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

Federico Zincenko*

Department of Economics

University of Pittsburgh

April 25, 2016

Abstract

This paper proposes nonparametric estimators for the bidders' utility function and density of private values in a first-price sealed-bid auction model with independent valuations. I study a setting with risk-averse bidders and adopt a fully nonparametric approach by not placing any restrictions on the shape of the utility function beyond regularity conditions. I propose a population criterion function that has a unique minimizer, which characterizes the utility function and density of private values. The resulting estimators emerge after replacing the population quantities by sample analogues. These estimators are uniformly consistent and their convergence rates are established. Monte Carlo experiments show that the proposed estimators perform well in practice.

JEL Classification: C14, D44.

Keywords: First-price auction, risk aversion, independent private values, nonparametric estimation, sieve spaces.

*Mailing address: 230 South Bouquet St, 4713 Wesley W. Posvar Hall, Pittsburgh, PA 15260, U.S.A. Phone: +1-(412)624-7461. E-mail: zincenko@pitt.edu.

1 Introduction

Risk aversion is essential to understanding economic decisions under uncertainty. In first-price sealed-bid auctions, risk aversion plays a fundamental role in explaining bidders' behavior. Although several families of utility functions have been employed to describe different attitudes toward risk, in practice, we do not know which one accurately explains bidders' behavior.

I consider a first-price sealed-bid auction with risk-averse bidders within the independent private values paradigm. In this setting, each potential buyer has his own private value for the item being sold and makes a sealed bid. The buyer who makes the highest bid wins the item and pays the seller the amount of that bid. This model is completely characterized by two objects. The first is the bidders' utility function, which describes bidders' risk preferences. The second is the density of private values, which describes the distribution of valuations for the auctioned item.

This paper develops consistent estimators for these two objects without imposing parametric specifications. Only standard regularity conditions are assumed. These assumptions are satisfied by linear (risk-neutral), constant relative risk aversion (CRRA), and constant absolute risk aversion (CARA) utility functions, as well as, many others. In this sense, my paper generalizes the empirical analysis of first-price auctions by nesting many existing estimators within a fully nonparametric framework. The distinguishing feature of this paper is that the proposed estimator is based on assumptions that are milder than those typically made in the literature. No restrictions on the shape of the bidders' utility function are imposed beyond strict monotonicity, concavity, and differentiability. In particular, the bidder's utility is not assumed to belong to a specific family of risk aversion –such as constant relative risk aversion (CRRA) or constant absolute risk aversion (CARA)–.

This paper has two objectives. The first is to nonparametrically estimate the bidders' utility function. Empirical and experimental evidence indicates that risk aversion is a fundamental component of bidders' behavior.[1] Despite its relevance, only a few articles have proposed an estimator for the bidders' utility function. Campo, Guerre, Perrigne, and Vuong (2011), for instance, adopt a semi-parametric approach and propose an estimator for the bidders' risk aversion parameter. Their approach requires that the researcher imposes a parametric specification –such as CRRA or CARA– on the bidders' utility function before estimating the risk aversion parameter and the density of private values. In real-world applications, the choice of the parametric specification may be arbitrary and not always realistic. There is no general agreement on which specification is the right one.

The second objective is to estimate the latent density of private values following a fully nonparametric perspective. To that end, I propose an estimator for the density of private values that does not rely on any parametric specification of the bidders' utility function. A common practice when estimating the valuation density is to first assume a specific family of risk aversion for the bidders' utility, and then, estimate the latent density. The main advantage of this procedure is its low implementation cost. However, it can be criticized because an incorrect choice of the family of risk aversion undermines the asymptotic properties of the valuation density estimator and may weaken the finite-sample performance.

Many papers have developed nonparametric estimators for the density of private values under the assumption that bidders are risk-neutral. The pioneering work of Guerre, Perrigne, and Vuong (2000) established the optimal rate of convergence for estimating this density and constructed an estimator that attains this rate. Marmer and Shneyerov (2012) proposed an alternative estimator that is asymptotically normal

[1] See e.g. Delgado (2008) whose findings are consistent with a role for risk aversion in the tendency to bid too high.

and also attains the optimal rate. Bierens and Song (2012) used integrated simulated moments to propose an estimator and construct a test for the validity of the first-price auction model. Recently, Hickman and Hubbard (2014) have proposed a boundary correction that achieves uniform consistency on the entire support of the private value density and improves the finite sample performance. Here, I build on previous work by allowing bidders to be risk-averse. My estimator for the density of private values is uniformly consistent and its rate of convergence is established. I derive this property by extending the approach of Guerre et al. (2000) to accommodate risk aversion.

To my knowledge, only two papers have analyzed the identification of the bidders' utility function from a nonparametric perspective. Lu and Perrigne (2008) identified and estimated such a function by exploiting two auction designs, ascending and first-price sealed-bid auctions. Guerre, Perrigne, and Vuong (2009) improved on Lu and Perrigne (2008) and identified the bidders' utility function by using the latter design only. They showed that the bidders' utility function is nonparametrically identified under certain exclusion restriction. Exploiting this restriction, Guerre et al. (2009) developed their constructive identification strategy. Such a strategy is recursive and based on an infinite series of differences in quantiles, so it does not lead to a natural estimator for the bidders' utility function. Guerre et al. (2009) did not develop a formal estimator and instead discussed various extensions such as endogenous participation and asymmetries. My contribution here is to develop a uniformly consistent estimator with its rate of convergence.

This paper is related to a vast literature on empirical industrial organization. First, it relates to the literature on structural econometrics of auction data. This literature is extensive and has expanded at an extraordinary rate; for example, see the surveys of Hendricks and Paarsch (1995), Laffont (1997), Perrigne and Vuong (1999), Athey and Haile (2007), and Hendricks and Porter (2007), as well as the textbook of Paarsch, Hong, and Haley (2006). I remark that nonparametric approaches have become very popular as auction data has become more available. Second, this paper is also related to the literature on recovering risk preferences from observed behavior. Within this line of research, I highlight the working papers of Lu (2004) and Ackerberg, Hirano, and Shahriar (2011). The former proposes a semiparametric method to estimate the risk aversion parameter, as well as the risk premium, in the context of a first-price sealed-bid auction with stochastic private values. The latter considers a buy price auction framework and nonparametrically identifies both time and risk preferences of the bidders. In a recent article, Kim (2015) suggests a method to nonparametrically estimate the utility function in a first-price sealed-bid auction with risk-averse bidders and its performance is studied by Monte Carlo experiments.

The results obtained in this paper are useful for public policy recommendations. First-price sealed-bid auctions are used in many socio-economic contexts such as timber sales, outer continental shelf wildcat auctions (Li, Perrigne, and Vuong (2003)), and competitive sales of municipal bonds (Tang (2011)). To establish the optimal reserve price (i.e., the one that maximizes the auctioneer's revenue), we need valid estimators for the bidders' utility function and the density of private values. The estimators proposed here can be useful to construct the set of optimal reserve prices when bidders are risk-averse; Hu, Matthews, and Zou (2010) provide a characterization of this set.

The rest of the paper is organized as follows. Section 2 describes the auction model together with the data degenerating process, and also, establishes the parameter of interest. Section 3 develops a nonparametric estimator for the parameter of interest. Section 4 provides estimators for the bidders' utility function and the density of private values. Section 5 provides an implementation guide and reports the results of Monte Carlo experiments. Section 6 concludes with a discussion of possible extensions. Proofs of lemmas, propositions, and theorems are relegated to the Appendix.

