(z,I) = \underline{v}$,

(ii) The function $s(\cdot;\cdot,I)$ admits up to $R+1$ continuous partial derivatives on $[\underline{v},\overline{v}] \times \mathcal{Z}$.

Appendix B

Appendix B gathers proofs of Lemma 1 and Propositions 1–3 stated in Sections 3 and 4.

Proof of Lemma 1: First, we prove that conditions (i), (ii) and (iii) are necessary. Because $b_i = s(v_i, U, F, I)$ and the v_is are i.i.d., it follows that the b_is are i.i.d. so that (i) must hold. Condition (ii) follows from applying Lemma 2 to the case with no conditioning variables (Z, I).

To prove that condition (iii) is also necessary, consider (4), where the function $\lambda(\cdot)$ is the ratio $U(\cdot)/U'(\cdot)$. 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. Moreover, $U(\cdot)$ admits $R+2$ continuous derivatives on $(0,+\infty)$. As $\lim_{x\downarrow 0} \lambda^{(r)}$ is finite for $r = 1,\ldots,R+1$, these imply that $\lambda(\cdot)$ 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 the function $\xi(\cdot)$ is increasing. 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 $\xi'(\cdot) = [s^{-1}(\cdot)]'$, which is strictly positive by Theorem 1.

Second, we have to show that conditions (i), (ii) and (iii) are together sufficient. Assume that bids are independently and identically distributed as $G(\cdot) \in \mathcal{G}_R$ and there exists a function $\lambda(\cdot)$ satisfying the properties of Lemma 1. First, we construct a pair $[U, F]$ belonging to $\mathcal{U}_R \times \mathcal{F}_R$. Let $U(\cdot)$ be such that $\lambda(\cdot) = U(\cdot)/U'(\cdot)$ or $U'(\cdot)/U(\cdot) = 1/\lambda(\cdot)$. By integrating, we obtain $U(x) = U(a)\exp[\int_a^x 1/\lambda(t)dt]$ for an arbitrary $a > 0$. With the normalization $U(1) = 1$, this gives $U(x) = \exp\int_1^x 1/\lambda(t)dt$. We verify that such a utility function belongs to $\mathcal{U}_R$. Because $\lambda(\cdot)$ admits $R+1$ continuous derivatives on $[0,+\infty)$, then condition (iii) of Definition 1 is clearly satisfied. Moreover, in the neighborhood of zero, $\lambda(t) \sim \lambda'(0)t$ with $1 \leq \lambda'(0) < \infty$. Thus the integral $\int_x^1 1/\lambda(t)dt$ diverges to infinity, which implies that $U(x)$ tends to zero as $x \downarrow 0$. Define $U(0) = 0$. The derivative $U'(x)$ is equal to $\exp\int_1^x 1/\lambda(t)dt/\lambda(x)$, where $\lambda(\cdot) > 0$ on $(0,+\infty)$ because $\lambda(0) = 0$ and $\lambda'(\cdot) \geq 1$. This implies that $U'(\cdot) > 0$ on $(0,+\infty)$. The second-order derivative gives $U''(x) = (-\lambda'(x)+1)\exp\int_1^x 1/\lambda(t)dt/\lambda^2(x)$. We know that $\lambda'(x) \geq 1$, which implies that $U''(\cdot) \leq 0$ on $(0,+\infty)$. It remains to show that $U(\cdot)$ admits $R+2$ continuous derivatives. By assumption, $\lambda(\cdot)$ has $R+1$ continuous derivatives on $[0,+\infty)$. It follows that

$U(\cdot)$ admits $R+2$ continuous derivatives on $(0,+\infty)$. Lastly, $U(\cdot)$ is continuous on $(0,+\infty)$ as $\lim_{x\downarrow 0} U(x) = U(0) = 0$.

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 such a distribution $F(\cdot)$ belongs to $\mathcal{F}_R$. We have $F(x) = \Pr(X \leq x) = \Pr(\xi(b) \leq x)$. The latter can be written as $\Pr(b \leq \xi^{-1}(x))$, which is equal to $G[\xi^{-1}(x)]$, because $\xi(\cdot)$ is strictly increasing by assumption. This implies that $F(\cdot) = G(\xi^{-1}(\cdot))$ 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, then $F(\cdot)$ is strictly increasing on $[\underline{v}, \overline{v}]$ and its support is $[\underline{v}, \overline{v}]$, which is a finite interval of $I\!R_+$. ¿From Definition 3, $G(\cdot)$ has $R+1$ continuous derivatives on $[\underline{b}, \overline{b}]$. 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 $[\underline{v}, \overline{v}]$. It remains to show that the corresponding density $f(\cdot)$ is strictly positive. We have $f(\cdot) = g(\xi^{-1}(\cdot))/\xi'(\xi^{-1}(\cdot))$, where $g(\cdot) > 0$ from Definition 3 and $\xi'(\cdot)$ is finite on $[\underline{b}, \overline{b}]$. Thus $f(\cdot) > 0$ on $[\underline{v}, \overline{v}]$.

Lastly, we have to show that the pair $[U, F]$ can rationalize $G(\cdot)$ in a first-price sealed-bid auction with risk averse bidders, 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 $F(\cdot)$, we have $G(\cdot) = F(\xi(\cdot))$. 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 $\xi(\underline{b}) = \underline{b} = \underline{v}$. ¿From the construction of $F(\cdot)$, we have $f(\cdot)/F(\cdot) = [\xi^{-1}(\cdot)]' g(\xi^{-1}(\cdot))/G(\xi^{-1}(\cdot))$. Thus $\xi^{-1}(\cdot)$ solves the differential equation (2) if

$$1 = (I-1)\frac{g(\xi^{-1}(v))}{G(\xi^{-1}(v))}\lambda(v - \xi^{-1}(v)),$$

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 the differential equation (2) with boundary condition $\xi^{-1}(\underline{v}) = \underline{v}$.

Proof of Proposition 1: (i) Consider a bid distribution $G(\cdot) \in \mathcal{G}_R$ generated by a structure $[U, F] \in \mathcal{U}_R \times \mathcal{F}_R$. We have to show that there exists a structure $[\tilde{U}, \tilde{F}]$, where $\tilde{U}(x) = x^{1-c}$, $0 \leq c < 1$ and $\tilde{F} \in \mathcal{F}_R$, that rationalizes the distribution $G(\cdot)$. In this case, $\tilde{\lambda}(x) = x/(1-c)$ so that $\tilde{\lambda}(0) = 0$ and $\tilde{\lambda}'(\cdot) \geq 1$. ¿From Lemma 1, it suffices to show that there exists a value $c \in [0,1)$ such that the function

$$\xi(b, c, G) = b + \frac{1-c}{I-1}\frac{G(b)}{g(b)},$$

has a strictly positive derivative on $[\underline{b},\overline{b}]$. Differentiating, this is equivalent to $[G(b)/g(b)]' > -(I-1)/(1-c)$ for all $b \in [\underline{b},\overline{b}]$. The latter is true if

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

Note that the left-hand side is finite because $G(\cdot)/g(\cdot)$ is $R+1$ continuously differentiable on $[\underline{b},\overline{b}]$, as noted after Definition 3. We consider two cases. If $\inf_b[G(b)/g(b)]' \geq 0$, then we can choose any value $c \in (0,1)$ to satisfy (B.1). Second, if $\inf_b[G(b)/g(b)]' < 0$, (B.1) can be written as $c > 1-(I-1)/(-\inf_b[G(b)/g(b)]')$, where the right-hand side is less than one. Thus we can always find a value for $c \in (0,1)$ such that $G(\cdot)$ can be rationalized by a CRRA model.

(ii) The proof for the CARA case is similar. Consider $\tilde{U} \in \mathcal{U}_R^{CARA}$. This gives the utility function $\tilde{U}(x) = (1-e^{-ax})/(1-e^{-a})$ with $a > 0$. Hence $\tilde{\lambda}(x) = (e^{ax}-1)/a$ and $\tilde{\lambda}^{-1}(x) = (1/a)\log(1+ax)$. This defines the following inverse bidding strategy

$$\xi(b) = b + \frac{1}{a}\log\left(1 + \frac{a}{I-1}\frac{G(b)}{g(b)}\right).$$

We have to show that there exists $a > 0$ such that $\xi'(b) > 0$ on $[\underline{b},\overline{b}]$. Differentiating gives the following inequality on a

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

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 $R \geq 1$ and $g(b) > 0$. Hence the preceding inequality holds at $\underline{b}$ for any $a > 0$. Thus the preceding inequality becomes

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

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 $R \geq 1$. 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 a value for a and hence a CARA model that can rationalize any bid distribution $G(\cdot)$.

Proof of Proposition 2: (i) Nonidentification of the general model. Consider a structure $[U,F] \in \mathcal{U}_R \times \mathcal{F}_R$, which generates a bid distribution $G(\cdot) \in \mathcal{G}_R$ by Lemma 1. Suppose first that $U(\cdot)$ is not of the form x^{1-c} for any c, $0 \leq c < 1$. ¿From Proposition 1, it follows that there exists a CRRA structure $[\tilde{U},\tilde{F}]$ with zero wealth and $\tilde{F} \in \mathcal{F}_R$ that leads to the same equilibrium bid distribution $G(\cdot)$. Because a CRRA utility function with zero wealth belongs to $\mathcal{U}_R$, the original structure $[U, F]$ is not identified. Suppose next that $U(\cdot)$ is of the form x^{1-c} for some c, $0 \leq c < 1$. From the second part of the proposition, which is proven below, there exists another CRRA structure with zero wealth, and hence another risk aversion structure $[\tilde{U}, \tilde{F}]$ with $\tilde{F} \in \mathcal{F}_R$ that is observationally equivalent to $[U, F]$. Hence $[U, F]$ is again not identified.

(ii) Nonidentification of the CRRA, CARA, DRRA and DARA models. We first show that the CRRA model is not identified. Consider a structure $[U, F]$ where $U(\cdot)$ is derived from a CRRA vNM utility function with some wealth $w \geq 0$ and $F \in \mathcal{F}_R$. This generates a bid distribution $G(\cdot) \in \mathcal{G}_R$. Proposition 1 shows that there exist a CRRA utility function $\tilde{U}(\cdot)$ with zero wealth and $0 \leq \tilde{c} < 1$ and a distribution $\tilde{F}(\cdot) \in \mathcal{F}_R$ leading to the bid distribution $G(\cdot)$. If $U(\cdot)$ is not of the form x^{1-c}, then $[U, F]$ is not identified. If $U(\cdot)$ is of the form x^{1-c}, the proof of Proposition 1 shows that there exists an infinity of values for $\tilde{c}$, $c < \tilde{c} < 1$, generating the same distribution $G(\cdot)$. Thus the CRRA model is unidentified. We can use a similar argument to show that the CARA model is unidentified from the proof of Proposition 1-(ii).

Next, consider a structure $[U, F] \in \mathcal{U}_R^{DRRA} \times \mathcal{F}_R$ defining a DRRA model and generating a bid distribution $G(\cdot) \in \mathcal{G}_R$. Note that $\mathcal{U}_R^{CRRA} \subset \mathcal{U}_R^{DRRA}$. If $U(\cdot)$ is generated from a vNM utility function with constant relative risk aversion, we know from above that there exists another CRRA structure that is observationally equivalent to $[U, F]$. On the other hand, if $U(\cdot)$ is generated from a vNM utility function with partly strictly decreasing relative risk aversion, we know from Proposition 1 that there exists an observationally equivalent CRRA structure with zero wealth. Thus the DRRA model is unidentified. A similar argument shows that the DARA model is unidentified.

Proof of Proposition 3: We distinguish two parts. The first part concerns the identification of the general semiparametric model composed of structures $[U, F]$ satisfying Assumption A1, while the second part concerns the identification of the CRRA and CARA models.

Part 1. 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 conditional bid distribution. We first show that (θ, γ) is identified, i.e $(\theta, \gamma) = (\tilde{\theta}, \tilde{\gamma})$. Writing (5) for each structure gives

$$\frac{1}{I-1} \frac{\alpha}{g(b_\alpha(z, I)|z, I)} = \lambda(v_\alpha(z, I; \gamma) - b_\alpha(z, I); \theta) = \lambda(v_\alpha(z, I; \tilde{\gamma}) - b_\alpha(z, I); \tilde{\theta}), \tag{B.2}$$

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

$$v = b + \lambda^{-1}\left[\frac{1}{I-1}\frac{G(b|z,I)}{g(b|z,I)};\theta\right] = \tilde{v},$$

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.

PART 2. We have $U(x) = x^{1-c}$ with $0 \leq c < 1$ for the CRRA model and $U(x) = (1 - \exp^{-ax})/(1-\exp^{-a})$ with $a > 0$ for the CARA model. Conditions (i)–(iii) of Assumption A1 are satisfied, where $v_\alpha(z,i) = \gamma$. Thus, it suffices to verify condition (iv).

For the CRRA model, we have $\lambda(\gamma - b_\alpha(z,I);\theta) = (\gamma - b_\alpha(z,I))/\theta$, where $\theta = 1-c$. By assumption there exist two pairs (z_1, I_1) and (z_2, I_2) belonging to $\mathcal{Z} \times \mathcal{I}$ such that $b_\alpha(z_1,I_1) \neq b_\alpha(z_2,I_2)$. Hence, $(\gamma - b_\alpha(z_1,I_1))/\theta \neq (\gamma - b_\alpha(z_2,I_2))/\theta$. On the other hand, knowing the function $\lambda(\gamma - b_\alpha(\cdot,\cdot);\theta)$ for every $(z,I) \in \mathcal{Z} \times \mathcal{I}$ and hence for (z_1,I_1) and (z_2,I_2) gives a system of two linear equations in two unknown parameters (θ,γ). Because the determinant of such a system is not equal to zero, there is a unique solution.

For the CARA model, we have $\lambda(\gamma - b_\alpha(z,I);\theta) = (\exp^{\theta(\gamma - b_\alpha(z,I))} - 1)/\theta$, where $\theta = a$. By assumption there exist two pairs (z_1,I_1) and (z_2,I_2) belonging to $\mathcal{Z} \times \mathcal{I}$ such that $b_\alpha(z_1,I_1) \neq b_\alpha(z_2,I_2)$. Hence, $\lambda_1 \neq \lambda_2$, where $\lambda_j = (\exp^{\theta(\gamma - b_\alpha(z_j,I_j))} - 1)/\theta$ for $j = 1,2$. Rearranging terms, eliminating γ and taking the logarithm give

$$\log\frac{1+\theta\lambda_2}{1+\theta\lambda_1} = \theta[b_\alpha(z_1,I_1) - b_\alpha(z_2,I_2)],$$

where $b_\alpha(z_1,I_1) > b_\alpha(z_2,I_2)$ without loss of generality and hence $\lambda_2 > \lambda_1 > 0$. Differentiating twice with respect to θ the left-hand side shows that the left-hand side is strictly increasing and concave in θ on $[0,+\infty)$. Because one root of the above equation is $\theta = 0$, there is at most one other strictly positive root. Thus, θ is uniquely determined, which gives a unique γ.