2 Auction Model and Data Generating Process

This section is divided into two subsections. Subsection 2.1 describes briefly the benchmark model, which is standard in the auction literature: a first-price sealed-bid auction with risk-averse bidders, independent private values, and a non-binding reserve price.[2] Subsection 2.2 presents data generating process together with its assumptions and resulting properties.

2.1 First-Price Auction Model

A single indivisible object is sold through a first-price sealed-bid auction with non-binding reserve price. In other words, the object is sold to the highest bidder who pays his bid to the seller and each bidder does not know others' bids when forming his bid. Within the independent private values (IPV) paradigm, each bidder knows his own private value V, but not other bidders' private values. There are $I \geq 2$ bidders and private values are drawn independently from a common cumulative distribution function (c.d.f.) $F_V(\cdot)$, which is independent of I. Such a distribution is twice continuously differentiable with density $f_V(\cdot)$ and has compact support $[\underline{v}, \bar{v}] \subseteq \mathbb{R}_{\geq 0}$. Both the number of bidders I and $F_V(\cdot)$ are common knowledge.

All bidders are identical ex ante and the game is symmetric. Each bidder has the same univariate utility function $U(\cdot)$ that is independent of I. If a bidder with value V wins and pays $B \geq 0$, his utility is $U(B - V)$, and if he loses, his utility is $U(0)$. Since any bidder's payment must be smaller or equal than his own valuation, the domain of $U(\cdot)$ is restricted to $\mathbb{R}_{\geq 0}$. Each bidder maximizes his expected utility with respect to his own bid. It is assumed that $U(\cdot)$ is twice continuously differentiable with $U(0) = 0$, $U'(\cdot) > 0$, and $U''(\cdot) \leq 0$.

[2]For a detailed description of this auction model and additional results, see sec. 2 of Guerre et al. (2009) and the references cited therein.

Only symmetric Bayesian Nash equilibria are considered. As a consequence, there exists a unique symmetric equilibrium bidding function $s(\cdot;I)$. Such a function is strictly increasing, continuous on $[\underline{v},\bar{v}]$, and continuously differentiable on $(\underline{v},\bar{v}]$. Moreover, it satisfies the differential equation

$$s'(v;I) = (I-1)\frac{f_V(v)}{F_V(v)}\lambda\left[v - s(v;I)\right] \tag{1}$$

for all $v \in (\underline{v},\bar{v}]$ with boundary condition $s(\underline{v};I) = \underline{v}$, where $s(v;I)$ is the optimal bid for a valuation v, $s'(v;I) = \partial s(v;I)/\partial v$, and $\lambda(\cdot) \equiv U(\cdot)/U'(\cdot)$. From equation (1), the equilibrium bid for a valuation V is

$$B = s(V;I) = V - \lambda^{-1}\left[\frac{s'(V;I)F_V(V)}{(I-1)f_V(V)}\right], \tag{2}$$

where $\lambda^{-1}(\cdot)$ stands for the inverse of $\lambda(\cdot)$. Observe that the equilibrium bid B and, consequently, its quantile function depend on the numbers of bidders I despite the fact that $F_V(\cdot)$ and $U(\cdot)$ do not. The reason is that an increase in the number of bidders results in higher and more aggressive bidding. Exploiting variations in the quantiles of B from changes in I, while the corresponding quantiles of V remain the same, Guerre et al. (2009) showed that $\lambda^{-1}(\cdot)$ is identified.

2.2 Data Generating Process

In practice, the auctioned object can be heterogeneous, so here I introduce an additional random vector X to account for the heterogeneity in the auctioned object. The econometrician observes a random sample $\{(B_{pl}, I_l, X_l) : p = 1,\ldots,I_l, l = 1,\ldots,L\}$ where B_{pl} is the bid placed by the pth individual in the lth auction, I_l is the number of bidders in the lth auction, X_l is a D-dimensional vector of continuous auction-specific covariates, and L denotes the number of auctions in the sample. The private values of the bidders $\{V_{pl} : p = 1, \dots, I_l; l = 1, \dots, L\}$ are unobservable, and also, the identity of the bidders is unknown. From now on I suppose that the the data generating process satisfies the following assumption. Let S be a positive integer related to the smoothness of certain functions.

Assumption 1. *The random vectors* $\{(V_{pl}, I_l, X_l) : p = 1, \dots, I_l, l = 1, \dots, L\}$ *satisfy the following conditions.*

1. $\{(V_{1l}, \dots, V_{I_l l}, I_l, X_l) : l = 1, \dots, L\}$ *are independent.*

2. $\{(I_l, X_l) : l = 1, 2, \dots, L\}$ *are identically distributed with joint density* $f_{IX}(\cdot, \cdot)$*. Its support is* $\mathcal{I} \times \mathcal{X} \subset \mathbb{N}_{\geq 2} \times \mathbb{R}^D$*, where* $2 \leq \#(\mathcal{I}) < +\infty$ *and* $\mathcal{X} = \prod_{d=1}^{D} [\underline{x}_d, \bar{x}_d]$ *is a rectangular compact set with nonempty interior.*

3. *For each* $i \in \mathcal{I}$*,* $f_{IX}(i, \cdot)$ *admits* $S + 1$ *continuous bounded partial derivatives on* $\mathcal{X}$ *and is bounded away from* 0*.*

4. *For each* $l = 1, \dots, L$*,* $\{V_{pl} : p = 1, \dots, I_l\}$ *are independent and identically distributed conditionally on* (I_l, X_l)*. The conditional density of* V_{pl} *given* (I_l, X_l) *is independent of* I_l *and denoted by* $f_{V|X}(\cdot|\cdot)$*.*

Conditions 1-3 are standard in the literature on empirical auctions. I remark that the case $D = 0$ corresponds to the absence of covariates case. The fourth item is the key condition. It establishes that private values are independent and imposes an exclusion restriction on the bidders' participation. More precisely, the conditional density of private values must be conditionally independent of the number of bidders. As shown by Guerre et al. (2009), Proposition 2, an exclusion restriction is necessary to identify an auction model with risk-averse bidders.

In addition to Assumption 1, I impose standard regularity conditions on the latent conditional density $f_{V|X}(\cdot|\cdot)$. With this aim, I define the following set of conditional densities that contains $f_{V|X}(\cdot|\cdot)$.

Definition 1. *Let $\mathcal{F}_S^*$ be the set of conditional densities $f(\cdot|\cdot)$ satisfying the next conditions: $f(\cdot|\cdot)$ has support $\mathscr{S}_{VX} \equiv \{(v,x) : v \in [\underline{v}(x), \bar{v}(x)], x \in \mathcal{X}\}$ with $0 \leq \underline{v}(x) < \bar{v}(x) \leq \bar{C}_v$ for some constant $\bar{C}_v < +\infty$; $f(\cdot|\cdot)$ is bounded away from 0 on $\mathscr{S}_{VX}$; and $f(\cdot|\cdot)$ admits S continuous bounded partial derivatives on $\mathscr{S}_{VX}$.*

The next assumption establishes $f_{V|X}(\cdot|\cdot) \in \mathcal{F}_S^*$ and formalizes the idea that the bids are generated by a first-price auction within the IPV paradigm. It also establishes the smoothness of the bidders' utility function. Denote the conditional c.d.f. of valuations by $F_{V|X}(v|x) = \int_{\underline{v}(x)}^{v} f_{V|X}(t|x)dt$ with $\underline{v}(x) \leq v \leq \bar{v}(x)$.