Appendix C

Appendix C gathers proofs of Lemma 2 and Theorem 2 stated in Section 5. Proofs of Lemmas used to prove Theorem 2 are given in Appendix E. Throughout, let $\xi(\cdot;z,I) = s^{-1}(\cdot;z,I)$.

Proof of Lemma 2: We have $\overline{b}(z,I) = s(\overline{v};z,I)$, which admits $R+1$ continuous derivatives on $\mathcal{Z}$ by Theorem A1-(iii). The other assertions in (i)–(iii) follow from Lemma A6-(i).

To prove (iv), we note that $\xi(b;z,I) = F^{-1}[G(b|z,I)|z,I]$. Hence, $\xi(\cdot;\cdot,I)$ is $R+1$ continuously differentiable on $S_I(G)$. Moreover, we have from (4)

$$g(b|z,I) = \frac{1}{I-1}\frac{G(b|z,I)}{\lambda\left(\xi(b;z,I) - b\right)} \tag{C.1}$$

with $\lambda\left(\xi(b;z,I)-b\right)>0$ because $s(v;z,I)<v$ whenever $v>\underline{v}$ by Theorem A1. Therefore, Lemma A6 and the composition rule for differentiation give that $g(\cdot|\cdot,I)$ admits up to $R+1$ continuous partial derivatives on $S_I^u(G)$, which establishes (iv).

It remains to show (v). Omit the dependence on (z,I) to simplify the notation. From Theorem 1, it follows that $s^{-1}(\cdot)$ admits $R+1$ continuous derivatives on $[\underline{b},\overline{b}]$ with $[s^{-1}(\cdot)]'>0$ on $[\underline{b},\overline{b}]$. Now, from (3), we have $G(b)/g(b)=(I-1)\lambda[s^{-1}(b)-b]$. Because $\lambda(\cdot)$ and $s^{-1}(\cdot)$ are $R+1$ continuously differentiable on $[0,+\infty)$ and $[\underline{b},\overline{b}]$, then $G(\cdot)/g(\cdot)$ is $R+1$ continuously differentiable on $[\underline{b},\overline{b}]$, and hence admits a finite limit as $b\downarrow\underline{b}$.

Proof of Theorem 2: We begin by studying the smoothness of $m(z,I;\beta)$, as summarized in the next lemma.

Lemma C1: *Let (U_0,F_0) satisfy Assumption A2-(i,ii) for some $\beta_0=(\theta_0,\overline{v}_0)\in\Theta^o\times(0,\infty)$ and $\mathcal{I}$ finite. Let $\overline{b}_0(z,I)$ be the upper boundary of the support of the corresponding bid distribution $G_0(\cdot|z,I)$, where $(z,I)\in\mathcal{Z}\times\mathcal{I}$. Then, for every $I\in\mathcal{I}$, the function $m(z,I;\beta)$ defined in (7) is $R+1$ continuously differentiable on $\mathcal{Z}\times\mathcal{B}$, where $\mathcal{B}=\{(\theta,\overline{v});\theta\in\Theta,\overline{v}>\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\overline{b}_0(z,I)\}$ with $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\overline{b}_0(z,I)<\overline{v}_0$.*

Now, let $t_N\geq 0$ be such that $t_N/\rho_N=o(1)$, and $\psi(\cdot):\mathbb{R}_-\rightarrow\mathbb{R}$ be an infinitely differentiable function on $\mathbb{R}_-$ with support $[-1,0]$, such that $\psi(0)=1$, $\int\psi(x)dx=0$. Let $\mathbb{1}_p=(1,\ldots,1)'\in\mathbb{R}^p$. For a fixed constant $\kappa>0$ to be chosen below, consider the following perturbations of θ_0 and $g_0(b|z,I)$, $I\in\mathcal{I}$,

$$\begin{aligned}
\beta_1 &= (\theta_1,\overline{v}_0)=(\theta_0+2t_N\mathbb{1}_p/\rho_N,\overline{v}_0),\\
g_1(b|z,I) &= g_0(b|z,I)+\pi_N(z,I)\psi\left[\kappa\rho_N^{1/(R+1)}\left(b-\overline{b}_0(z,I)\right)\right],\\
\pi_N(z,I) &= m\left(z,I;\beta_1\right)-m(z,I;\beta_0)=\frac{\partial m(z,I;\beta_0)}{\partial\beta}\left(\beta_1-\beta_0\right)+o(\|\beta_1-\beta_0\|)=O(1/\rho_N).
\end{aligned}$$

Note that $\{(\theta,\overline{v}),\theta=\theta_0+2t_N\mathbb{1}_p/\rho_N,N=1,2,\ldots,\overline{v}=\overline{v}_0\}$ can be assumed to be in $\mathcal{B}$ since $\theta_0\in\Theta^o$ 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 C1, and hence uniformly continuous on the product of $\mathcal{Z}$ and any compact subset of $\mathcal{B}$ containing $\{(\theta,\overline{v}),\theta=\theta_0+2t_N\mathbb{1}_p/\rho_N,N=1,2,\ldots,\overline{v}=\overline{v}_0\}$.

¿From Lemma 2-(i,iii) it follows that $g_1(\cdot|z,I)$ is a conditional density with support $[\underline{b}_0,\overline{b}_0(z,I)]$ for N large enough. It is important to verify that such a density corresponds to a structure $[U(\cdot;\theta_1),F_1]$ in our semiparametric model. This is established in the next Lemma. For $j=0,1$, define

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

Lemma C2: *Let (U_0, F_0) satisfy Assumption 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 $(z, I) \in \mathcal{Z} \times \mathcal{I}$, $G_1(\cdot|z, I)$ is rationalized by the IPV auction structure with risk aversion $[U(\cdot; \theta_1), F_1(\cdot|z, I)]$, where $F_1(\cdot|\cdot, I) \in \mathcal{F}_R^*$ with support $[\underline{v}_0, \overline{v}_0]$,
(ii) The conditional distribution function $F_1(\cdot|\cdot, \cdot)$ is such that $(\beta_1, f_1) \in \mathcal{V}_\epsilon(\beta_0, f_0)$.

We now turn to the statistical part of the proof. Because $\|\rho_N(\beta_1 - \beta_0)\|_\infty = 2t_N$, using the triangular inequality we have for any $\tilde{\beta}$

$$\begin{aligned}
\Pr_{\beta_1,f_1}\Big(\|\rho_N(\tilde{\beta} - \beta_1)\|_\infty \geq t_N\Big) &\geq \Pr_{\beta_1,f_1}\Big(\|\rho_N(\beta_1 - \beta_0)\|_\infty - \|\rho_N(\tilde{\beta} - \beta_0\|_\infty \geq t_N\Big) \\
&\geq \Pr_{\beta_1,f_1}\Big(\|\rho_N(\tilde{\beta} - \beta_0)\|_\infty < t_N\Big).
\end{aligned}$$

Thus, because (β_0, f_0) and (β_1, f_1) are in $\mathcal{V}_\epsilon(f_0, \beta_0)$ for L large enough, we obtain

$$\begin{aligned}
&\sup_{(\beta,f)\in\mathcal{V}_\epsilon(\beta_0,f_0)} \Pr_{\beta,f}\Big(\|\rho_N(\tilde{\beta} - \beta)\|_\infty \geq t_N\Big) \\
&\geq \frac{1}{2}\Big[\Pr_{\beta_0,f_0}\Big(\|\rho_N(\tilde{\beta}-\beta_0)\|_\infty \geq t_N\Big) + \Pr_{\beta_1,f_1}\Big(\|\rho_N(\tilde{\beta}-\beta_1)\|_\infty \geq t_N\Big)\Big] \\
&\geq \frac{1}{2}\mathrm{E}\Big[\Pr_{\beta_0,f_0}\Big(\|\rho_N(\tilde{\beta} - \beta_0)\|_\infty \geq t_N \mid \mathcal{F}_L\Big) + \Pr_{\beta_1,f_1}\Big(\|\rho_N(\tilde{\beta} - \beta_0)\|_\infty < t_N \mid \mathcal{F}_L\Big)\Big] \\
&\equiv \frac{1}{2}\mathrm{E}[\Pr_e(\mathcal{F}_L)],
\end{aligned} \tag{C.2}$$

where $\mathcal{F}_L$ is the σ-field generated by $\{(Z_\ell, I_\ell), 1 \leq \ell \leq L\}$.

Let $\Pr_{jN}$ be the probability of the $B_{i\ell}$ given $\mathcal{F}_L$ under $g_j(\cdot|\cdot,\cdot)$, for $j = 0, 1$. 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

$$\begin{aligned}
\Pr_e(\mathcal{F}_L) &= 1 - \Big(\Pr_{0N}(\|\rho_N(\tilde{\beta} - \beta_0)\|_\infty < t_N) - \Pr_{1N}(\|\rho_N(\tilde{\beta} - \beta_0)\|_\infty < t_N)\Big) \\
&\geq 1 - \sup_A |\Pr_{0N}(A) - \Pr_{1N}(A)| = 1 - \frac{1}{2}\int |d\Pr_{0N} - d\Pr_{1N}| \\
&\geq 1 - \left[\int \left(\sqrt{d\Pr_{0N}} - \sqrt{d\Pr_{1N}}\right)^2\right]^{1/2} = 1 - \sqrt{2}\left(1 - \int \sqrt{d\Pr_{0N} d\Pr_{1N}}\right)^{1/2}.
\end{aligned}$$

Therefore

$$\Pr_e(\mathcal{F}_L) \geq 1 - \sqrt{2}\left(1 - \prod_{\ell=1}^{L}\prod_{i=1}^{I_\ell}\int_{\underline{b}_0}^{\bar{b}_0(Z_\ell,I_\ell)} \sqrt{g_0(b_{i\ell}|Z_\ell, I_\ell)g_1(b_{i\ell}|Z_\ell, I_\ell)}db_{i\ell}\right)^{1/2}. \tag{C.3}$$

But, because $g_j(\cdot|\cdot,\cdot)$, $j = 0, 1$, 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 $g_1(\cdot|\cdot,\cdot)$ and a standard Taylor expansion

$$\int_{\underline{b}_0}^{\bar{b}_0(Z_\ell,I_\ell)} \sqrt{g_0(b|Z_\ell, I_\ell)g_1(b|Z_\ell, I_\ell)}db$$

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

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

$$
\begin{aligned}
&\prod_{\ell=1}^{L}\prod_{i=1}^{I_\ell}\int_{\underline{b}_0}^{\bar{b}_0(Z_\ell,I_\ell)}\sqrt{g_0(b_{i\ell}|Z_\ell,I_\ell)g_1(b_{i\ell}|Z_\ell,I_\ell)}db_{i\ell} \\
&= \left[1+O\left(\rho_N^{-\frac{2R+3}{R+1}}\right)\right]^N=\exp\left[N\log\left(1+O\left(\rho_N^{-(2R+3)/(R+1)}\right)\right)\right] \\
&= \exp\left[NO\left(\rho_N^{-(2R+3)/(R+1)}\right)\right]=1+O\left(N\rho_N^{-(2R+3)/(R+1)}\right)=1+o(1).
\end{aligned}
$$

Hence, (C.2) and (C.3) yield

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

Appendix D

Appendix D establishes Theorem 3 stated in Section 6. Proofs of Lemmas used to prove Theorem 3 are given in Appendix E. Throughout, let $\mathcal{F}_L$ be the σ-field generated by $\{(Z_\ell,I_\ell),1\leq\ell\leq L\}$, let $a\asymp b$ mean that $a/b\to c$ with $0<c<\infty$, and for $u=(u_{i\ell})\in I\!R^N$, set

$$
\|u\|_p=\left(\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}|u_{i\ell}|^p\right)^{1/p},\ \|u\|_\infty=\max_{1\leq\ell\leq L}\max_{1\leq i\leq I_\ell}|u_{i\ell}|.
$$

We first state a series of lemmas. The first lemma studies the bias and error terms of (10).

Lemma D1: *Let Assumptions A2, A3 and A4-(iii,iv) hold.*

(i) The variables $Y_{i\ell}$ (or $\epsilon_{i\ell}$), $1 \le i \le I_\ell$, $1 \le \ell \le L$ 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)

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

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

Lemma D2: *Let Assumptions 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*

$$\frac{\sum_{\ell=1}^L \sum_{i=1}^{I_\ell} u_{i\ell}\epsilon_{i\ell}}{\mathrm{Var}^{1/2}[\sum_{\ell=1}^L \sum_{i=1}^{I_\ell} u_{i\ell}\epsilon_{i\ell}|\mathcal{F}_L]} \xrightarrow{d} \mathcal{N}(0,1)$$

conditionally on $\mathcal{F}_L$ and thus unconditionally.

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

Lemma D3: *Let Assumptions A2, A3 and A4-(iii–v) hold. For any $u \in \mathbb{R}^N$ that is $\mathcal{F}_L$-measurable, we have*

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

Lemma D4: *Let Assumptions A2-(i,ii), A3-(i) and A4-(i,v) hold. We have*

$$\begin{aligned}
\sup_{\beta\in\mathcal{B}_\delta}\max_{1\le\ell\le L} |\hat{m}(Z_\ell, I_\ell;\beta) - m(Z_\ell, I_\ell;\beta)| &= O_P(a_N), \\
\sup_{\beta\in\mathcal{B}_\delta}\max_{1\le\ell\le L} \left\|\frac{\partial\hat{m}(Z_\ell, I_\ell;\beta)}{\partial\beta} - \frac{\partial m(Z_\ell, I_\ell;\beta)}{\partial\beta}\right\|_\infty &= O_P(a_N).
\end{aligned}$$

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 D5: *Let Assumptions A2, A3 and A4-(i,ii) hold. Define*

$$\overline{Q}(\beta) = \mathrm{E}\left[I\omega(Z,I)\left(m(Z,I;\beta) - m(Z,I;\beta_0)\right)^2\right] .$$

Then, for any $\epsilon > 0$, there exists $C_\epsilon > 0$ such that $\inf_{\beta\in\mathcal{B}_\delta;\|\beta-\beta_0\|_\infty\geq\epsilon}\overline{Q}(\beta) > C_\epsilon$. Moreover, the matrix $A(\beta)$ and $B(\beta)$ defined in (14) and (15) are of full rank in a neighborhood of β_0.

Lemma D6: *Let Assumptions A2, A3 and A4-(i,ii) hold. We have*

$$\sup_{\beta\in\mathcal{B}_\delta}\left|\frac{1}{L}\overline{Q}_N(\beta) - \overline{Q}(\beta)\right| = O_P\left(\frac{1}{\sqrt{L}}\right) = o_P(1).$$

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

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

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 D7: *Let Assumptions A2, A3 and A4-(i–iv) hold. If $\tilde{\beta}_N = \beta_0 + o_P(1)$, then we have $\sup_{\beta\in\mathcal{B}_\delta}|W_N(\beta)| = O_P(1)$ and $W_N^{(1)}(\tilde{\beta}_N) = o_P(1)$.*

Proof of Theorem 3: The proof is in 2 steps.
STEP 1: *Consistency.* Note that $|\max_{1\leq\ell\leq L}\hat{\bar{b}}_N(Z_\ell,\ell) - \sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\bar{b}_0(z,I)| < \delta/4$ with probability approaching one by Assumptions 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 and the bounds on the $g(z,I)$s.

Thus, $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\overline{b}_0(z,I)+\delta/4<\max_{1\leq\ell\leq L}\hat{\overline{b}}_N(Z_\ell,\ell)+\delta/2<\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}}\overline{b}_0(z,I)+3\delta/4<\overline{v}_0<\overline{v}_{\sup}$ with high probability, using Assumption A4-(i). That is, $\overline{v}_0\in\mathcal{B}_N\subset\mathcal{B}_{\delta/4}$ with high probability for N large enough.

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 D3 and D4 together with Assumption A4-(ii) yield

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

since $a_N=o(\sqrt{a_N}/h_N)$ by Assumption 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

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

uniformly in $\beta\in\mathcal{B}_{\delta/4}$. The second equality follows from Assumption A4-(ii,iv), Lemma C1 and Lemma D1, the latter implying that $\sum_\ell\sum_i\epsilon_{i\ell}^2=O_P(N/h_N)$ by Markov inequality. Thus, combining (D.1) and (D.2) gives

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

since $a_N=o(h_N^3)$ by Assumption 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). Hence, Lemmas D1 and D7 together with Assumption A4-(iv) yield

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

Thus, using (D.3), (D.4), Lemma D6 and $L \asymp N$ gives

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

Combining this with Lemma D5 and recalling that $\overline{v}_0 \in \mathcal{B}_N \subset \mathcal{B}_{\delta/4}$ with high probability 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 2: *Asymptotic Normality.* Given Assumption A4-(i), we have $\sup_{(z,I)\in\mathcal{Z}\times\mathcal{I}} \overline{b}_0(z, I) + \delta < \overline{v}_0 < \overline{v}_{\sup}$. Thus, β_0 is an inner point of $\mathcal{B}_N$ with high probability. Hence, because it is a consistent estimator of β_0, with probability tending to 1, $\hat{\beta}_N$ solves the first-order condition:

$$\begin{aligned}0 &= \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell, I_\ell)\left(\hat{Y}_{i\ell} - \hat{m}(Z_\ell, I_\ell;\hat{\beta}_N)\right)\frac{\partial\hat{m}(Z_\ell, I_\ell;\hat{\beta}_N)}{\partial\beta} \\ &= \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell, I_\ell)\left(\hat{Y}_{i\ell} - \hat{m}(Z_\ell, I_\ell;\beta_0)\right)\frac{\partial\hat{m}(Z_\ell, I_\ell;\hat{\beta}_N)}{\partial\beta} \\ &\quad - \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell, I_\ell)\frac{\partial\hat{m}(Z_\ell, I_\ell;\hat{\beta}_N)}{\partial\beta}\int_0^1\frac{\partial\hat{m}(Z_\ell, I_\ell;\beta_0 + t(\hat{\beta}_N - \beta_0))}{\partial\beta'}dt\ (\hat{\beta}_N - \beta_0).\end{aligned}$$

Therefore,

$$\begin{aligned}\hat{\beta}_N - \beta_0 &= \left[\sum_{\ell=1}^{L} I_\ell\omega(Z_\ell, I_\ell)\frac{\partial\hat{m}(Z_\ell, I_\ell;\hat{\beta}_N)}{\partial\beta}\int_0^1\frac{\partial\hat{m}(Z_\ell, I_\ell;\beta_0 + t(\hat{\beta}_N - \beta_0))}{\partial\beta'}dt\right]^{-1} \\ &\quad \times\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell, I_\ell)\left(\hat{Y}_{i\ell} - \hat{m}(Z_\ell, I_\ell;\beta_0)\right)\frac{\partial\hat{m}(Z_\ell, I_\ell;\hat{\beta}_N)}{\partial\beta}. \quad \text{(D.5)}\end{aligned}$$

Let $\hat{J}_N$ be the term within brackets. Lemmas C1, D4, D6 and the consistency of $\hat{\beta}_N$ yield

$$\begin{aligned} \hat{J}_N &= \sum_{\ell=1}^{L} I_\ell \omega(Z_\ell, I_\ell) \frac{\partial m(Z_\ell, I_\ell; \hat{\beta}_N)}{\partial \beta} \int_0^1 \frac{\partial m(Z_\ell, I_\ell; \beta_0 + t(\hat{\beta}_N - \beta_0))}{\partial \beta'} dt + O_P(N a_N) \\ &= A_N(\beta_0) + o_P(N) = N A(\beta_0) + o_P(N), \end{aligned} \tag{D.6}$$

where $A(\beta_0)$ has an inverse by Lemma D5.

Next, we study the second term in (D.5), i.e. the score $\hat{S}_N$. We have

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

¿From (10) we have $\sum_\ell \sum_i |Y_{i\ell}| = O_P(N/\sqrt{h_N})$ by Lemmas C1 and D1, using Markov and Cauchy-Schwartz inequalities to get $\sum_\ell \sum_i |\epsilon_{i\ell}| = O_P(N/\sqrt{h_N})$, which is the leading term given Assumption A4-(iv). Therefore, using Lemmas C1, D3 and D4 together with Assumption A4-(ii), we obtain

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

Using (10), the consistency of $\hat{\beta}_N$, Lemmas C1, D1 and D7 with Assumption A2-(ii) imply that the second term is an $o_P\left(N h_N^{R+1}\right) + o_P\left(\sqrt{N/h_N}\right)$. Note that $N a_N / h_N = o\left(N h_N^{R+1}\right)$ and $N a_N / h_N^2 = o(N/h_N)$ under Assumption A4-(v). Hence, (10) and Lemmas D1 and D6 imply

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

$$= h_N^{R+1}\mathbf{b}_N(\beta_0, g_0) + \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell, I_\ell)\epsilon_{i\ell}\frac{\partial m(Z_\ell, I_\ell;\beta_0)}{\partial\beta} + o_P\left(Nh_N^{R+1} + \sqrt{\frac{N}{h_N}}\right)$$
$$= Nh_N^{R+1}\mathbf{b}(\beta_0, g_0) + \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell, I_\ell)\epsilon_{i\ell}\frac{\partial m(Z_\ell, I_\ell;\beta_0)}{\partial\beta} + o_P\left(Nh_N^{R+1} + \sqrt{\frac{N}{h_N}}\right). \quad \text{(D.7)}$$

Let $u_{i\ell} = \omega(Z_\ell, I_\ell)\partial m(Z_\ell, I_\ell;\beta_0)/\partial\beta$. Lemmas C1, D1, D6 and assumption A4-(ii) imply

$$\begin{aligned}\operatorname{Var}\left[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}u_{i\ell}\epsilon_{i\ell} \mid \mathcal{F}_L\right] &= \sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega^2(Z_\ell, I_\ell)\frac{m(Z_\ell, I_\ell;\beta_0)}{h_N}\frac{\partial m(Z_\ell, I_\ell;\beta_0)}{\partial\beta}\frac{\partial m(Z_\ell, I_\ell;\beta_0)}{\partial\beta'}\int\Phi^2(x)dx \\ &\quad + o_P(N/h_N) \\ &= \frac{1}{h_N}B_N(\beta_0)\int\Phi^2(x)dx + o_P(N/h_N) \\ &= \frac{N}{h_N}\left[B(\beta_0)\int\Phi^2(x)dx + o_p(1)\right].\end{aligned}$$

Because $\|u\|_\infty/(\|u\|_2\sqrt{h_N}) = O_P(1/\sqrt{Nh_N}) = o_P(1)$ by Assumption A4-(iv), Lemma D2 implies

$$\sqrt{\frac{h_N}{N}}\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\omega(Z_\ell, I_\ell)\epsilon_{i\ell}\frac{\partial m(Z_\ell, I_\ell;\beta_0)}{\partial\beta} \xrightarrow{d} \mathcal{N}\left(0, B(\beta_0)\int\Phi^2(x)dx\right). \quad \text{(D.8)}$$

Collecting (D6)–(D.8) and using $\hat{\beta}_N - \beta_0 = \hat{J}_N^{-1}\hat{S}_N$ from (D.5) give

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

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 (D.6).

Appendix E

Appendix E gathers proofs of Lemmas stated in Appendices A, C, and D.

E.1 Proofs of Lemmas A1–A6

Proof of Lemma A1: Fix $(z, I) \in \mathcal{Z} \times \mathcal{I}$, and omit (z, I) to simplify the notation.

First, we show that property (i) holds. Recall from Maskin and Riley (1984) that $s(\cdot)$ is continuous and strictly increasing on $[\underline{v}, \overline{v}]$ with $s(\underline{v}) = \underline{v}$. Consider the optimization problem

(1) of an arbitrary bidder with a private value $v \in (\underline{v}, \overline{v}]$. If $b \geq v$, we have $\mathrm{E}\Pi \leq 0$ since $U(v-b) \leq 0$.[29] On the other hand, if $b < v$, we have $U(v-b) > 0$ because $U(\cdot) > 0$ on $(0, +\infty)$ by Definition 1. Moreover, if $b \in (\underline{v}, v)$, which implies $b > \underline{v} = s(\underline{v})$, we have $\Pr(\text{bidder wins}) = \mathbb{1}[b \leq s(\overline{v})]F^{I-1}[s^{-1}(b)] + \mathbb{1}[b > s(\overline{v})]$, which is strictly positive because $F[s^{-1}(b)] > F[\underline{v}] = 0$. Hence, $\mathrm{E}\Pi > 0$ if $b \in (\underline{v}, v)$. This shows that the optimal bid must satisfy $b < v$ whenever $v \in (\underline{v}, \overline{v}]$.

Next, we show that $s(\cdot)$ is continuously differentiable on $[\underline{v}, \overline{v}]$ and that $s'(\cdot)$ satisfies property (ii). Because $s(\cdot)$ solves (A.1) on $(\underline{v}, \overline{v}]$, property (i) implies that $s'(\cdot)$ is continuous and strictly positive on $(\underline{v}, \overline{v}]$. Thus, it remains to prove that $s(\cdot)$ is continuously differentiable at $\underline{v}$ with $0 < s'(\underline{v}) < 1$. Let

$$\Psi(v) = (I-1)\frac{f(v)(v-\underline{v})}{F(v)}\frac{\lambda(v-s(v))}{v-s(v)} \text{ if } v \in (\underline{v}, \overline{v}], \qquad \Psi(v) = (I-1)\lambda'(0) \text{ if } v = \underline{v}.$$

Note that $\Psi(\cdot)$ is continuous on $[\underline{v}, \overline{v}]$. For $v_0 \in [\underline{v}, \overline{v}]$, define

$$\gamma(v_0) = \inf_{v \in [\underline{v}, v_0]} \Psi(v) > 0, \qquad \Gamma(v_0) = \sup_{v \in [\underline{v}, v_0]} \Psi(v) < \infty,$$

where the first inequality follows from $f(\cdot|z, I) > 0$, property (i) and $\lambda'(0) > 0$. Moreover, $\gamma(\cdot)$ and $\Gamma(\cdot)$ are continuous on $[\underline{v}, \overline{v}]$ with $\gamma(\underline{v}) = \Gamma(\underline{v}) = (I-1)\lambda'(0) > 0$ and $\Gamma(\underline{v})/(1+\gamma(\underline{v})) < 1$. Hence, there exists a $v_0 \in (\underline{v}, \overline{v}]$ with $0 < \Gamma_0/(1+\gamma_0) < 1$, where $0 < \gamma_0 \equiv \gamma(v_0) \leq \Gamma(v_0) \equiv \Gamma_0$.

Thus, from (A.1) and property (i), we have for any $v \in (\underline{v}, v_0]$

$$\gamma_0 \frac{v-s(v)}{v-\underline{v}} \leq s'(v) = \Psi(v)\frac{v-s(v)}{v-\underline{v}} \leq \Gamma_0 \frac{v-s(v)}{v-\underline{v}}.$$

We use these inequalities to find upper and lower bounds on $s'(v)$. Let $\overline{\kappa}(v) = (v-\underline{v})^{\Gamma_0}[s(v)-\underline{v}]$ for $v \in [\underline{v}, v_0]$. Differentiating and using the above inequalities, we obtain for $v \in (\underline{v}, v_0]$

$$\gamma_0(v-\underline{v})^{\Gamma_0} \leq \overline{\kappa}'(v) = (v-\underline{v})^{\Gamma_0-1}[\Gamma_0(s(v)-\underline{v}) + (v-\underline{v})s'(v)] \leq \Gamma_0(v-\underline{v})^{\Gamma_0},$$

since $\gamma_0 \leq \Gamma_0$. In particular, this shows that $\lim_{v\downarrow\underline{v}} \overline{\kappa}'(v) = 0$, which implies that $\overline{\kappa}'(\underline{v})$ exists and is equal to 0 using the Mean Value Theorem. Hence, $\gamma_0(v-\underline{v})^{\Gamma_0} \leq \overline{\kappa}'(v) \leq \Gamma_0(v-\underline{v})^{\Gamma_0}$ for any $v \in [\underline{v}, v_0]$, which implies that $\overline{\kappa}'(\cdot)$ is integrable on $[\underline{v}, v_0]$. Moreover, because $\overline{\kappa}(\underline{v}) = 0$, we obtain $\overline{\kappa}(v) = \int_{\underline{v}}^{v} \overline{\kappa}'(x)dx \leq \Gamma_0(v-\underline{v})^{\Gamma_0+1}/(\Gamma_0+1)$ for $v \in [\underline{v}, v_0]$. Similarly, letting $\underline{\kappa}(v) =$

[29] Though Definition 1 requires $U(\cdot)$ to be defined on $\mathbb{R}_+$ only, we could have defined it on $\mathbb{R}$, as $U(x) = U_{vNM}(w+x) - U_{vNM}(w)$ for $x \in \mathbb{R}$. In this case $U(x) < 0$ whenever $x < 0$ because $U_{vNM}(\cdot)$ and hence $U(\cdot)$ are strictly increasing, while $U(0) = 0$.