Assumption 2. *The bids $\{B_{pl} : p = 1, \ldots, I_l, l = 1, \ldots, L\}$ are generated by the auction model of subsection 2.1 with $[U(\cdot), f_{V|X}(\cdot|\cdot)] \in \mathcal{U}_S \times \mathcal{F}_S^*$, where $U(\cdot)$ is a real function with domain $\mathbb{R}_{\geq 0}$ and $\mathcal{U}_S$ is defined in Guerre et al. (2009). To be specific, $B_{pl} = s(V_{pl}; I_l, X_l)$ where, for each $(i,x) \in \mathcal{I} \times \mathcal{X}$, $s(\cdot; i, x) : [\underline{v}(x), \bar{v}(x)] \to \mathbb{R}_{\geq 0}$ satisfies the differential equation*

$$s(v; i, x) = v - \lambda^{-1}\left[\frac{s'(v; i, x)F_{V|X}(v|x)}{(i-1)f_{V|X}(v|x)}\right]$$

for all $v \in (\underline{v}(x), \bar{v}(x)]$ with boundary condition $s[\underline{v}(x); i, x] = \underline{v}(x)$, $\lambda(\cdot) \equiv U(\cdot)/U'(\cdot)$, and $s'(v; i, x) = \partial s(v; i, x)/\partial v$.

Our parameter of interest is the function $\lambda^{-1}(\cdot)$. For instance, when the utility function exhibits CRRA with parameter $\eta \in [0, 1)$, $U(y) = y^{1-\eta}/(1-\eta)$, we have that $\lambda(y) = y/(1-\eta)$ and $\lambda^{-1}(u) = (1-\eta)u$. When the utility function exhibits CARA with parameter $\eta > 0$, $U(y) = [1 - \exp(-\eta y)]/[1 - \exp(-\eta)]$, we have $\lambda(y) = [\exp(\eta y) -$

$1]/\eta$ and $\lambda^{-1}(u) = \log(\eta u + 1)/\eta$. The distinguishing feature of this paper is that no parametric restrictions are imposed on $\lambda^{-1}(\cdot)$ beyond standard regularity assumptions. In particular, I do not assume that the bidder's utility belongs to a specific family of risk aversion –such as CRRA or CARA–.

Observe that $U(\cdot)$ can be recovered from $\lambda^{-1}(\cdot)$ as the closed-form solution of the differential equation $\lambda(\cdot)U'(\cdot) - U(\cdot) = 0$, with boundary condition $U(0) = 0$ and a normalizing restriction such as $U(1) = 1$ for some . Now denote the conditional c.d.f. of B_{pl} given (I_l, X_l) by $G(\cdot|\cdot,\cdot)$ and let $g(\cdot|\cdot,\cdot)$ be the corresponding density. Then, private values can be written as

$$V_{pl} = B_{pl} + \lambda^{-1}\left[\frac{1}{(I_l - 1)}\frac{G(B_{pl}|I_l, X_l)}{g(B_{pl}|I_l, X_l)}\right], \tag{3}$$

and consequently, we can recover the conditional p.d.f. $f_{V|X}(\cdot|\cdot)$ from $\lambda^{-1}(\cdot)$ and the conditional distribution of bids.

Note that $\lambda^{-1}(\cdot)$, as well as the utility function, does not depend on the number of bidders nor the auction-specific covariates. This assumption is standard in the literature; see e.g. Lu and Perrigne (2008) and Campo et al. (2011). Observe also that $\lambda(0) = 0$ and $\lambda'(\cdot) \geq 1$ because $U(0) = 0$, $U'(\cdot) > 0$, and $U''(\cdot) \leq 0$; consequently, the first derivative of $\lambda^{-1}(\cdot)$ satisfies $0 < \lambda^{-1\prime}(\cdot) \leq 1$.

Before proceeding, I mention some useful results that can be derived from Lemma 3 of Guerre et al. (2009) and Lemma 1 of Campo et al. (2011). First, the conditional density $g(\cdot|i,\cdot)$ is continuously differentiable on its support $\mathscr{S}_{BX}(i) \equiv \{(b,x) \in \mathbb{R}_{\geq 0} \times \mathbb{R}^D : b \in [\underline{b}(x), \bar{b}(i,x)], x \in \mathcal{X}\}$ for $i \in \mathcal{I}$, where $\underline{v}(x) = \underline{b}(x) < \bar{b}(i,x) \leq \bar{C}_B$ and $\bar{C}_B > 0$ is a finite constant. Second, it can be shown that $\inf\{\bar{b}(i,x) - \underline{b}(x) : x \in \mathcal{X}\} > 0$. Third, $g(\cdot|i,\cdot)$ is bounded away from 0 on $\mathscr{S}_{BX}(i)$; more specifically, there exists a constant $\underline{c}_g > 0$ such that $g(\cdot|i,\cdot) \geq \underline{c}_g$ on $\mathscr{S}_{BX}(i)$ for all $i \in \mathcal{I}$. Fourth, Observation

1 below summarizes additional useful results. Let $b(\cdot|i,x):[0,1]\to[\underline{b}(x),\bar{b}(i,x)]$ and $v(\cdot|x):[0,1]\to[\underline{v}(x),\bar{v}(x)]$ denote the (conditional) quantile functions of $G(\cdot|i,x)$ and $F_{V|X}(\cdot|x)$, respectively. Define the set $\mathcal{I}^*=\{(i_1,i_2)\in\mathcal{I}^2:i_1<i_2\}$ and the function $R(\alpha|i,x)=\alpha b'(\alpha|i,x)/(i-1)$ for $(\alpha,i,x)\in[0,1]\times\mathcal{I}\times\mathcal{X}$, being $b'(\alpha|i,x)$ the conditional quantile density function, i.e., $b'(\alpha|i,x)=\partial b(\alpha|i,x)/\partial\alpha$.

Observation 1. *Under Assumptions 1-2, the following statements hold.*

1. *For each $i\in\mathcal{I}$, the next properties are satisfied.*

   (a) *$G(\cdot|i,\cdot)$ admits $S+1$ continuous bounded partial derivatives on $\mathscr{S}_{BX}(i)$.*

   (b) *$g(\cdot|i,\cdot)$ admits $S+1$ continuous partial derivatives on $\{(b,x)\in\mathbb{R}_{\geq 0}\times\mathbb{R}^D : b\in(\underline{b}(x),\bar{b}(i,x)],x\in\mathcal{X}\}$.*

   (c) *$\lim_{b\downarrow\underline{b}(x)}\partial^s\left[\frac{G(b|i,x)}{g(b|i,x)}\right]/\partial^s b$ exists and is finite for $s=1,\dots,S+1$ and $x\in\mathcal{X}$.*

2. *For all $(\alpha,x)\in(0,1]\times\mathcal{X}$ and $(i_1,i_2)\in\mathcal{I}^*$, $b(\alpha|i_1,x)<b(\alpha|i_2,x)$. Moreover, $\underline{b}(0|i,x)$ does not depend on $i\in\mathcal{I}$.*

3. (a) *For all $(\alpha,x)\in[0,1]\times\mathcal{X}$ and $(i_1,i_2)\in\mathcal{I}^*$, the compatibility condition*

$$b(\alpha|i_2,x)-b(\alpha|i_1,x)=\lambda^{-1}[R(\alpha|i_1,x)]-\lambda^{-1}[R(\alpha|i_2,x)] \tag{4}$$

   is satisfied.

   (b) *For each $i\in\mathcal{I}$, the function $\xi_i:\mathscr{S}_{BX}(i)\to\mathbb{R}_{\geq 0}$ defined by*

$$\xi_i(b,x)=b+\lambda^{-1}\left[\frac{1}{(i-1)}\frac{G(b|i,x)}{g(b|i,x)}\right]$$

   satisfies $\partial\xi_i(b,x)/\partial b>0$ for every (b,x) and $v(\alpha|x)=\xi_i[b(\alpha|i,x),x]$.