$(v-\underline{v})^{\gamma_0}[s(v)-\underline{v}]$ for $v \in [\underline{v}, v_0]$, we obtain $\underline{\kappa}(v) \geq \gamma_0(v-\underline{v})^{\gamma_0+1}/(\gamma_0+1)$ for $v \in [\underline{v}, v_0]$. Thus, using the definitions of $\overline{\kappa}(\cdot)$ and $\underline{\kappa}(\cdot)$, we obtain

$$\frac{\gamma_0}{\gamma_0+1}[v-\underline{v}] \leq s(v)-\underline{v} \leq \frac{\Gamma_0}{\Gamma_0+1}[v-\underline{v}] \text{ for any } v \in [\underline{v}, v_0].$$

Therefore, the inequalities on $s'(v)$ yield

$$0 < \frac{\gamma_0}{\Gamma_0+1} \leq s'(v) \leq \frac{\Gamma_0}{\gamma_0+1} < 1 \text{ for any } v \in (\underline{v}, v_0].$$

In particular, these inequalities hold at v_0. Using $\gamma_0 \equiv \gamma(v_0)$ and $\Gamma_0 \equiv \Gamma(v_0)$ converge to $\gamma(\underline{v}) = \Gamma(\underline{v}) = (I-1)\lambda'(0) > 0$ as $v_0 \downarrow \underline{v}$, we obtain that $\lim_{v_0 \downarrow \underline{v}} s'(v_0) = (I-1)\lambda'(0)/[(I-1)\lambda'(0)+1]$, which is strictly between 0 and 1. Application of the Mean Value Theorem between $\underline{v}$ and v establishes that $s'(\underline{v})$ exists and is equal to $(I-1)\lambda'(0)/[(I-1)\lambda'(0)+1]$. Moreover, $s'(\cdot)$ is continuous at $\underline{v}$, which implies that (A.1) also holds at $\underline{v}$ by taking the limit as $v \downarrow \underline{v}$ because $\Psi(\cdot)$ is continuous at $\underline{v}$.

Proof of Lemma A2: Without loss of generality, assume that $\underline{v} = 0$. For (i), let $x > 0$ and $v > 0$. The Liebnitz-Newton formula yields

$$\begin{aligned}\frac{\partial^{r_2}\Lambda(x;v)}{\partial v^{r_2}} &= \frac{\partial^{r_2}}{\partial v^{r_2}}\left(\frac{\lambda(vx)}{v}\right) = \sum_{j=0}^{r_2} \frac{r_2!}{j!(r_2-j)!}\frac{\partial^j \lambda(vx)}{\partial v^j}\frac{\partial^{r_2-j}}{\partial v^{r_2-j}}\left(\frac{1}{v}\right) \\ &= \frac{(-1)^{r_2} r_2!}{v^{r_2+1}}\sum_{j=0}^{r_2}\frac{\lambda^{(j)}(vx)}{j!}(-vx)^j.\end{aligned}$$

On the other hand, a Taylor expansion of $\lambda(0) = \lambda(vx - vx) = 0$ around vx with integral remainder (see e.g Zeidler (1985, p.77)) shows that

$$0 = \sum_{j=0}^{r_2}\frac{\lambda^{(j)}(vx)}{j!}(-vx)^j + \frac{1}{r_2!}\int_0^1 (1-t)^{r_2}(-vx)^{r_2+1}\lambda^{(r_2+1)}(vx - tvx)\,dt.$$

Hence, after a change of variables, we obtain

$$\frac{1}{x}\frac{\partial^{r_2}\Lambda(x;v)}{\partial v^{r_2}} = x^{r_2}\int_0^1 t^{r_2}\lambda^{(r_2+1)}(tvx)\,dt\,,\quad \frac{\partial^{r_1+r_2}\Lambda(x;v)}{\partial x^{r_1}\partial v^{r_2}} = \frac{\partial^{r_1}}{\partial x^{r_1}}\left(x^{r_2+1}\int_0^1 t^{r_2}\lambda^{(r_2+1)}(tvx)\,dt\right).$$

The Lebesgue Dominated Convergence Theorem shows that these functions are continuous on $I\!R_+ \times [\underline{v}, \overline{v}]$ because $r_1 + r_2 \leq R$ and $\lambda(\cdot)$ is $R+1$ continuously differentiable on $I\!R_+$.

For (ii), using the definition of $\Phi_I(\cdot;\cdot,\cdot)$, and note that the partial derivatives up to order R of $(v,z) \to f(vu|z,I)$ are continuous with respect to (u,v,z). Note also that $vu/F(vu|z,I)$ is bounded away from infinity uniformly in (u,v,z) as $\min_{v,u,z} f(vu|z,I) > 0$. The rules of differentiation then shows that (ii) is proved if the partial derivatives up to order R with respect to (v,z) of $F(vu|z,I)/(vu)$ are continuous with respect to (u,v,z). Since $\partial^{r_3}F(0|z,I)/\partial z^{r_3}=0$ and using the same argument as for $\Lambda(\cdot;\cdot)$ replaced by $(1/v)\partial^{r_3}F(vu|z,I)/\partial z^{r_3}$ with u playing the role of x, we have

$$\frac{\partial^{r_2+r_3}}{\partial v^{r_2}\partial z^{r_3}}\left(\frac{F(vu|z,I)}{vu}\right)=\frac{\partial^{r_2}}{\partial v^{r_2}}\left(\frac{1}{vu}\frac{\partial^{r_3}F(vu|z,I)}{\partial z^{r_3}}\right)=u^{r_2}\int_0^1 t^{r_2}\frac{\partial^{r_2+r_3+1}F(tuv|z,I)}{\partial y^{r_2+1}\partial z^{r_3}}dt,$$

where ∂y indicates differentiation with respect to the first argument of $F(\cdot|z,I)$. The Lebesgue Dominated Convergence Theorem shows that (ii) is proved when $[U,F]\in\mathcal{U}_R\times\mathcal{F}_R^*$ since $r_2+r_3+1\leq R+1$.

Proof of Lemma A3: Note that (i) has been shown after (A.8).

(ii) When $r_1=0$, the desired property is easy to prove using Lemma A2. For $r_1\geq 1$, we first show that $\mathbf{E}_I^{r_1r_2r_3}(\sigma;v,z)$ is a continuous r_1-multilinear operator from $(C_1^0)^{r_1}$ to C_0 when $(\sigma,v,z)\in\Sigma_1^0\times[\underline{v},\overline{v}]\times\mathcal{Z}$. As noted above, this operator maps $(C_1^0)^{r_1}$ into a subset of C_0. The norm of r_1-multilinear operators $\mathbf{A}(\eta_1,\ldots,\eta_{r_1})$ from $(C_1^0)^{r_1}$ to C_0 is

$$\rho_{r_1}(\mathbf{A})=\sup_{\|\eta_1\|_{1,\infty}=\cdots=\|\eta_{r_1}\|_{1,\infty}=1}\|\mathbf{A}(\eta_1,\ldots,\eta_{r_1})\|_\infty,$$

and $\mathbf{A}$ is continuous if $\rho_{r_1}(\mathbf{A})<\infty$ (see Zeidler (1985, p.773)). Consider $\mathbf{E}_I^1(\sigma;v,z)$. By definition of $\|\cdot\|_\infty$ and $\|\cdot\|_{1,\infty}$, we have from (A.9)

$$\begin{aligned}\left\|\mathbf{E}_I^1(\sigma;v,z)(\eta_1)\right\|_\infty &\leq \left\|\eta_1^{(1)}\right\|_\infty+\sup_{u,v,z}|\Phi_I(u;v,z)|\sup_{(x,v)\in[0,1]\times[\underline{v},\overline{v}]}\left|\Lambda^{(1)}(x;v)\right|\left\|\frac{\eta_1(u)}{u}\right\|_\infty\\ &\leq \left[1+\sup_{u,v,z}|\Phi_I(u;v,z)|\sup_{(x,v)\in[0,1]\times[\underline{v},\overline{v}]}\left|\Lambda^{(1)}(x;v)\right|\right]\|\eta_1\|_{1,\infty},\end{aligned}$$

since $\sigma\in\Sigma_1^0$ implies $0\leq u-\sigma(u)\leq u\leq 1$ for all $u\in[0,1]$, and since $\eta_1(0)=0$ implies $\|\eta_1(u)/u\|_\infty=\|u^{-1}\int_0^u\eta_1^{(1)}(y)dy\|_\infty\leq\|\eta_1\|_{1,\infty}$. Now, the compactness of $[0,1]$, $[\underline{v},\overline{v}]$, $\mathcal{Z}$ and Lemma A2 give $\rho_1(\mathbf{E}_I^1(\sigma;v,z))<\infty$. The other operators $\mathbf{E}_I^{r_1r_2r_3}(\sigma;v,z)$ are treated similarly.

We now show the uniform continuity of $(\sigma,v,z)\mapsto\mathbf{E}_I^{r_1r_2r_3}(\sigma;v,z)$ relative to $\rho_{r_1}(\cdot)$, $r_1\geq 1$. We consider $(r_1,r_2,r_3)=(1,0,0)$, the other partial derivatives being treated similarly using Lemma A2. For any (σ_0,v_0,z_0), (σ,v,z) in $\Sigma_1^0\times[\underline{v},\overline{v}]\times\mathcal{Z}$, we have as above

$$\begin{aligned}&\rho_1\left(\mathbf{E}_I^1(\sigma;v,z)-\mathbf{E}_I^1(\sigma_0;v_0,z_0)\right)\\ &\quad\leq\sup_{u\in[0,1]}\left|\Phi_I(u;v,z)\Lambda^{(1)}(u-\sigma(u);v)-\Phi_I(u;v_0,z_0)\Lambda^{(1)}(u-\sigma_0(u);v_0)\right|,\end{aligned}$$

which can be made arbitrarily small independently of (σ_0, v_0, z_0) by taking $\|\sigma - \sigma_0\|_\infty \leq \|\sigma - \sigma_0\|_{1,\infty}$, $|v - v_0|$ and $\|z - z_0\|$ small by Lemma A2, because σ and σ_0 take values in the compact set $[0,1]$, and (v,z), (v_0, z_0) are also in compact sets.

To find the Fréchet derivatives of $\mathbf{E}_I(\sigma; v, z)$, we note that the above operators $\mathbf{E}_I^{r_1 r_2 r_3}(\sigma; v, z)$ are actually the Gâteaux derivatives of $\mathbf{E}_I(\sigma; v, z)$ over $\Sigma_1^0 \times [\underline{v}, \overline{v}] \times \mathcal{Z}$. It then follows from the continuity of these Gâteaux partial derivatives on $\Sigma_1^0 \times [\underline{v}, \overline{v}] \times \mathcal{Z}$ that the $\mathbf{E}_I^{r_1 r_2 r_3}(\sigma; v, z)$s are also the Fréchet partial derivatives of $\mathbf{E}_I(\sigma; v, z)$ by Proposition 4.8 in Zeidler (1985).

(iii) Fix $(\sigma, v, z, I) \in \Sigma_1^0 \times [\underline{v}, \overline{v}] \times \mathcal{Z} \times \mathcal{I}$. To show that $\mathbf{E}_I^1(\sigma; v, z)$ is a one-to-one operator from C_1^0 to C_0, consider $\zeta \in C_0$. ¿From (A.9), solving $\mathbf{E}_I^1(\sigma; v, z)(\eta) = \zeta$ for $\eta \in C_1^0$ leads to solving the following first-order linear differential equation

$$E_I^1(v,z) \quad : \quad \eta^{(1)}(u) + \left[\Phi_I(u; v, z) \frac{\Lambda^{(1)}(u - \sigma(u); v)}{u}\right] \eta(u) = \zeta(u) \text{ with } \eta(0) = 0.$$

Because the term in brackets diverges as $u \downarrow 0$, standard approaches to solve $E_I^1(v,z)$ on $[0,1]$ do not apply. In particular, solving this equation with $\zeta = 0$ would lead to consider the integral

$$\int_u^1 \Phi_I(y; v, z) \frac{\Lambda^{(1)}(y - \sigma(y); v)}{y} dy = \int_u^1 \Phi_I(y; v, z) \frac{\lambda'((v - \underline{v})(y - \sigma(y)))}{y} dy \to +\infty \text{ when } u \downarrow 0,$$

since $\Phi_I(0; v, z) = i - 1$ and $\lambda'(\cdot) \geq 1$. Nevertheless, it can be readily shown that the *unique* solution η_0 of $E_I^1(v,z)$ on $(0,1]$ with $\lim_{u \to 0} \eta_0(u) = 0$ is

$$\eta_0(u) = \int_0^u \zeta(y) \frac{R_\sigma(y)}{R_\sigma(u)} dy, \quad \text{where} \quad R_\sigma(u) = \exp\left[-\int_u^1 \Phi_I(y; v, z) \frac{\Lambda^{(1)}(y - \sigma(y); v)}{y} dy\right] \tag{E.1}$$

is defined on $[0,1]$ with $R_\sigma(0) = 0$. Moreover, $0 \leq R_\sigma(y) \leq R_\sigma(u)$ for $0 \leq y \leq u$ implies $|\eta_0(u)| \leq \|\zeta\|_\infty u$, so that $\eta_0(\cdot)$ is defined on $[0,1]$ with $\eta_0(0) = 0$.

The proof of (iii) is complete by showing that $\eta_0 \in C_1^0$ and that η_0 solves $E_I^1(v,z)$ at 0. This is done via the following two steps: a) η_0 is differentiable at 0; b) $\eta_0^{(1)}$ is continuous at 0 and satisfies $E_I^1(v,z)$ at $u = 0$. For the sake of simplification, assume that $\underline{v} = 0$ and $I = 2$.

(iii.a) η_0 *is differentiable at* 0. For $0 \leq y \leq u$ and $u \to 0$, consider the expansion of

$$\frac{R_\sigma(y)}{R_\sigma(u)} = \exp\left(-\int_y^u \Phi_2(x; v, z) \frac{\Lambda^{(1)}(x - \sigma(x); v)}{x} dx\right).$$

¿From the definition of $\Phi_2(\cdot; \cdot, \cdot)$ and $\Lambda(\cdot; v)$ combined with $\inf_z f(0|z, 2) > 0$ under Definition 4, it is easily seen that, when $x \downarrow 0$,

$$\begin{aligned}
\frac{\Phi_2(x; v, z)}{x} &= \frac{v f(vx|z,2)}{F(vx|z,2)} = \frac{1}{x}\left(1 + \frac{vx}{2} \frac{f'(0|z,2)}{f(0|z,2)} + o(x)\right), \\
\Lambda^{(1)}(x - \sigma(x); v) &= \lambda'(v(x - \sigma(x))) = \lambda'(0) + v\lambda''(0)\left(1 - \sigma^{(1)}(0)\right) x + o(x),
\end{aligned}$$

so that

$$\Phi_2(x;v,z)\frac{\Lambda^{(1)}\left(x-\sigma(x);v\right)}{x}=\frac{\lambda'(0)}{x}+O(1) \quad \text{when} \quad x\downarrow 0. \tag{E.2}$$

Consequently, since $0\leq y\leq u$, we have, when $u\to 0$,

$$\frac{R_\sigma(y)}{R_\sigma(u)}=\exp\left[-\int_y^u\left(\frac{\lambda'(0)}{x}+O(1)\right)dx\right]=\exp\left(-\lambda'(0)\ln\frac{u}{y}+O(u-y)\right)=\left(\frac{y}{u}\right)^{\lambda'(0)}(1+o(1)).$$

We now expand $\eta_0(u)$ when u is close to 0. The above expansion gives

$$\eta_0(u)=\int_0^u(\zeta(0)+o(1))\frac{R_\sigma(y)}{R_\sigma(v)}dy=\frac{\zeta(0)+o(1)}{u^{\lambda'(0)}}\int_0^u y^{\lambda'(0)}(1+o(1))dy=\frac{\zeta(0)}{\lambda'(0)+1}u+o(u), \tag{E.3}$$

where $\lambda'(0)/geq1$. Hence $\eta_0(\cdot)$ is differentiable at 0.

(iii.b) $\eta_0^{(1)}$ *is continuous at* 0 *and satisfies* $E_I^1(v,z)$ *at* $u=0$. We show that the value $\eta_0^{(1)}(0)=\zeta(0)/(\lambda'(0)+1)$ given by (E.3) is consistent with $E_2^1(v,z)$. The differential equation $E_2^1(v,z)$, (E.2) and (E.3) yield, for $u\downarrow 0$,

$$\begin{aligned}\eta_0^{(1)}(u) &= \zeta(u)-\frac{\Phi_2(u;v,z)}{u}\Lambda^{(1)}\left(u-\sigma(u);v\right)\eta_0(u)\\ &= \zeta(0)+o(1)-\left(\frac{\lambda'(0)}{u}+O(1)\right)\left(\frac{\zeta(0)}{\lambda'(0)+1}u+o(u)\right)=\frac{\zeta(0)}{\lambda'(0)+1}+o(1),\end{aligned}$$

as required by (E.3). Hence η_0 solves $E_2^1(v,z)$ and is C_1 on $[0,1]$.

Proof of Lemma A4: Fix $(z,I)\in\mathcal{Z}\times\mathcal{I}$. The differential equation (A.1) is not defined if $s(v;z,I)>v$ because $\lambda(\cdot)$ is defined on $\mathbb{R}_+$. We thus extend (A.1) to

$$\widetilde{E}(z,I)\ :\ s'(v;z,I)=(I-1)\frac{f(v|z,I)}{F(v|z,I)}\tilde{\lambda}\left(v-s(v;z,I)\right)\ ,\ v\in[\underline{v},\overline{v}]\ , \tag{E.4}$$

with $s(\underline{v};z,I)=\underline{v}$, where $\tilde{\lambda}(x)=\lambda(x)$ if $x\geq 0$, and $\tilde{\lambda}(x)=\lambda'(0)x$ if $x\leq 0$. Note that $\tilde{\lambda}(\cdot)$ is continuously differentiable on $\mathbb{R}$, and that $\tilde{\lambda}(\cdot)$ is consistent with our previous definition of $\Lambda(x;v)$, which can be equivalently defined as $\Lambda(x;v)=\tilde{\lambda}[(v-\underline{v})x]/(v-\underline{v})$ if $v>\underline{v}$, and $\Lambda(x;v)=\tilde{\lambda}'(0)x$ if $v=\underline{v}$, for $x\in\mathbb{R}$. Moreover, because $\tilde{\lambda}(v-s(v))=\lambda(v-s(v))$ as soon as $s(v)\leq v$, then $s(\cdot;z,I)$ is a continuously differentiable solution of (A.1) on $[\underline{v},\overline{v}]$ with initial condition $s(\underline{v};z,I)=\underline{v}$ if and only if it is a solution in $S_1(\underline{v})$ of (E.4). To prove Lemma A4, it thus suffices to show that (E.4) has a solution, which is unique in $S_1(\underline{v})$. This is done in three steps.

Step 1: *The continuously differentiable solutions of (E.4) with initial condition* $s(\underline{v};z,I)=\underline{v}$ *are in* $S_1(\underline{v})$. Let $s(\cdot;z,I)$ be such a solution, if it exists.

We first check that $s(\cdot;z,I)$ verifies the conditions of $S_1(\underline{v})$ in a neighborhood $[\underline{v},v_0]$ of $\underline{v}$. For $v>\underline{v}$ close to $\underline{v}$, we have

$$\begin{aligned}
0 &= F(\underline{v}|z,I) = F(v|z,I) + (f(v|z,I)+o(1))\,(\underline{v}-v),\\
0 &= \widetilde{\lambda}\,(\underline{v}-s(\underline{v};z,I)) = \widetilde{\lambda}\,(v-s(v;z,I)) + \Big[\big(1-s'(v;z,I)\big)\,\widetilde{\lambda}'\,(v-s(v;z,I))+o(1)\Big]\,(\underline{v}-v),
\end{aligned}$$

where the remainder terms $o(1)=o_v(1)$ are uniform in v since $f(v|z,I)$ and $s'(v;z,I)$ are continuous on $[\underline{v},\overline{v}]$, while $\widetilde{\lambda}(\cdot)$ is continuously differentiable on $\mathbb{R}$. Substituting in (E.4) and simplifying give

$$s'(v;z,I) = \frac{I-1}{1+o(1)}\Big[\big(1-s'(v;z,I)\big)\,\tilde{\lambda}'\,(v-s(v;z,I))+o(1)\Big],$$

so that

$$s'(\underline{v};z,I) = \lim_{v\downarrow\underline{v}} s'(v;z,I) = \frac{(I-1)\lambda'(0)}{(I-1)\lambda'(0)+1} \in (0,1).$$

Therefore, there is a $v_0\in(\underline{v},\overline{v}]$ such that $s(v;z,I)<v$ and $s'(v;z,I)>0$ for all $v\in(\underline{v},v_0]$.

We now check that $s(\cdot;z,I)$ verifies the conditions of $S_1(\underline{v})$ over $[v_0,\underline{v}]$. Observe that $s(v;z,I)<v$ should be true at least on an interval $[v_0,v^*)$, $v^*>v_0$. Let $v_1>\underline{v}$ be the largest possible $v^*\le\overline{v}$. Note that $s(v_1;z,I)=v_1$ and $s'(v_1;z,I)=0$ by (E.4). Because $R\ge 1$, taking left derivatives of (E.4) at $v=v_1$ gives

$$\begin{aligned}
\frac{s''(v_1;z,I)}{I-1} &= \frac{\partial}{\partial v}\left(\frac{f(v_1|z,I)}{F(v_1|z,I)}\right)\lambda\,(v_1-s(v_1;z,I)) + \frac{f(v_1|z,I)}{F(v_1|z,I)}\lambda'\,(v_1-s(v_1;z,I))\,\big(1-s'(v_1;z,I)\big)\\
&= \frac{f(v_1|z,I)}{F(v_1|z,I)}\lambda'(0)>0,
\end{aligned}$$

since $\lambda'(\cdot)\ge 1$ and $f(\cdot|z,I)>0$ on $[\underline{v},\overline{v}]$. Consequently, a second-order Taylor expansion for $\epsilon>0$ small enough yields $s(v_1-\epsilon;z,I)=v_1+\epsilon^2(s''(v_1;z,I)/2+o(1))$, which, since $s''(v_1;z,I)>0$, contradicts the condition $s(v_1-\epsilon;z,I)<v_1-\epsilon$ coming from the definition of v_1. Thus $s(v;z,I)<v$ for all $v\in[v_0,\overline{v}]$, which also gives $s'(v;z,I)>0$ for all $v\in[v_0,\overline{v}]$. This shows that all continuously differentiable solutions of (E.4) with $s(\underline{v};z,I)=\underline{v}$ are in $S_1(\underline{v})$. Equivalently, all the zeros in C_1^0 of the operator $\mathbf{E}_I(\cdot;v,z)$ defined in (A.8) are in $\Sigma_1^0\subset C_1^0$.

Step 2: *(E.4) has a unique continuously differentiable solution on a small interval* $[\underline{v},v_0]$, $v_0>\underline{v}$ *with initial condition* $s(\underline{v};z,i)=\underline{v}$. Equivalently, we want to show that the operator $\mathbf{E}_I(\cdot;v,z)$ given in (A.8) has a unique zero in C_1^0 for $v=v_0$. Existence and uniqueness of this solution follow using a continuation argument given in Proposition 6.10 in Zeidler (1985) as $\mathbf{E}_I(\cdot;\underline{v},z)$ has a unique zero in C_1^0, namely the solution given by (A.7).

For $v_0 \in [\underline{v}, \overline{v}]$, define

$$\begin{aligned}
\gamma(v_0) &= \inf_{v\in[\underline{v},v_0],u\in[0,1]} \Phi_I(u;v,z) \times \inf_{v\in[\underline{v},v_0],x\in[0,1]} \frac{\Lambda(x;v)}{x} > 0, \\
\Gamma(v_0) &= \sup_{v\in[\underline{v},v_0],u\in[0,1]} \Phi_I(u;v,z) \times \sup_{v\in[\underline{v},v_0],x\in[0,1]} \frac{\Lambda(x;v)}{x} < \infty,
\end{aligned}$$

where the inequalities follow from $f(\cdot|z,I) > 0$ and $\Lambda(0;v)/0 = \lambda'(0) > 0$. Moreover, Lemma A2 implies that $\Gamma(\cdot)$ and $\gamma(\cdot)$ are continuous. Thus, $\Gamma(\underline{v}) = \gamma(\underline{v}) = (I-1)\lambda'(0) > 0$ and $\Gamma(\underline{v})/(1+\gamma(\underline{v})) < 1$. Hence, there exists a $v_0 \in (\underline{v}, \overline{v}]$ with $0 < \Gamma_0/(1+\gamma_0) < 1$, where $0 < \gamma_0 \equiv \gamma(v_0) \leq \Gamma(v_0) \equiv \Gamma_0$.

Now, if $\sigma_I(\cdot;v,z)$ is a zero in C_1^0 of $\mathbf{E}_I(\cdot;v,z)$, then $\sigma_I(\cdot;v,z) \in \Sigma_1^0$ by Step 1. Thus, using (A.8) and the definitions of Γ_0, γ_0 and Σ_1^0, we have for any $(u,v) \in [0,1] \times [\underline{v}, v_0]$

$$\gamma_0 \frac{u-\sigma(u;v)}{u} \leq \sigma^{(1)}(u;v) = \Phi(u;v) \frac{\Lambda\left(u-\sigma(u;v);v\right)}{u-\sigma(u;v)} \frac{u-\sigma(u;v)}{u} \leq \Gamma_0 \frac{u-\sigma(u;v)}{u},$$

dropping the dependence upon (z,I) to simplify the notation. Thus, the same argument as in the proof of Lemma A1-(ii) with $\underline{v}=0$ and $[v,s(v)]$ replaced by $[u,\sigma(u,v)]$ gives

$$\frac{\gamma_0}{\gamma_0+1} u \leq \sigma(u;v) \leq \frac{\Gamma_0}{\Gamma_0+1} u \text{ for any } u \in [0,1] \text{ and every } v \in [\underline{v}, v_0],$$

$$0 < \frac{\gamma_0}{\Gamma_0+1} \leq \sigma^{(1)}(u;v) \leq \frac{\Gamma_0}{\gamma_0+1} < 1 \text{ for any } u \in [0,1] \text{ and every } v \in [\underline{v}, v_0].$$

This implies that there is an $\epsilon > 0$ such that

$$\{\zeta \in C_1^0 \,;\, \|\zeta(\cdot) - \sigma(\cdot;v)\|_{1,\infty} < \epsilon\} \subset \Sigma_1^0 \text{ for any } \sigma(\cdot;v) \text{ zero in } C_1^0 \text{ of } \mathbf{E}_I(\cdot;v,z),\ v \in [\underline{v}, v_0].$$

Because Σ_1^0 is open, it follows that the a priori condition (ii) of the continuation argument in Proposition 6.10 of Zeidler (1985) is satisfied.

We now check condition (iv) of Proposition 6.10 in Zeidler (1985), and in particular, that the norm of the (linear) operator $\left[\mathbf{E}_I^1(\sigma;v)\right]^{-1}$ is uniformly bounded for $(\sigma,v) \in \Sigma_1^0 \times [\underline{v}, v_0]$. For any $\zeta \in C_0$, (E.1) gives $\eta_0 = \left[\mathbf{E}_I^1(\sigma;v)\right]^{-1}(\zeta)$. Hence $\left|\mathbf{E}_I^1(\sigma;v)\right]^{-1}(\zeta)\|_{1,\infty} = \|\eta_0\|_{1,\infty} = \|\eta_0^{(1)}\|_\infty$, where $\eta_0^{(1)} + \Phi(u;v)\Lambda^{(1)}\left(u-\sigma(u);v\right)\eta_0(u)/u = \zeta(u)$ by (A.9). Hence

$$\begin{aligned}
\left\|\left[\mathbf{E}_I^1(\sigma;v)\right]^{-1}(\zeta)\right\|_{1,\infty} &\leq \|\zeta\|_\infty + \left\|\Phi(u;v)\Lambda^{(1)}\left(u-\sigma(u);v\right)\frac{1}{u}\int_0^u \zeta(y)\frac{R_\sigma(y)}{R_\sigma(u)}dy\right\|_\infty \\
&\leq \|\zeta\|_\infty + \left\|\Phi(u;v)\lambda'\left((v-\underline{v})(u-\sigma(u))\right)\right\|_\infty \|\zeta\|_\infty \\
&\leq \left[1 + \sup_{(u,v)\in[0,1]\times[\underline{v},v_0]} \Phi(u;v) \times \sup_{v\in[\underline{v},v_0]} \lambda'(v)\right] \|\zeta\|_\infty,
\end{aligned}$$

for any $(\sigma, v) \in \Sigma_1^0 \times [\underline{v}, v_0]$, because $R_\sigma(y) \leq R_\sigma(u)$ and $\Lambda^{(1)}(x; v) = \lambda'((v - \underline{v})x)$, recalling that $0 \leq u - \sigma(u) \leq 1$. Thus, the norm of the operator $\left[\mathbf{E}_I^1(\sigma; v)\right]^{-1}$ is uniformly bounded, as desired.