This observation also follows directly by combining the results of Guerre et al. (2009) and Campo et al. (2011), hence its proof is omitted. Its innovation just consists in adding an auction covariate to Lemma 3 of Guerre et al. (2009). Part 1 states smoothness properties of the bids distribution and follows directly from Lemma 1 in Campo et al. (2011). Parts 2, 3, and 4 follow by adding an auction covariate to Lemma 3 in Guerre et al. (2009). As Part 2 does not depend on unknown functions, $\lambda^{-1}(\cdot)$ and $\xi_i(\cdot,\cdot)$, it provides a testable implication that can be verified with a first-order stochastic dominance test.

The next lemma constitutes the main contribution of this section as it establishes crucial properties of $R(\cdot|\cdot,\cdot)$. Write $R'(\alpha|i,x) = \partial R(\alpha|i,x)/\partial\alpha$.

Lemma 1. *Under Assumptions 1-2, there exist finite constants* $\underline{c}_R, \underline{c}'_R > 0$ *and* $0 < \tilde{\alpha}'' \leq \tilde{\alpha}' \leq \tilde{\alpha} \leq 1$ *such that*

1. $R'(\alpha|i,x) \geq \underline{c}_R$ *for all* $(\alpha,i,x) \in [0,\tilde{\alpha}] \times \mathcal{I} \times \mathcal{X}$;

2. $R'(\alpha|i_1,x) - R'(\alpha|i_2,x) \geq \underline{c}'_R$ *for all* $(\alpha,x) \in [0,\tilde{\alpha}'] \times \mathcal{X}$ *and* $(i_1,i_2) \in \mathcal{I}^*$;

3. *for every* $(i_1,i_2) \in \mathcal{I}^*$,

$$0 < \max_{x\in\mathcal{X}} \left[\frac{\max\{R'(\alpha|i_2,x) : \alpha \in [0,\tilde{\alpha}'']\}}{\min\{R'(\alpha|i_1,x) : \alpha \in [0,\tilde{\alpha}'']\}} \right] < 1.$$

Proof. See Appendix A.1. □

The function $R(\alpha|i,x)$ can be interpreted as the equilibrium markup associated with the α-quantile of $F_{V|X}(\cdot|x)$ when there are i bidders; specifically, $R(\alpha|i,x) = \lambda[v(\alpha|x) - b(\alpha|i,x)]$. The first item of Lemma 1 implies that this markup is strictly increasing for small values of α. The second establishes that the markup increases faster when the number of bidders decreases. The third implies that the ratio of the speeds at which markups increase is uniformly bounded on some neighborhood of 0. This last result is technical and will be used later in the proof of Lemma 2 to determine the effects of deviating from the compatibility condition.

3 Nonparametric Estimation

The purpose of this section is to build a nonparametric estimator of $\lambda^{-1}(\cdot)$, the parameter of interest. All asymptotic properties are established taking $L \to +\infty$, while the set $\mathcal{I}$ is fixed.

Consider the interval $[0, \bar{R}]$ as the domain of $\lambda^{-1}(\cdot)$, where $\bar{R} = \max_{\alpha \in [0,1]} R(\alpha|\underline{i}, x)$, $\underline{i} = \min\{\mathcal{I}\}$, and $x \in \text{interior}(\mathcal{X})$.[3] The value of x is fixed and chosen by the researcher. The parameter space is defined as follows. Denote the sup-norm of a function $\phi(\cdot)$ over a set $\mathcal{Z}$ by $\|\phi\|_{\mathcal{Z},\infty} = \sup_{z \in \mathcal{Z}} |\phi(z)|$ and, when ϕ is a real function, write its sth derivative by $\phi^{(s)}$ whenever exists.

Definition 2. *Let $\mathscr{H}_S$ be the space of functions $\phi : [0, \bar{R}] \to \mathbb{R}_{\geq 0}$ that satisfy the next conditions: $\phi(0) = 0$, $\phi(\cdot)$ admits $S + 1$ continuous derivatives on $[0, \bar{R}]$, $0 < \phi'(\cdot) \leq 1$, and $\|\phi^{(s)}\|_{[0,\bar{R}],\infty} \leq \bar{C}_{\mathscr{H}}$ for $s = 2, \ldots, S + 1$ and some large constant $\bar{C}_{\mathscr{H}} > 0$.*

As a measure of distance between a function $\phi(\cdot)$ and $\lambda^{-1}(\cdot)$, I consider the sup-norm over the interval $[0, \bar{u}]$, where $\bar{u} \in (0, \bar{R})$ is fixed and arbitrarily close to $\bar{R}$. The constant $\bar{C}_{\mathscr{H}}$ is taken to be large enough so that $\lambda^{-1} : [0, \bar{R}] \to \mathbb{R}_{\geq 0}$, belongs to $\mathscr{H}_S$. This is a technical requirement that can be ignored for practical purposes, i.e., working with real-world data.

[3] By Proposition 3 in Guerre et al. (2009), $\lambda^{-1}(\cdot)$ is nonparametrically identified on $[0, \bar{R}]$ and it can be shown that $\lambda^{-1}(u)$ is not identified when $u > \bar{R}$. The reason is that $\lambda^{-1}(\bar{R}) = \max\{v - s(v; \underline{i}, x) : v \in [\underline{v}(x), \bar{v}(x)]\}$, where $v - s(v; \underline{i}, x)$ represents the monetary gain of bidding $s(v; \underline{i}, x)$. Intuitively, the identification region $[0, \bar{R}]$ cannot be improved because bidders cannot obtain a monetary gain greater than its maximum, $\lambda^{-1}(\bar{R})$.

3.1 Population Criterion Function

In this subsection, I propose a population criterion function that will allow us to build a consistent estimator of $\lambda^{-1}(\cdot)$ and obtain its convergence rate.

As a starting point, define the functional $Q_\varepsilon(\cdot|i) : \mathscr{H}_S \to \mathbb{R}_{\geq 0}$ as

$$Q_\varepsilon(\phi|i) = \max_{\alpha\in[\varepsilon,1-\varepsilon]} \left|b(\alpha|i,x) - b(\alpha|\underline{i},x) + \phi[R(\alpha|i,x)] - \phi[R(\alpha|\underline{i},x)]\right|,$$

where $\varepsilon \in [0,1)$, $i \in \mathcal{I}\backslash\{\underline{i}\}$, and $x \in \text{interior}(\mathcal{X})$ is fixed. The functional form of $Q_\varepsilon(\cdot|i)$ is based on the compatibility condition (4). Since $\mathcal{I}$ may contain more than two elements, I define the population criterion function $Q_\varepsilon : \mathscr{H}_S \to \mathbb{R}_{\geq 0}$ as

$$\begin{aligned} Q_\varepsilon(\phi) &= \sum_{i\in\mathcal{I}\backslash\{\underline{i}\}} Q_\varepsilon(\phi|i) \\ &= \sum_{i\in\mathcal{I}\backslash\{\underline{i}\}} \max_{\alpha\in[\varepsilon,1-\varepsilon]} \left|b(\alpha|i,x) - b(\alpha|\underline{i},x) + \phi[R(\alpha|i,x)] - \phi[R(\alpha|\underline{i},x)]\right|. \end{aligned}$$

Note that $Q_\varepsilon(\lambda^{-1}) = 0$ for any $\varepsilon \in [0,1)$ due to eq. (4). Furthermore, the criterion function satisfies the following property.