Because the other assumptions of Proposition 6.10 of Zeidler (1985) have been established in Lemma A3, the continuation argument yields that $\mathbf{E}_I(\cdot; v_0, z)$ has a unique zero $\sigma_I(\cdot; v_0, z)$ in C_1^0 since $\mathbf{E}_I(\cdot; \underline{v}, z)$ has a unique zero (A.7) on C_1^0. Moreover, $\sigma_I(\cdot; v_0, z) \in \Sigma_1^0$. Thus, (A.4) implies that

$$s_0(v; z, I) = \underline{v} + (v_0 - \underline{v})\sigma_I\left(\frac{v - \underline{v}}{v_0 - \underline{v}}; v_0, z\right), \quad v \in [\underline{v}, v_0]$$

is the unique continuously differentiable solution of (E.4) on $[\underline{v}, v_0]$ with $s_0(\underline{v}; z, I) = \underline{v}$. Moreover, $0 < s_0'(\underline{v}; z, I) < 1$, and $s_0(v; z, I) < v$, $s_0'(v; z, I) > 0$ for $v \in [\underline{v}, v_0]$.

STEP 3: *(E.4) has a unique continuously differentiable solution on $[\underline{v}, \overline{v}]$ with initial condition $s(\underline{v}; z, I) = \underline{v}$.* We extend $s_0(\cdot; z, I)$ on $[\underline{v}, \overline{v}]$. Let $\tilde{\lambda}^*(x) = \tilde{\lambda}(x)$ if $x \leq \overline{v}$, and $\tilde{\lambda}^*(x) = \tilde{\lambda}'(\overline{v})x$ if $x \geq \overline{v}$. Consider the first-order differential equation that coincides with (A.1) for $s(\cdot) \in S_1(\underline{v})$

$$\widetilde{E}^*(z, I) \;:\; s'(v) = (I-1)\frac{f(v|z, I)}{F(v|z, I)}\tilde{\lambda}^*\left(v - s(v)\right), \; s(v_0) = s_0(v_0; z, I), \; v \in [v_0, \overline{v}], \qquad \text{(E.5)}$$

where $0 < s_0(v_0; z, I) < v_0$ by Step 2. Define

$$\mathcal{L} = (I-1)\sup_{v\in[v_0,\overline{v}]}\frac{f(v|z, I)}{F(v|z, I)}\sup_{x\in[0,\overline{v}]}\left|\lambda'(x)\right| < \infty, \qquad Q_0 = [v_0, \overline{v}] \times I\!R,$$

so for any (v, b_1), (v, b_2) in Q_0,

$$\left|(I-1)\frac{f(v|z, I)}{F(v|z, I)}\tilde{\lambda}^*\left(v - b_1\right) - (I-1)\frac{f(v|z, I)}{F(v|z, I)}\tilde{\lambda}^*\left(v - b_2\right)\right| \leq \mathcal{L}|b_1 - b_2|,$$

by the Mean Function Theorem. It now follows from the Picard-Lindelöf Theorem (see Corollary 1.10 in Zeidler (1985)) that $\widetilde{E}^*(z, I)$ has exactly one continuously differentiable solution, say $s_1(v; z, I)$ for $v \in [v_0, \overline{v}]$.

Define $s(v; z, I) = s_0(v; z, i)$ for $v \in [\underline{v}, v_0]$, $s(v; z, I) = s_1(v; z, I)$ for $v \in [v_0, \overline{v}]$. We check that $s(\cdot; z, I)$ is the unique continuously differentiable solution of (E.4) on $[\underline{v}, \overline{v}]$ with $s(\underline{v}; z, I) = \underline{v}$. Note first that $s(v; z, I)$ is continuously differentiable at v_0 by (E.4) and (E.5), the initial condition $s_1(v_0; z, I) = s_0(v_0; z, I)$, and $0 < v_0 - s_0(v_0; z, I) < \overline{v}$. To establish that $s(\cdot; z, I)$ solves (E.4) for $v \in [\underline{v}, \overline{v}]$ with $s(\underline{v}; z, I) = \underline{v}$, it suffices to show that $s_1(v; z, I)$ solves (E.4) for $v \in [v_0, \overline{v}]$ with $s_1(v_0; z, I) = s_0(v_0; z, I)$, since $s_0(v; z, I)$ already solves (E.4) for $v \in [\underline{v}, v_0]$ with $s_0(\underline{v}; z, I) = \underline{v}$. In view of (E.5), it suffices to show that $v - s_1(v; z, I) \leq \overline{v}$ for $v \in [v_0, \overline{v}]$. The latter holds if $s_1(\cdot; z, I) > 0$, which is implied by $s_1(v_0; z, I) = s_0(v_0; z, I) > 0$ and $s_1'(\cdot; z, I) > 0$ on $[v_0, \overline{v}]$, which follows by repeating the second part of Step 1 to establish that $s_1(v; z, I) < v$

for $v \in [v_0, \overline{v}]$. Uniqueness of $s(\cdot; z, I)$ on $[\underline{v}, \overline{v}]$ follows from the uniqueness of $s_0(\cdot; z, I)$ and $s_1(\cdot; z, I)$ on $[\underline{v}, v_0]$ and $[v_0, \overline{v}]$, respectively.

Proof of Lemma A5: Fix $I \in \mathcal{I}$. We first show that the mapping $(v, z) \to \sigma_I(\cdot; v, z)$ from $[\underline{v}, \overline{v}] \times \mathcal{Z}$ to the Banach space $\left(C_1^0, \|\cdot\|_{1,\infty}\right)$ is R continuously differentiable. By (A.8) and Lemma A4, $\sigma_I(\cdot; v, z)$ is a (unique) zero of the operator $\mathbf{E}_I(\cdot; v, z)$ in $\Sigma_1^0 \subset C_1^0$, i.e.

$$\mathbf{E}_I(\sigma_I(\cdot; v, z); v, z) = 0 \text{ for all } (v, z) \in [\underline{v}, \overline{v}] \times \mathcal{Z}.$$

It follows from Lemma A3 that $\mathbf{E}_I(\cdot; v, z)$ verifies the conditions of the Functional Implicit Function Theorem 4.B-(d) in Zeidler (1985). Therefore $(v, z) \to \sigma_I(\cdot; v, z)$ is R continuously differentiable on the interior of $[\underline{v}, \overline{v}] \times \mathcal{Z}$. Now note that the partial derivatives of $\mathbf{E}_I(\cdot; v, z)$ are continuous over $[\underline{v}, \overline{v}] \times \mathcal{Z}$. It then follows from the expression of the implicit partial derivatives (see the proof of Theorem 4.B in Zeidler (1985)) and arguing as in the proof of Lemma C1 in Guerre, Perrigne and Vuong (2000), that the partial derivatives of $(v, z) \to \sigma_I(\cdot; v, z)$ have limits at the boundaries, so that this mapping is R continuously differentiable with respect to $(v, z) \in [\underline{v}, \overline{v}] \times \mathcal{Z}$.

To prove the desired result, note that $\sigma \in (C_1^0, \|\cdot\|_{1,\infty}) \mapsto \sigma(1)$ is a linear continuous mapping, which is therefore infinitely continuously differentiable. It then follows from the differentiation rule with respect to composition that $\sigma_I(1; v, z)$ is R continuously differentiable with respect to $(v, z) \in [\underline{v}, \overline{v}] \times \mathcal{Z}$. Hence, by (A.4), $s(v; z, I) = \underline{v} + (v - \underline{v})\sigma_I(1; v, z)$ is R continuously differentiable with respect to $(v, z) \in [\underline{v}, \overline{v}] \times \mathcal{Z}$. Next, for $u = 1$, (A.5) yields

$$s'(v; z, I) = \sigma_I^{(1)}(1; v, z) = \Phi_I(1; v, z)\Lambda\left(1 - \sigma_I(1; v, z); v\right), \; (v, z) \in [\underline{v}, \overline{v}] \times \mathcal{Z}.$$

Hence, $s'(v; z, I)$ is R continuously differentiable with respect to $(v, z) \in [\underline{v}, \overline{v}] \times \mathcal{Z}$ by Lemma A2, the derivation rules for composition and product of functions, and $0 \leq \sigma_I(1; v, z) < 1$.

Proof of Lemma A6: Fix $I \in \mathcal{I}$ and let $\xi(\cdot; z, I) = s^{-1}(\cdot; z, I)$.

(i) We have $\underline{b}(z, I) = s(\underline{v}; z, I) = \underline{v}$ and $\overline{b}(z, I) = s(\overline{v}; z, I) > \underline{v}$. Hence, $\inf_{z \in \mathcal{Z}} \left(\overline{b}(z, I) - \underline{b}(z, I)\right) > 0$ by continuity of $s(\overline{v}; \cdot, I)$ and compactness of $\mathcal{Z}$. The conditional distribution $G(\cdot|z, I) = F(\xi(\cdot; z, I)|z, I)$ has support $[\underline{b}(z, I), \overline{b}(z, I)]$ because $F(\cdot|z, I)$ and $s(\cdot; z, I)$ are strictly increasing and continuous on $[\underline{v}, \overline{v}]$. Its density is $g(b|z, I) = f(s^{-1}(b; z, I)|z, I)/\ s'(s^{-1}(b; z, I); z, I)$. Thus $\inf_{(b,z) \in S_I(G)} g(b|z, I) > 0$ using Definition 4-(iii).

We now study the smoothness of $G(\cdot|\cdot, I)$. Note that Lemma A5 shows that $G(b|z, I)$ and $g(b|z, I)$ admit up to R continuous partial derivatives on $S_I(G)$ since $\inf_{(v,z) \in [\underline{v}, \overline{v}] \times \mathcal{Z}} s'(v; z, I) > 0$ by definition of $S_1(\underline{v})$ and Lemmas A4–A5. Thus, it remains to prove that $\partial G^{R+1}(\cdot|\cdot, I)/\partial z^{R+1}$

exists and is continuous on $S_I(G)$. Let $G^{(1)}(b|z,I) = g(b|z,I)$. Using the expression of the latter and the fact that $s(\cdot;z,I)$ solves (A.1) gives

$$G^{(1)}(b|z,I) = \frac{F\left[s^{-1}(b;z,I)|z,I\right]}{(I-1)\lambda\left[s^{-1}(b;z,I)-b\right]} = \frac{G(b|z,I)}{(I-1)\lambda\left[F^{-1}\left(G(b|z,I)|z,I\right)-b\right]}, \tag{E.6}$$

since $G(\cdot|z,I) = F(s^{-1}(\cdot;z,I)|z,I)$ and $F^{-1}(G(b|z,I)|z,I) = s^{-1}(b;z,I) > b$ on $(\underline{b}(z,I),\overline{b}(z,I)]$ as $s(\cdot;z,I) \in S_1(\underline{v})$. Thus, the distribution $G(\cdot|z,I)$ solves the first-order differential equation

$$E_I^G(z) \ : \ G^{(1)}(b) = \frac{G(b)}{(I-1)\lambda\left[F^{-1}\left(G(b)|z,I\right)-b\right]}, \ G(\overline{b}(z,I)) = 1, \ b \in (\underline{b}(z,I),\overline{b}(z,I)],$$

excluding $\underline{b}(z,I)$ in order to avoid the singularity at that point. Consider (b_0,z_0) in the interior of $S_I(G)$. The function

$$(b,G,z) \to \frac{G}{(I-1)\lambda\left[F^{-1}\left(G|z,I\right)-b\right]}$$

admits up to $R{+}1$ continuous partial derivatives on a neighborhood of $(b_0, G(b_0|z_0,I), z_0)$. Hence the Implicit Function Theorem 4.D for first-order differential equations in Zeidler (1985, p.165) yields that the solution $G(b|z,I)$ of $E_I^G(z)$ has an $R+1$ continuous z-derivative at (b_0,z_0), and thus in the interior of $S_I(G)$. Arguing as in the first part of the proof of Lemma A5, the proof of Theorem 4.D in Zeidler (1985) shows that this extends to the upper boundary $\overline{b}(\cdot,I)$.

We now extend this result to the lower boundary $\underline{b}(\cdot,I) = \underline{b} = \underline{v}$, assuming $d=1$ for sake of simplicity. Because $\mathcal{Z}$ is rectangular, differentiating (E.6) with respect to z for any $b > \underline{b}$ yields

$$\begin{aligned}
&(I-1)\frac{\partial}{\partial b}\left(\frac{\partial G(b|z,I)}{\partial z}\right) \\
&= \frac{\partial G(b|z,I)/\partial z}{\lambda\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)} \\
&\quad - \frac{G(b|z,I)}{\lambda^2\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)}\lambda'\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)\frac{\partial F^{-1}\left(G(b|z,I)|z,I\right)}{\partial z}\frac{\partial G(b|z,I)}{\partial z} \\
&= \left[1-(I-1)g(b|z,I)\lambda'\big(F^{-1}(G(b|z,I)|z,I)-b\big)\frac{\partial F^{-1}(G(b|z,I)|z,I)}{\partial z}\right]\frac{\partial G(b|z,I)/\partial z}{\lambda\left(F^{-1}(G(b|z,I)|z,I)-b\right)} \\
&\equiv \frac{A(b;z,I)}{\lambda\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)}\frac{\partial G(b|z,I)}{\partial z}.
\end{aligned}$$

Note that $A(\cdot;\cdot,I)$ has up to R continuous partial derivatives on $S_I(G)$. Moreover, $\lim_{b\downarrow\underline{b}}\sup_{z\in\mathcal{Z}} |A(b;z,I)-1| = 0$ since $F^{-1}(0|z,I) = \underline{v}$ is independent of z. Thus, there exists a b_0 close enough to $\underline{b}$ such that $A(b;z,I) > 0$ on $[\underline{b},b_0]\times\mathcal{Z}$. For $1 \le r \le R+1$ and $b > \underline{v}$, define the property

$$\mathcal{P}_r \ : \ (I-1)\frac{\partial}{\partial b}\left(\frac{\partial^r G(b|z,I)}{\partial z^r}\right) = D_r(b;z,I) + \frac{A(b;z,I)}{\lambda\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)}\frac{\partial^r G(b|z,I)}{\partial z^r},$$

where $D_r(\cdot;\cdot,I)$ admits up to $R+1-r$ partial continuous derivatives on $[\underline{b},b_0]\times\mathcal{Z}$. Note that $\mathcal{P}_1$ is true. We prove by induction that $\mathcal{P}_{R+1}$ is also true.

Suppose that $\mathcal{P}_r$ is true for some r with $1\le r\le R$. Then $\mathcal{P}_{r+1}$ is true since $\mathcal{P}_1$ and $\mathcal{P}_r$ yield

$$
\begin{aligned}
&(I-1)\frac{\partial}{\partial b}\left(\frac{\partial^{r+1}G(b|z,I)}{\partial z^{r+1}}\right)\\
&= \frac{\partial D_r(b;z,I)}{\partial z}+\frac{\partial A(b;z,I)}{\partial z}\frac{\partial^r G(b|z,I)/\partial z^r}{\lambda\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)}\\
&\quad -A(b;z,I)\frac{\partial^r G(b|z,I)}{\partial z^r}\frac{\partial F^{-1}\left(G(b|z,I)|z,I\right)}{\partial z}\frac{\partial G(b|z,I)}{\partial z}\frac{\lambda'\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)}{\lambda^2\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)}\\
&\quad +\frac{A(b;z,I)}{\lambda\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)}\frac{\partial^{r+1}G(b|z,I)}{\partial z^{r+1}}\\
&= \frac{\partial D_r(b;z,I)}{\partial z}+\frac{1}{A(b;z,I)}\frac{\partial A(b;z,I)}{\partial z}\left((I-1)\frac{\partial^r g(b|z,I)}{\partial z^r}-D_r(b;z,I)\right)\\
&\quad -(I-1)\frac{\partial^r G(b|z,I)}{\partial z^r}\frac{\partial F^{-1}\left(G(b|z,I)|z,I\right)}{\partial z}\frac{\partial g(b|z,I)}{\partial z}\frac{\lambda'\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)}{\lambda\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)}\\
&\quad +\frac{A(b;z,I)}{\lambda\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)}\frac{\partial^{r+1}G(b|z,I)}{\partial z^{r+1}}\\
&\equiv D_{r+1}(b;z,I)+\frac{A(b;z,I)}{\lambda\left(F^{-1}\left(G(b|z,I)|z,I\right)-b\right)}\frac{\partial^{r+1}G(b|z,I)}{\partial z^{r+1}},
\end{aligned}
$$

where $D_{r+1}(b;z,I)$ has up to $R-r\le R-1$ continuous partial derivatives at $\underline{b}$ and thus on $[\underline{b},b_0]\times\mathcal{Z}$ by proceeding as in Lemma A2 and observing that $\partial F^{-1}\left(G(\underline{b}|z,I)|z,I\right)/\partial z=0$ for all $z\in\mathcal{Z}$, and that $G(\cdot|\cdot,I)$, $g(\cdot|\cdot,I)$ and $A(\cdot;\cdot,I)$ are R-continuously differentiable on $S_I(G)$.

Therefore, using $\xi(b;z,I)=F^{-1}(G(b|z,I)|z,I)$, $\mathcal{P}_{R+1}$ yields

$$
(I-1)\frac{\partial}{\partial b}\left(\frac{\partial^{R+1}G(b|z,I)}{\partial z^{R+1}}\right)=D_{R+1}(b;z,I)+\frac{A(b;z,I)}{\lambda\left(\xi(b;z,I)-b\right)}\frac{\partial^{R+1}G(b|z,I)}{\partial z^{R+1}}
$$

for $b>\underline{b}$. Solving this differential equation on $(\underline{b},b_0]$ gives the unique solution

$$
\begin{aligned}
\frac{\partial^{R+1}G(b|z,I)}{\partial z^{R+1}} &= \exp\left(-\int_b^{b_0}\frac{A(y;z,I)}{(I-1)\lambda\left(\xi(y;z,I)-y\right)}dy\right)\\
&\quad\times\left[\int_{b_0}^{b}\frac{D_{R+1}(y;z,I)}{I-1}\exp\!\left(\int_y^{b_0}\frac{A(u;z,I)}{(I-1)\lambda\left(\xi(u;z,I)-u\right)}du\right)dy+\frac{\partial^{R+1}G(b_0|z,I)}{\partial z^{R+1}}\right]\\
&= \int_{b_0}^{b}\frac{D_{R+1}(y;z,I)}{I-1}\exp\left(-\int_b^{y}\frac{A(u;z,I)}{(I-1)\lambda\left(\xi(u;z,I)-u\right)}du\right)dy\\
&\quad+\frac{\partial^{R+1}G(b_0|z,I)}{\partial z^{R+1}}\exp\left(-\int_b^{b_0}\frac{A(y;z,I)}{(I-1)\lambda\left(\xi(y;z,I)-y\right)}dy\right). \qquad \text{(E.7)}
\end{aligned}
$$

Without loss of generality, assume $\underline{b}=0$. ¿From Lemma A1, we have $\xi'(\underline{b};z,I)=[(I-1)\lambda'(0)+$

$1]/[(I-1)\lambda'(0)]$. Hence, when $b \downarrow 0$,

$$\begin{aligned} A(b;z,I) &= 1+o(1), \\ \frac{1}{(I-1)\lambda\left(\xi(b;z,I)-b\right)} &= \frac{1}{b}+O(1), \\ \exp\left(-\int_b^{b_0} \frac{A(y;z,I)}{(I-1)\lambda\left(\xi(y;z,I)-y\right)}dy\right) &= b\exp(O(1)) = O(b), \end{aligned}$$

uniformly in $z \in \mathcal{Z}$ because $\lambda'(0) \geq 1$. By definition of b_0, which is such that $A(b;z,I) > 0$ for any $(b,z) \in (\underline{b}, b_0] \times \mathcal{Z}$, (E.7) yields for every $b \in (\underline{b}, b_0]$

$$\sup_{z\in\mathcal{Z}} \left|\frac{\partial^{R+1}G(b|z,I)}{\partial z^{R+1}}\right| \leq \frac{b_0-b}{I-1} \sup_{(b,z)\in(\underline{b},b_0]\times\mathcal{Z}} |D_{R+1}(b;z,I)| + O(b)\sup_{z\in\mathcal{Z}}\left|\frac{\partial^{R+1}G(b_0|z,I)}{\partial z^{R+1}}\right|.$$

Consider now an arbitrary $\epsilon > 0$. Note first that b_0 can be chosen small enough so that the first term is smaller than $\epsilon/2$. Next, observe that there is a $0 < \eta < b_0$ such that the second term is smaller than $\epsilon/2$ whenever $b < \eta$. Consequently, for any $\epsilon > 0$ there is a $\eta > 0$ such that $\sup_{(b,z)\in(0,\eta]\times\mathcal{Z}} \left|\partial^{R+1}G(b|z,I)/\partial z^{R+1}\right| \leq \epsilon$. Hence, $\lim_{b\downarrow 0}\sup_{z\in\mathcal{Z}}\left|\partial^{R+1}G(b|z,I)/\partial z^{R+1}\right| = 0$. Collecting results, $G(\cdot|\cdot,I)$ admits up to $R+1$ continuous partial derivatives on $S_I(G)$.

(ii) We have $s(v;z,I) = G^{-1}\left(F(v|z,I)|z,I\right)$. Thus (i) shows that $s(\cdot;\cdot,I)$ admits up to $R+1$ continuous partial derivatives on $[\underline{v},\overline{v}] \times \mathcal{Z}$.

E.2 Proofs of Lemmas C1–C2

Proof of Lemma C1: We have

$$m(z,I;\beta) = \frac{1}{I-1}\frac{1}{\lambda\left(\overline{v}-\overline{b}_0(z,I);\theta\right)} = \frac{1}{I-1}\frac{U'\left(\overline{v}-\overline{b}_0(z,I);\theta\right)}{U\left(\overline{v}-\overline{b}_0(z,I);\theta\right)},$$

where $(\theta,\overline{v}) \in \mathcal{B}$. By Lemma 2-(i), $\overline{b}_0(\cdot,I)$ has $R+1$ continuous derivatives on $\mathcal{Z}$ with $R \geq 1$. The desired result follows from the $R+2$ continuous differentiability of $U(\cdot;\cdot)$ 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 A1-(i) applied at $v = \overline{v}_0$.

Proof of Lemma C2: The proof is in four steps.

STEP 1: *$G_1(\cdot|\cdot,\cdot)$ satisfies the properties of Lemma 2.* Let $\Psi(b) = \int_{-\infty}^b \psi(x)dx$. We have

$$\begin{aligned} G_1(b|z,I) &= G_0(b|z,I) + \pi_N(z,I)\int_{-\infty}^b \psi\left(\kappa\rho_N^{\frac{1}{R+1}}\left(x-\overline{b}_0(z,I)\right)\right)dx \\ &= G_0(b|z,I) + \pi_N(z,I)\kappa^{-1}\rho_N^{-\frac{1}{R+1}}\Psi\left(\kappa\rho_N^{\frac{1}{R+1}}\left(b-\overline{b}_0(z,I)\right)\right). \end{aligned}$$

In particular, $G_1(\cdot|z,I)$ and $g_1(\cdot|z,I)$ are equal to $G_0(\cdot|z,I)$ and $g_0(\cdot|z,I)$ on $[\underline{v}_0, \overline{b}_0(z,I) - \rho_N^{-1/(R+1)}]$, while differing from the latter by an $O(1/\rho_N)$ on $[\overline{b}_0(z,I) - \rho_N^{-1/(R+1)}, \overline{b}_0(z,I)]$. 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 C1. Therefore, for N large enough, $G_1(\cdot|\cdot,\cdot)$ satisfies the properties of Lemma 2 with $\underline{b}_1(z,I) = \underline{b}_0(z,I) = \underline{v}_0$ and $\overline{b}_1(z,I) = \overline{b}_0(z,I)$ 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 $G_1(\cdot|\cdot,I)/g_1(\cdot|\cdot,I)$ is $R+1$ continuously differentiable on $S_I(G_0)$ for every $I \in \mathcal{I}$.

Step 2: We study the difference $\xi_1(\cdot|z,I) - \xi_0(\cdot|z,I)$ and its derivatives. Lemma 2 yields

$$\frac{\partial^r G_1(b|z,I)}{\partial b^r} - \frac{\partial^r G_0(b|z,I)}{\partial b^r} = \pi_N(z,I)\kappa^{r-1}\rho_N^{\frac{r-1}{R+1}}\Psi^{(r)}\Big(\rho_N^{\frac{1}{R+1}}\big(b - \overline{b}_0(z,I)\big)\Big), b \in [\underline{v}_0, \overline{b}_0(z,I)],$$

$$\frac{\partial^r g_1(b|z,I)}{\partial b^r} - \frac{\partial^r g_0(b|z,I)}{\partial b^r} = \pi_N(z,I)\kappa^{r}\rho_N^{\frac{r}{R+1}}\psi^{(r)}\Big(\rho_N^{\frac{1}{R+1}}\big(b - \overline{b}_0(z,I)\big)\Big), b \in [\underline{v}_0, \overline{b}_0(z,I)],$$

$$\frac{\partial^{R+1} g_1(b|z,I)}{\partial b^{R+1}} - \frac{\partial^{R+1} g_0(b|z,I)}{\partial b^{R+1}} = \pi_N(z,I)\kappa^{R+1}\rho_N\psi^{(R+1)}\Big(\rho_N^{\frac{1}{R+1}}\big(b - \overline{b}_0(z,I)\big)\Big), b \in (\underline{v}_0, \overline{b}_0(z,I)],$$

where the first and second equalities hold for $0 \le r \le R+1$ and $0 \le r \le R$, respectively. By Lemma 2-(i) there is a b_* with $\underline{v}_0 < b_* < \overline{b}_0(z,I)$ for all (z,I). Because $\sup_{z,I}|\pi_N(z,I)| = O(1/\rho_N)$, this gives for $r = 0,\ldots,R+1$,

$$\sup_{(b,z,I)\in\cup_{\mathcal{Z},\mathcal{I}}[b_*,\overline{b}_0(z,I)]\times\{z,I\}} \left|\frac{\partial^r G_1(b|z,I)}{\partial b^r} - \frac{\partial^r G_0(b|z,I)}{\partial b^r}\right| = \kappa^{r-1} O\left(\rho_N^{\frac{r-1-(R+1)}{R+1}}\right), \quad \text{(E.8)}$$

$$\sup_{(b,z,I)\in\cup_{\mathcal{Z},\mathcal{I}}[b_*,\overline{b}_0(z,I)]\times\{z,I\}} \left|\frac{\partial^r g_1(b|z,I)}{\partial b^r} - \frac{\partial^r g_0(b|z,I)}{\partial b^r}\right| = \kappa^{r} O\left(\rho_N^{\frac{r-(R+1)}{R+1}}\right), \quad \text{(E.9)}$$

where the remainder terms are independent of κ.

Observe now that, for L large enough and $(b,z,I) \in [\underline{v}_0, \overline{b}_0(z,I)] \times \mathcal{Z} \times \mathcal{I}$,

$$\begin{aligned}
&\left|\frac{\partial^r \xi_1(b;z,I)}{\partial b^r} - \frac{\partial^r \xi_0(b;z,I)}{\partial b^r}\right| \\
&\le \left|\frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_1(b|z,I)}{(I-1)g_1(b|z,I)};\theta_1\right)\right] - \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)};\theta_1\right)\right]\right| \\
&\quad + \left|\frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)};\theta_1\right)\right] - \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)};\theta_0\right)\right]\right| \\
&\le \sup_{\theta\in\Theta}\left|\frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_1(b|z,I)}{(I-1)g_1(b|z,I)};\theta\right)\right] - \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)};\theta\right)\right]\right| \\
&\quad + \left|\frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)};\theta_1\right)\right] - \frac{\partial^r}{\partial b^r}\left[\lambda^{-1}\left(\frac{G_0(b|z,I)}{(I-1)g_0(b|z,I)};\theta_0\right)\right]\right|. \quad \text{(E.10)}
\end{aligned}$$

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 $k =$

$0, \dots, r$. Therefore, for $r = 0, \dots, R+1$ the first term in (E.10) vanishes when $b \in [\underline{v}_0, b_*]$ and is of order $\kappa^r O\left(\rho_N^{(r-(R+1))/(R+1)}\right)$ uniformly on $[b_*, \overline{b}_0(z, I)]$ by (E.8)-(E.9). Regarding the second term in (E.10), note that $\lambda^{-1}(x;\theta)$ is $R+1$ continuously differentiable on $[0, +\infty) \times \Theta$ because $\lambda(x;\theta)$ 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 $\partial^r \lambda^{-1}\left(G_0(b|z,I)/[(I-1)g_0(b|z,I)];\theta\right)/\partial b^r$ is continuous on $S_I(G_0) \times \Theta$. Hence, the second term is of order $\|\beta_1 - \beta_0\|_\infty = O(t_N/\rho_N) = o(1)$ uniformly on $S_I(G_0)$, for $r = 0, \dots, R+1$, because $t_N/\rho_N = o(1)$. Collecting results and using the finiteness of $\mathcal{I}$, (E.10) yields

$$\sup_{I \in \mathcal{I}} \sup_{(b,z) \in S_I(G_0)} \left| \frac{\partial^r \xi_1(b;z,I)}{\partial b^r} - \frac{\partial^r \xi_0(b;z,I)}{\partial b^r} \right| = \kappa^r O\left(\rho_N^{\frac{r-(R+1)}{R+1}}\right) + o(1) \ , \ r = 0, \dots, R+1. \quad \text{(E.11)}$$

Step 3: *Proof of (i).* Because $R \geq 1$, applying (E.11) for $r = 1$ yields that $\xi_1'(\cdot; z, I) > 0$ on $[\underline{v}_0, \overline{b}_0(z,I)]$ for every $(z, I) \in \mathcal{Z} \times \mathcal{I}$. Because $G_1(\cdot|z,I)$ satisfies Definition 3, Lemma 1 shows that $G_1(\cdot|z,I)$ is rationalized by $[U(\cdot;\theta_1), F_1(\cdot|z,I)]$ for every $(z,I) \in \mathcal{Z} \times \mathcal{I}$, where $F_1(\cdot|z,I)$ is the distribution of $\xi_1(b;z,I)$ with $b \sim G_1(\cdot|z,I)$.