Lemma 2. *Suppose that Assumptions 1-2 hold. There exists a constant* $\underline{c}_Q > 0$ *so that the next implication holds: for every* $\phi \in \mathscr{H}_S$ *and* $\varepsilon > 0$ *sufficiently small,*

$$\|\phi(\cdot) - \lambda^{-1}(\cdot)\|_{[0,\bar{u}],\infty} \geq \varepsilon \;\Rightarrow\; Q_\varepsilon(\phi) \geq \underline{c}_Q \frac{\varepsilon}{\log(\varepsilon^{-1})}. \tag{5}$$

Proof. See Appendix A.2. □

This lemma is essential to construct a valid estimator for $\lambda^{-1}(\cdot)$ with its rate of convergence. We already know that $\lambda^{-1}(\cdot)$ is the unique function satisfying $Q_0(\lambda^{-1}) = 0$ due to Guerre et al. (2009)'s identification result, which proves that there exists only one function that satisfies eq. (4). Lemma 2 establishes that $Q_\varepsilon(\cdot)$ is bounded below, over a neighborhood of $\lambda^{-1}(\cdot)$, by certain function in the distance from $\lambda^{-1}(\cdot)$. This result is crucial to derive the convergence rate as it parallels, e.g., Condition 3.1 in Chen (2007) and Condition C.2 in Chernozhukov, Hong, and Tammer (2007). An implication of Lemma 2 is that the parameter of interest can be approximated by any sequence of functions $(\phi_L)_{L\in\mathbb{N}} \subset \mathscr{H}_S$ satisfying $Q_\varepsilon(\phi_L) \to 0$ as $L \to +\infty$ and $\varepsilon \to 0$, i.e., $Q_\varepsilon(\phi_L) \to 0$ implies $\phi_L \to \lambda^{-1}$. Intuitively, this feature helps us develop an estimator for $\lambda^{-1}(\cdot)$ as estimation involves approximating a parameter using sample analogues.

Before proceeding to the next subsection, I highlight that there are different alternatives to the shape of the criterion function $Q_\varepsilon(\cdot)$. For instance, one can use different weights for $Q_\varepsilon(\cdot|i)$. Note also that $Q_\varepsilon(\cdot)$ employs the sup-norm to penalize deviations from the compatibility condition. So another alternative would be to use an L_p-norm with $1 \leq p < +\infty$ to penalize these deviations. However, in such a case the resulting rate of convergence would be slower due to the inequalities between norms; see e.g. Theorem 1 in Gabushin (1967).

3.2 Preliminary Nonparametric Estimators

This subsection constructs nonparametric estimators for the quantile function $b(\cdot|i,x)$ and its derivative, $b'(\cdot|i,x)$. Since $R(\alpha|i,x) = \alpha b'(\alpha|i,x)/(i-1)$, $R(\cdot|i,x)$ will be estimated by plugging in the estimator for $b'(\cdot|i,x)$. The proposed estimators for $b(\cdot|i,x)$ and $R(\cdot|i,x)$ will be used later to compute the empirical counterpart of $Q_\varepsilon(\cdot)$.

Let $k(\cdot)$ be a univariate kernel, h_X and h_μ bandwidths, $(J_L)_{L\in\mathbb{N}}$ an increasing sequence of positive integers, and h_ε a positive sequence that converges to zero as $L \to +\infty$. These objects will be employed throughout this subsection to construct preliminary nonparametric estimators. I make the following assumptions about them. Denote the ceiling function by $\lceil y \rceil = \min\{n \in \mathbb{Z} : n \geq y\}$ for $y \in \mathbb{R}$.

Assumption 3. *The kernel $k(\cdot)$ is symmetric with $S+1$ continuous derivatives on $\mathbb{R}$, has support $[-1,1]$, and satisfies $\int k(v)dv = 1$. The order of $k(\cdot)$ is $S+1$, i.e., $\int v^s k(v)dv = 0$ for $s = 1,\ldots,S$ and $0 < \int v^{S+1}k(v)dv < +\infty$.*

Assumption 4. *Let γ_X, γ_μ, γ_J, and γ_ε be positive constants. The bandwidths h_X and h_μ are given by $h_X = \gamma_X[\log(L)/L]^{1/(2S+D+2)}$ and $h_\mu = \gamma_\mu L^{-1/(2S+D+2)}$, respectively. The sequence $(J_L)_{L\in\mathbb{N}}$ is of the form*

$$J_L = \left\lceil \gamma_J L^{\frac{2(S+1)}{(2S+3)(2S+D+2)}} \right\rceil$$

and h_ε satisfies $h_\varepsilon = \gamma_\varepsilon L^{-(S+1)/(2S+D+3)}$.

Following closely Marmer and Shneyerov (2012), we employ a kernel approach to estimate $f_{IX}(\cdot,\cdot)$ and $G(\cdot|\cdot,\cdot)$:

$$\begin{aligned} \hat{f}_{IX}(i,x) &= \frac{1}{Lh_X^D}\sum_{l=1}^{L}\mathbb{1}\{I_l = i\}K\left(\frac{x-X_l}{h_X}\right) \text{ and} \\ \hat{G}(b|i,x) &= \frac{1}{\hat{f}_{IX}(i,x)Lh_X^D}\sum_{l=1}^{L}\frac{1}{I_l}\sum_{p=1}^{I_l}\mathbb{1}\{B_{lp}\le b, I_l = i\}K\left(\frac{x-X_l}{h_X}\right), \end{aligned} \tag{6}$$

respectively, where $(b,i,x)\in\mathbb{R}_{\ge 0}\times\mathcal{I}\times\mathbb{R}^D$ and $K(\cdot)$ is the product kernel, i.e., $K(x) = \prod_{j=1}^{D}k(x_j)$. The conditional quantile function $b(\alpha|i,x)$ is estimated by

$$\hat{b}(\alpha|i,x) = \inf\{b\ge 0 : \hat{G}(b|i,x)\ge\alpha\} \tag{7}$$

for $\alpha\in(0,1)$. The order of $k(\cdot)$ and the form of the bandwidth h_X have been chosen according to the smoothness of $f_{IX}(\cdot,\cdot)$ and $G(\cdot|\cdot,\cdot)$. Such choices allow the estimators $\hat{f}_{IX}(i,x)$ and $\hat{G}(\cdot|i,x)$ to attain the fastest possible convergence rates.

Due to the shape of the criterion function, we need a uniformly consistent estimator for $b'(\cdot|i,x) = 1/g[b(\cdot|i,x)|i,x]$ over the interval $[\varepsilon, 1-\varepsilon]$, where $\varepsilon > 0$ approaches zero in a suitable manner as $L \to +\infty$; we will set $\varepsilon = h_\varepsilon$ in the next subsection. To build such an estimator, it suffices to construct a uniformly consistent estimator for $g(\cdot|i,x)$ over $[\underline{b}(x), \bar{b}(i,x)]$. Estimating $g(\cdot|i,x)$ by kernel methods would generate a boundary bias problem because h_ε will converge to zero faster than the optimal bandwidth (Assumption 4). In contrast, a series approach would not suffer from this problem.

The conditional density of equilibrium bids, $g(\cdot|i,x)$, is estimated by extending Barron and Sheu (1991)'s approach to conditional densities. To simplify the discussion, this section assumes that the boundaries $\underline{b}(x)$ and $\bar{b}(i,x)$ are known for $(i,x) \in \mathcal{I} \times \mathcal{X}$; subsection 5.1 below suggests estimators for $\underline{b}(x)$ and $\bar{b}(i,x)$. Let $g^*(\cdot|i,x)$ be the conditional density of

$$B^*_{lp} \equiv \frac{B_{lp} - \underline{b}(X_l)}{\bar{b}(I_l, X_l) - \underline{b}(X_l)}$$

given $(I_l, X_l) = (i,x)$. Note that the support of $g^*(\cdot|i,x)$ is $[0,1]$ and

$$g(b|i,x) = \left[\bar{b}(i,x) - \underline{b}(x)\right]^{-1} g^* \left(\frac{b - \underline{b}(x)}{\bar{b}(i,x) - \underline{b}(x)} \middle| i,x \right)$$

by the transformation formula. Then, the conditional density $g(\cdot|i,x)$ is estimated by

$$\hat{g}(b|i,x) = \left[\bar{b}(i,x) - \underline{b}(x)\right]^{-1} \hat{g}^* \left(\frac{b - \underline{b}(x)}{\bar{b}(i,x) - \underline{b}(x)} \middle| i,x \right),$$

being $\hat{g}^*(\cdot,\cdot|i)$ a conditional exponential series estimator of $g^*(\cdot|i,x)$. To be specific,

$$\hat{g}(b^*|i,x) = \frac{\exp\left[\sum_{1 \le j \le J_L} \hat{\theta}_j(i,x)\pi_j\left(b^*\right)\right]}{\int_0^1 \exp\left[\sum_{1 \le j \le J_L} \hat{\theta}_j(i,x)\pi_j(y)\right] dy},$$

where $b^* \in [0,1]$ and $\{\pi_j(\cdot) : 1 \le j \le J_L\}$ are the orthonormal Legendre polynomials on

$[0,1]$ with respect to the Lebesgue measure:

$$\pi_j(y) = (-1)^j \sqrt{2j+1} \sum_{\tau=0}^{j} \binom{j}{\tau} \binom{j+\tau}{j} (-y)^\tau$$

with $y \in [0,1]$. The coefficients $\{\hat{\theta}_j(i,x) : 1 \le j \le J_L\}$ are obtained by solving the following system of nonlinear equations:

$$\frac{\int_0^1 \pi_j(y) \exp\left[\sum_{1 \le j \le J_L} \hat{\theta}_j(i,x) \pi_j(y)\right] dy}{\int_0^1 \exp\left[\sum_{1 \le j \le J_L} \hat{\theta}_j(i,x) \pi_j(y)\right] dy} = \hat{\mu}_j(i,x) \tag{8}$$

for $j = 1, 2, \ldots, J_L$, where

$$\hat{\mu}_j(i,x) = \sum_{l=1}^{L} \left[\frac{1}{I_l} \sum_{p=1}^{I_l} \pi_j\left(B^*_{lp}\right) \right] \omega_l(i,x)$$

and the weights are

$$\omega_l(i,x) = \frac{\mathbb{1}\{I_l = i\} K\left(\frac{x - X_l}{h_\mu}\right)}{\sum_{m=1}^{L} \mathbb{1}\{I_m = i\} K\left(\frac{x - X_m}{h_\mu}\right)}.$$

Note that $\hat{\mu}_j(i,x)$ estimates the conditional expectation of $\pi_j\left(B^*_{lp}\right)$ given $(I_l, X_l) = (i,x)$, which is denoted by $\mu_j(i,x) \equiv E\left[\pi_j\left(B^*_{lp}\right) \middle| I_l = i, X_l = x\right]$. I remark that $\hat{g}^*(\cdot|\cdot)$ coincides with Barron and Sheu (1991)'s estimator when $D = 0$.

From the precedent discussion, the quantile density function $b'(\cdot|i,x)$ is estimated by $\hat{b}'(\alpha|i,x) = 1/\hat{g}[\hat{b}(\alpha|i,x)|i,x]$. Naturally, $\hat{R}(\alpha|i,x) = \alpha\hat{b}'(\alpha|i,x)/(i-1)$ becomes the estimator of $R(\alpha|i,x)$. The asymptotic properties of $\hat{b}(\cdot|i,x)$, $\hat{g}(\cdot|i,x)$, and $\hat{R}(\cdot|i,x)$ are stated in the next lemma. We abbreviate 'with probability approaching 1' and write 'w.p.a.1' instead.

Lemma 3. *Under Assumptions 1-4, the following statements hold for any* $(i,x) \in \mathcal{I} \times \text{interior}(\mathcal{X})$.

1. $\|\hat{b}(\cdot|i,x) - b(\cdot|i,x)\|_{[h_\varepsilon,1-h_\varepsilon],\infty} = O_P\left(\left[\frac{\log(L)}{L}\right]^{\frac{S+1}{2S+D+2}}\right)$.

2. *There exists a unique solution* $\{\hat{\theta}_j(i,x) : 1 \le j \le J_L\}$ *to eqs. (8) w.p.a.1 and*

$$\|\hat{g}(\cdot|i,x) - g(\cdot|i,x)\|_{[\underline{b}(x),\bar{b}(i,x)],\infty} = O_P\left(L^{-\frac{2S(S+1)}{(2S+3)(2S+D+2)}}\right).$$

3. $\|\hat{R}(\cdot|i,x) - R(\cdot|i,x)\|_{[h_\varepsilon,1-h_\varepsilon],\infty} = O_P\left(L^{-\frac{2S(S+1)}{(2S+3)(2S+D+2)}}\right)$.

Proof. See Appendix A.3. □

The first part of this lemma is similar to Lemma 1.(d) of Marmer and Shneyerov (2012). The difference is that here the support of the sup-norm is expanding, while in Marmer and Shneyerov (2012) is fixed. The second part, the main contribution of Lemma 3, establishes existence and uniform consistency of $\hat{g}(\cdot|i,x)$. In case of having $D = 0$, the obtained rate can be improved by using Theorem 7 in Wu (2010); see also subsection 3.3.1 for a discussion about improving the convergence rate when $D \in \{0,1\}$. The third part states that $\hat{R}(\cdot|i,x)$ inherits the rate of $\hat{g}(\cdot|i,x)$. This result follows by exploiting certain inequalities derived from the shape of $R(\cdot|i,x)$.

3.3 The Estimator: Definition and Uniform Consistency

From the discussion of subsection 3.1, the parameter of interest can be characterized as the unique argument that minimizes $Q_0(\cdot)$ over $\mathscr{H}_S$:

$$\lambda^{-1}(\cdot) = \arg\min_{\phi\in\mathscr{H}_S} Q_0(\phi).$$

Given this characterization and Lemmas 2-3, in this subsection, I build the estimator of $\lambda^{-1}(\cdot)$ as the argument that minimizes the empirical counterpart of $Q_\varepsilon(\cdot)$ over a sieve space (a finite-dimensional approximation space) with $\varepsilon \to 0$.

First, we construct the empirical criterion function $\hat{Q}(\cdot)$, which is the empirical counterpart of $Q_\varepsilon(\cdot)$ after replacing ε by the sequence h_ε of Assumption 4:

$$\hat{Q}(\phi) = \sum_{i \in \mathcal{I} \setminus \{\underline{i}\}} \max_{\alpha \in [h_\varepsilon, 1-h_\varepsilon]} \left| \hat{b}(\alpha|i,x) - \hat{b}(\alpha|\underline{i},x) + \phi[\hat{R}(\alpha|i,x)] - \phi[\hat{R}(\alpha|\underline{i},x)] \right|$$

with $\phi \in \mathscr{H}_S$. Lemma 4 below states that $\hat{Q}(\cdot)$ converges uniformly in probability to its population counterpart. More specifically, $\hat{Q}(\cdot)$ inherits the rate of convergence associated with the slowest term, $\hat{g}(\cdot|i,x)$. The proof of this lemma exploits the restriction $0 < \phi'(\cdot) \leq 1$ and combines it with Lemma 3.