We now check that $F_1(\cdot|\cdot,\cdot) \in \mathcal{F}_R^*$, as required by Assumption A2-(ii). The support of $F_1(\cdot|z,I)$ is the same as $F_0(\cdot|z,I)$, namely $[\underline{v}_0, \overline{v}_0]$. Indeed, we have $\underline{v}_1(z,I) = \xi_1(\underline{v}_0; z, I) = \underline{v}_0$. Moreover, by definitions of $\lambda(\cdot;\theta)$ and $g_1(\cdot|z,I)$, and $\psi(0) = 1$

$$\begin{aligned}
\overline{v}_1(z,I) &= \xi_1(\overline{b}_0(z,I); z, I) \\
&= \overline{b}_0(z,I) + \lambda^{-1}\left(\frac{1}{I-1}\frac{1}{g_0(\overline{b}_0(z,I)|z,I) + m(z,I;\beta_1) - m(z,I;\beta_0)};\theta_1\right) \\
&= \overline{b}_0(z,I) + \lambda^{-1}\left(\frac{1}{I-1}\frac{1}{m(z,I;\beta_1)};\theta_1\right) \\
&= \overline{b}_0(z,I) + \lambda^{-1}\left(\lambda\left(\overline{v}_0 - \overline{b}_0(z,I);\theta_1\right);\theta_1\right) = \overline{v}_0,
\end{aligned}$$

because $g_0(\overline{b}_0(z,I)|z,I) = m(z,I;\beta_0)$ by (6) and the definition of $m(z,I;\beta)$. Also, it can be seen that $F(\cdot|\cdot,I)$ 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 $\xi_1(\cdot;\cdot,I)$ and $G_1(\cdot|\cdot,I)$ as $F_1(\cdot|z,I) = G_1(\xi_1^{-1}(\cdot;z,I)|z,I)$.

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 $s_1(v;z,I) - s_0(v;z,I)$ also satisfies (E.11), where $\xi_j(b;z,I)$ is replaced by $s_j(v;z,I)$, for $j = 0, 1$. Now, $f_1(v|z,I) = g_1(s_1(v;z,I)|z,I)s_1'(v;z,I)$. Thus, following Step 3 of that lemma we obtain

$$\sup_{I \in \mathcal{I}} \sup_{(v,z) \in [\underline{v}_0, \overline{v}_0] \times \mathcal{Z}} \left| \frac{\partial^r f_1(v|z,I)}{\partial v^r} - \frac{\partial^r f_0(v|z,I)}{\partial v^r} \right| = \kappa^{r+1} O\left(\rho_N^{\frac{r-R}{R+1}}\right) + o(1), \quad r = 0, 1, \dots, R.$$

In particular, for $r = 0$ we obtain $\|f_1(\cdot|\cdot,\cdot) - f_0(\cdot|\cdot,\cdot)\|_\infty < \epsilon$ for N sufficiently large. Moreover, for $r = R$ the triangular inequality gives $\|\partial^R f_1(\cdot|\cdot,\cdot)/\partial v^R\| < M$ because $\|\partial^R f_0(\cdot|\cdot,\cdot)/\partial v^R\| < M$ by assumption, provided κ is sufficiently small.

E.3 Proofs of Lemmas D1–D7

Proof of Lemma D1: (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 $\epsilon_{i\ell}$ as $\epsilon_{i\ell} = Y_{i\ell} - \mathrm{E}[Y_{i\ell}|\mathcal{F}_L]$.

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

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

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

(iii) Similarly, using (ii) and Lemma 2, we have

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

Proof of Lemma D2: It suffices to check the Lyapounov condition of Theorem 7.3 in Billingsley (1968) given $\mathcal{F}_L$. From Lemma D1 we have

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

Thus, $\mathrm{E}[|u_{i\ell}\epsilon_{i\ell}|^3|\mathcal{F}_L] \leq \mathrm{E}[|u_{i\ell}\epsilon_{i\ell}|^2|\mathcal{F}_L]O(\|u\|_\infty/h_N)$. 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

$$\begin{aligned}
\frac{1}{\mathrm{Var}^{3/2}[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell} u_{i\ell}\epsilon_{i\ell}|\mathcal{F}_L]}\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell}\mathrm{E}[|u_{i\ell}\epsilon_{i\ell}|^3|\mathcal{F}_L] &\leq \frac{O(\|u\|_\infty/h_N)}{\mathrm{Var}^{1/2}[\sum_{\ell=1}^{L}\sum_{i=1}^{I_\ell} u_{i\ell}\epsilon_{i\ell}|\mathcal{F}_L]} \\
&= O\left(\frac{\|u\|_\infty}{\sqrt{h_N}\|u\|_2}\right),
\end{aligned}$$

which is an $o_P(1)$ by assumption.

Proof of Lemma D3: In view of Assumption A4-(v), let $\overline{a}_N \asymp a_N$ be such that the event $\mathcal{E}_N = \{\max_{i,\ell}|\hat{\overline{b}}(Z_\ell) - \overline{b}_0(Z_\ell)| \leq \overline{a}_N\}$ has a probability larger than $1-\epsilon$, where $\epsilon > 0$ can be chosen arbitrary small. Hereafter, we consider that $\mathcal{E}_N$ occurs. By Assumption 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 $B_{i\ell} \leq \overline{b}_0(Z_\ell, I_\ell) - \overline{a}_n - h_n$. Note also that $\overline{a}_N < h_N$ for N sufficiently large by Assumption 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 $\overline{b}_0(Z_\ell, I_\ell) - \overline{a}_n - h_n \leq B_{i\ell} \leq \overline{b}_0(Z_\ell, I_\ell) - 2\overline{a}_n$, while we use that $\Phi(\cdot)$ is bounded when $\overline{b}_0(Z_\ell, I_\ell) - 2\overline{a}_n \leq B_{i\ell} \leq \overline{b}_0(Z_\ell, I_\ell)$. This gives

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

Let $\zeta_{i\ell}$ denote the right-hand side. Now, because $\overline{a}_N = o(h_N)$ by Assumption 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

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

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 $E[|X|] \leq E^{1/2}[X^2]$, the Markov inequality conditional on $\mathcal{F}_L$ and that $\Pr(\mathcal{E}_N) \geq 1 - \epsilon$ establishes the desired results.

Proof of Lemma D4: By definition of $\mathcal{B}_\delta$, we have $\delta \leq \overline{v} - \overline{b}_0(Z_\ell, I_\ell) \leq \overline{v}_{\sup}$ for all ℓ. By Assumption A4-(v), we have $|\hat{\overline{b}}(Z_\ell, I_\ell) - b_0(Z_\ell, I_\ell)| < \delta/2$ for all ℓ with probability tending to 1. Thus, $\delta/2 \leq \overline{v} - \hat{\overline{b}}(Z_\ell, I_\ell) \leq \overline{v}_{\sup} + \delta/2$ for all ℓ with probability tending to 1. 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 tending to 1. The desired result follows from Assumption A2-(i) since $\lambda(x;\theta)$ is R+1 continuously differentiable on $\in [0, +\infty] \times \Theta$, and hence uniformly continuous on the compact $[\delta/2, \overline{v}_{\sup} + \delta/2] \times \Theta$. The study of the derivatives is similar.

Proof of Lemma D5: By Assumptions A3-(i) and A4-(ii), there exists some $\underline{C}$ such that

$$\overline{Q}(\beta) \geq \underline{C} \sum_{I \in \mathcal{I}} \int_{\mathcal{Z}} \left(m(z, I; \beta) - m(z, I; \beta_0)\right)^2 dz.$$

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 $m(z, I; \beta) - m(z, I; \beta_0)$ with respect to z whenever $\beta \in \mathcal{B}_\delta$ in view of Lemma C1 and $\mathcal{B}_\delta \subset \mathcal{B}$. By definition of $m(z, I; \beta_0)$, this gives

$$\lambda\left(\overline{v} - \overline{b}_0(z, I); \theta\right) = \lambda\left(\overline{v}_0 - \overline{b}_0(z, I); \theta_0\right)\ ,\ \forall (z, I) \in \mathcal{Z} \times \mathcal{I}.$$

Hence, Assumption A2-(iii) yields $\beta = \beta_0$. Therefore, $\overline{Q}(\beta) = 0$ if and only if $\beta = \beta_0$. Moreover, because $\mathcal{B}_\delta$ is compact, Lemma C1 yields that $\overline{Q}(\cdot)$ is continuous on $\mathcal{B}_\delta$ and hence on $\mathcal{B}_\delta \cap \{\|\beta - \beta_0\| \geq \epsilon\}$ by the Lebesgue Dominated Convergence Theorem. This implies the first claim.

Consider the matrix $A(\beta)$, the case of $B(\beta)$ being similar. We have

$$\begin{aligned}\frac{\partial m(z, I; \beta)}{\partial \beta} &= -\frac{1}{I - 1}\frac{1}{\lambda^2\left(\overline{v} - \overline{b}_0(z, I); \theta\right)}\frac{\partial \lambda\left(\overline{v} - \overline{b}_0(z, I); \theta\right)}{\partial \beta} \\ &= -(I - 1)m^2(z, I; \beta)\frac{\partial \lambda\left(\overline{v} - \overline{b}_0(z, I); \theta\right)}{\partial \beta},\end{aligned}$$

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 $m(z, I; \beta)$ does not vanish and is continuous on $\mathcal{Z} \times \mathcal{B}_\delta$ by Lemma C1. Moreover, the Lebesgue Dominated Convergence Theorem and Lemma C1 yield that $\det(A(\beta))$ is a continuous function of $\beta \in \mathcal{B}_\delta$. Thus, it is sufficient to show that $A(\beta_0)$ is of full rank. Assume that $A(\beta_0)$ is not of full rank. Then, there exists $t \in \mathbb{R}^{p+1} \setminus \{0\}$ with $t' A(\beta_0) t = 0$. Arguing as for $\overline{Q}(\cdot)$, we have

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

which implies

$$t' \cdot \frac{\partial \lambda \left(\overline{v}_0 - \overline{b}_0(z, I); \theta_0 \right)}{\partial \beta} = 0 \text{ , for almost all } z \in \mathcal{Z} \text{ and all } I \in \mathcal{I}.$$

This contradicts Assumption A4-(i) by continuity of $\lambda(\cdot; \theta_0)$.

Proof of Lemma D6: 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 C1 and the compactness of $\mathcal{B}_\delta$ and $\mathcal{Z}$. 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 e.g. 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 D7: 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 $w_{\ell L}$ iid within rows. Because L/N is bounded, it suffices to establish that $\sup_{\beta \in \mathcal{B}_\delta} \left| L^{-1/2} \ \sum_{\ell=1}^{L} w_{\ell L}(\beta) \right| = O_P(1)$. Now, by Lemma C1 and D1, Assumption A4-(ii) and the compactness of $\mathcal{Z}$, $\mathcal{I}$ and $\mathcal{B}_\delta$ imply that there exists some constant C such that

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

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

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

where C' and C'' are positive constants independent of L. Applying Markov inequality, this gives $\sup_{\beta \in \mathcal{B}_\delta} |W_N(\beta)| = O_P(1)$.

Regarding $W_N^{(1)}(\tilde{\beta}_N)$, note that $\tilde{\beta}_N \in \mathcal{B}_\delta$ with probability tending to 1, because $\beta_0 \in \mathcal{B}_\delta^o$ and $\tilde{\beta}_N = \beta_0 + o_P(1)$. Therefore, $W_N^{(1)}(\tilde{\beta}_N) = o_P(1)$ holds if we can show that $\sup_{\beta \in \mathcal{B}_\delta} |W_N^{(1)}(\beta)| = O_P(1)$. This can be done as above.

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

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. (2002): "Attitudes Toward Risk and Asymmetric Bidding: Evidence from Construction Procurement," Working Paper, University of North Carolina.

Campo, S., **I. Perrigne** and **Q. Vuong** (2002): "Asymmetry in First-Price Auctions with Affiliated Private Values," *Journal of Applied Econometrics,* forthcoming.

Chamberlain, G. (1987): "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions," *Journal of Econometrics*, 34, 305-334.

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

Cox, J., **V. Smith** and **J. Walker** (1988): "Theory and Individual Behavior of First-Price Auctions," *Journal of Risk and Uncertainty*, 1, 61-99.

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** (2001): "Precautionary Bidding in Auctions," Working Paper, Harvard University.

Goere, 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.

Haile, P. (2001): "Auctions with Resale Markets: An Application to US 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.

Harris, M. and **Raviv, A.** (1981): "Allocation Mechanisms and the Design of Auctions," *Econometrica*, 49, 1477-1499.

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.

Klemperer, P. (1999): "Auction Theory: A Guide to the Literature," *Journal of Economic Surveys*, 13, 227-260.

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

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, H. and **G. Tan** (2000): "Hidden Reserve Prices with Risk Averse Bidders," Working Paper, University of British Columbia.

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.

Lizzeri, A. and **N. Persico** (2000): "Uniqueness and Existence of Equilibrium in Auctions with a Reserve Price," *Games and Economic Behavior*, 30, 83-114.

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

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** (1996): "Uniqueness in Sealed High Bid Auctions," Working Paper, UCLA.

Maskin, E. and **J. Riley** (2000): "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. (2001): "Random Reserve Prices and Risk Aversion in Timber Auctions," Working Paper, University of Southern California.

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.

Riley, J. and **W. Samuelson** (1981): "Optimal Auctions," *American Economic Review,* 71, 381-392.

Robinson, P. (1988): "Root-N-Consistent Semiparametric Regression," *Econometrica*, 56, 931-954.

Stone, C.J. (1982): "Optimal Rates of Convergence for Nonparametric Regressions," *The 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 Aanlysis,* Econometric Society Monograph, 22, Cambridge University Press.

Zeidler, E. (1985): *Nonlinear Functional Analysis and its Applications, I. Fixed-Point Theorems*, Springer Verlag.