Lemma 4. *Under Assumptions 1-4,*

$$\sup_{\phi \in \mathscr{H}_S} |\hat{Q}(\phi) - Q_{h_\varepsilon}(\phi)| = O_P\left(L^{-\frac{2S(S+1)}{(2S+3)(2S+D+2)}} \right).$$

Proof. See Appendix A.4. □

Second, we consider a sequence of sieve spaces $\{\mathscr{H}^{(L)} : L \in \mathbb{N}\}$ to approximate $\mathscr{H}_S$ and impose the following assumption.

Assumption 5. $\{\mathscr{H}^{(L)} : L \in \mathbb{N}\}$ *is a sequence of finite-dimensional approximation spaces that satisfy* $\mathscr{H}^{(L)} \subseteq \mathscr{H}_S$ *for every* L *sufficiently large and*

$$\inf_{\phi \in \mathscr{H}^{(L)}} \left\| \phi - \tilde{\phi} \right\|_{[0,\bar{R}],\infty} = O\left(L^{-\frac{2S(S+1)}{(2S+3)(2S+D+2)}} \right) \tag{9}$$

for any $\tilde{\phi} \in \mathscr{H}_S$ *as* $L \to +\infty$.

Regarding the choice of the sieve basis, I suggest using Bernstein polynomials.[4]

[4]See ch. 10 of DeVore and Lorentz (1993) for a detailed discussion. Chen (2007) also suggests

Denote the Bernstein basis over the interval $[0, \bar{U}]$ by

$$P_{M,m}(u) = \binom{M}{m} \left(\frac{u}{\bar{U}}\right)^m \left(1 - \frac{u}{\bar{U}}\right)^{M-m},$$

where $u \in [0, \bar{U}]$, $\bar{U}$ is a constant (chosen by the researcher) such that $\bar{U} \geq \bar{R}$, and $M \geq m$ are positive integers. Let M_L be a sequence of positive integers such that

$$M_L = \left\lceil \gamma_M L^{\frac{2S(S+1)}{(2S+3)(2S+D+2)}} \right\rceil \tag{10}$$

for some constant $\gamma_M > 0$.[5] The sieve $\mathscr{H}^{(L)}$ can be taken to be the space of functions $\phi : [0, \bar{R}] \to \mathbb{R}_{\geq 0}$ of the form

$$\phi(u) = \sum_{m=1}^{M_L} \beta_m P_{M_L,m}(u),$$

where the coefficients $\{\beta_m : m = 1, \ldots, M_L\}$ satisfy the next conditions:

$$\bar{U} L^{-2} \leq \beta_{m+1} - \beta_m \leq \bar{U} M_L^{-1} \tag{11}$$

for $0 \leq m \leq M_L - 1$ with $\beta_0 \equiv 0$, and

$$\left| \sum_{\tau=0}^{S+1} (-1)^{S+1-\tau} \binom{S+1}{\tau} \beta_{m+\tau} \right| \leq \left(\frac{\bar{U}}{M_L} \right)^{S+1} \bar{C}_{\mathscr{H}} \tag{12}$$

for $0 \leq m \leq M_L - S - 1$. Note that M_L indicates the dimension of the sieve space. Conditions (11)-(12) guarantee that the proposed sieves satisfy Assumption 5. To be specific, condition (11) implies $0 < \phi'(\cdot) \leq 1$, while (12) implies $\|\phi^{(s)}\|_{[0,\bar{u}],\infty} \leq \bar{C}_{\mathscr{H}}$ for all

Cardinal B-splines wavelets as shape-preserving sieves. As noted by Matzkin (1994), shape-preserving estimators have several advantages, e.g., they decrease the variance and improve the quality of an extrapolation beyond the support of the data.

[5]Suggestions for $\bar{U}$ and γ_M are given in subsection 5.1 below.

$\phi \in \mathscr{H}^{(L)}$ and $s = 2, \ldots, S+1$. Since $\phi(0) = 0$ by construction, $\mathscr{H}^{(L)} \subseteq \mathscr{H}_S$. Requirement (9) is satisfied by the form of M_L in eq. (10); see Theorem 3.1 in ch. 10 of DeVore and Lorentz (1993).

Third, we define the estimator of $\lambda^{-1}(\cdot)$ as

$$\hat{\lambda}^{-1}(\cdot) = \arg\min_{\phi \in \mathscr{H}^{(L)}} \hat{Q}(\phi). \tag{13}$$

Computational aspects are discussed in subsection 5.1 below. Theorem 1 below establishes the uniform consistency of $\hat{\lambda}^{-1}(\cdot)$ with its rate of convergence. Its proof uses arguments similar to that of Theorem 3.1 in Chen (2007). Let $\varphi^{-1}(\cdot)$ be the inverse of the function $\varphi(x) \equiv x\log(x)$ and denote the convergence rate by $r_L^* = \varphi^{-1}\left(L^{\frac{2S(S+1)}{(2S+3)(2S+D+2)}}\right)$.

Theorem 1. *Under Assumptions 1-5,* $r_L^* \left\|\hat{\lambda}^{-1}(\cdot) - \lambda^{-1}(\cdot)\right\|_{[0,\bar{u}],\infty} = O_P(1)$.

Proof. See Appendix A.5. □

Regarding existing literature results, in the absence of covariates ($D = 0$), Guerre et al. (2009) characterized $\lambda^{-1}(u)$ as an infinite series of differences in quantiles:

$$\lambda^{-1}(u) = \sum_{t=0}^{+\infty}[b(a_t|i) - b(a_t|\underline{i})] \tag{14}$$

where $u \in (0, \bar{u}]$, $i \in \mathcal{I}\backslash\{\underline{i}\}$, and $(a_t)_{t\in\mathbb{N}} \subseteq (0,1)$ is a strictly decreasing sequence that satisfies the nonlinear recursive relation $R(a_t|\underline{i}) = R(a_{t-1}|i)$ with initial condition $R(a_0|\underline{i}) = u$. Since $R(\cdot|\underline{i})$ is not necessarily increasing outside $[0,\tilde{\alpha}]$ (see Lemma 1.1), the sequence $(a_t)_{t\in\mathbb{N}}$ is not necessarily unique. At this point, it is not known whether expression (14) can lead to a consistent estimator of $\lambda^{-1}(u)$.[6] Obtaining the asymptotic properties of an estimator based on (14) would have several difficulties.

[6]Guerre et al. (2009), pp. 1216, state that "a nonparametric estimation for $[U, F_{V|X}]$ or, equivalently, $[\lambda^{-1}, F_{V|X}]$ clearly needs to be developed" but do not develop a formal estimator. Then, they discuss possible estimation strategies with their difficulties.

First, expression (14) does not provide a rate at which $\lambda^{-1}(u)$ can be estimated. Even if the empirical counterpart of each summand converges at a given rate, the summation may not converge as there are infinitely many terms. Second, since there is no "polynomial minorant" for $R(\cdot|\underline{i})$, we cannot establish the rate at which $\tilde{a}_0$ and, consequently, $\lambda^{-1}(u))$ can be estimated; see e.g., Condition C.2 and Theorem 3.1 in Chernozhukov et al. (2007). Third, an estimator based on (14) would have the problem of accumulating estimation errors because $(a_t)_{t\in\mathbb{N}}$ is recursively defined.

More recently, Kim (2015) proposed a method to nonparametrically estimate $\lambda^{-1}(\cdot)$ when there are no covariates. He characterized $\lambda^{-1}(\cdot)$ as the unique fixed point of a mapping $\mathcal{T}:\mathscr{H}_1 \to \mathscr{H}_1$ implicitly defined by the transformation

$$\mathcal{T}(\phi)[R(\alpha|\underline{i})] = b(\alpha|i) - b(\alpha|\underline{i}) + \phi[R(\alpha|i)]$$

for $\alpha \in [0,1]$, and also he established that there exists a metric over $\mathscr{H}_1$ for which $\mathcal{T}(\cdot)$ is a contraction. Then, $\lambda^{-1}(\cdot)$ is estimated by replacing $[b(\cdot|\cdot), R(\cdot|\cdot)]$ with their sample analogues and iterating the contraction operator $\mathcal{T}(\cdot)$. The performance of this method is studied by Monte Carlo simulations and, at this point, the asymptotic properties have not been established. I remark two differences between Kim (2015)'s approach and mine. First, the former characterizes $\lambda^{-1}(\cdot)$ as a fixed point of a contraction mapping ($\lambda^{-1} = \mathcal{T}(\lambda^{-1})$), whereas the latter represents $\lambda^{-1}(\cdot)$ as the argument that minimizes a criterion function. Second, Kim (2015)'s estimation method is iterative, while the estimation procedure proposed here involves minimizing the empirical counterpart of a criterion function (to avoid accumulating estimation errors).

3.3.1 Discussion: Improving Theorem 1's Rate of Convergence

This subsection considers the case $D \in \{0,1\}$ and discusses how Theorem 1's convergence rate can be improved.

When there is no covariates, $D = 0$, the conditional bids density $g(\cdot|i)$ can be estimated following Stone (1990)'s log-spline approach. In such a case, it can be shown that

$$\sup_{\phi \in \mathscr{H}_S} |\hat{Q}(\phi) - Q_{h_\varepsilon}(\phi)| = O_P\left(\left[\frac{\log(L)}{L}\right]^{\frac{S+1}{2S+3}}\right).$$

The reason is that $\hat{Q}(\phi)$ inherits the rate of convergence associated with the estimator of $g(\cdot|i)$ and Stone (1990)'s estimator achieves the optimal rate of uniform convergence, $[L/\log(L)]^{(S+1)/(2S+3)}$. Similarly, when $D = 1$, the conditional density of bids $g(\cdot|i,x)$ can estimated following Stone (1991)'s procedure and we can show that

$$\sup_{\phi \in \mathscr{H}_S} |\hat{Q}(\phi) - Q_{h_\varepsilon}(\phi)| = O_P\left(\left[\frac{\log(L)}{L}\right]^{\frac{S+1}{2S+4}}\right).$$

The next observation establishes the rate of convergence of $\hat{\lambda}^{-1}(\cdot)$ when $D \in \{0,1\}$ and the conditional density of bids is estimated as suggested above.

Observation 2. *Suppose that Assumptions 1-4 hold, $D \in \{0,1\}$, and the sequence of sieves $\{\mathscr{H}^{(L)} \subset \mathscr{H}_S : L \in \mathbb{N}\}$ satisfies*

$$\inf_{\phi \in \mathscr{H}^{(L)}} \|\phi - \tilde{\phi}\|_{[0,\bar{R}],\infty} = O\left(\left[\frac{\log(L)}{L}\right]^{\frac{S+1}{2R+D+3}}\right)$$

for any $\tilde{\phi} \in \mathscr{H}_S$. Using an appropriate estimator for the conditional density of bids, i.e.

Stone (1990) or Stone (1991), we obtain

$$\varphi^{-1}\left\{\left[\frac{L}{\log(L)}\right]^{\frac{S+1}{2S+D+3}}\right\}\left\|\hat{\lambda}^{-1}(\cdot)-\lambda^{-1}(\cdot)\right\|_{[0,\bar{u}],\infty}=O_P(1).$$

The proof of this observation is omitted as it follows immediately from the proof of Theorem 1 and the precedent discussion. Note that $\varphi^{-1}(x) < x$ when $x > e$, so the obtained rate of convergence is slower than $[L/\log(L)]^{(S+1)/(2R+D+3)}$, which is Stone (1982)'s optimal uniform rate when there are D covariates. But for any fixed $c \in (0,1)$, we still have that $\varphi^{-1}(x) > x^c$ whenever $x > 0$ is sufficiently large. So when $D \in \{0,1\}$, up to a exponent, Theorem 1's rate can be arbitrarily close to Stone (1982)'s rate. Formally, for any fixed $c \in (0,1)$ and $D \in \{0,1\}$, we have that

$$\left[\frac{L}{\log(L)}\right]^{\frac{c(S+1)}{2S+D+3}}\left\|\hat{\lambda}^{-1}(\cdot)-\lambda^{-1}(\cdot)\right\|_{[0,\bar{u}],\infty}=o_P(1).$$

To extend this result to the general case $D > 1$, we require an estimator for $g(\cdot|i,x)$ that attains a uniform rate of $[L/\log(L)]^{(S+1)/(2S+D+3)}$ over $[\underline{b}(x),\bar{b}(x,i)]$. To my knowledge, such an estimator has not been developed yet and it is not known whether it is possible.

4 Estimating the First-Price Auction Model

The previous section developed an estimator for $\lambda^{-1}(\cdot)$, this section applies it to the auction model of Section 2. Exploiting the uniform consistency of $\hat{\lambda}^{-1}(\cdot)$, I propose estimators for the bidders' utility function and density of private values.

To estimate the utility function $U(\cdot)$, pick any $\bar{y} \in (0,\lambda^{-1}(\bar{u}))$ with $0 < \bar{u} < \bar{R}$.[7] Consider the normalizing restrictions $U(0)=0$ and $U(\bar{y})=1$. Then, $U(\cdot)$ becomes the

[7] Since $\bar{R}$ and $\lambda^{-1}(\bar{u})$ are unknown but can be estimated from data, next section suggests how to choose their values.

unique solution of the differential equation $\lambda(\cdot)U'(\cdot) - U(\cdot) = 0$ on $[0, \bar{y}]$; specifically, $U(y) = \exp\{-\int_y^{\bar{y}}[1/\lambda(t)]dt\}$ for $y \in [0, \bar{y}]$. As $\hat{\lambda}^{-1}(\cdot)$ is strictly increasing on $[0, \bar{u}]$, the proposed estimator for $\lambda(\cdot)$ is simply the inverse of $\hat{\lambda}^{-1}(\cdot)$, i.e., $\hat{\lambda}(y) = (\hat{\lambda}^{-1})^{-1}(y)$. Then, $U(y)$ is estimated by

$$\hat{U}(y) = \begin{cases} 0 & \text{if } y = 0, \\ \exp\left\{-\int_y^{\bar{y}}[1/\hat{\lambda}(t)]dt\right\} & \text{if } 0 < y < \bar{y}, \\ 1 & \text{if } y \geq \bar{y}. \end{cases} \tag{15}$$

Before proceeding, I remark that $\hat{\lambda}(\cdot)$ is well-defined and $\hat{\lambda}'(\cdot) \geq 1$ by $0 < \hat{\lambda}^{-1\prime}(\cdot) \leq 1$ on $[0, \bar{u}]$. Observe that $\hat{U}(\cdot)$ is continuous because $\hat{\lambda}(\cdot)$ is continuous, $\hat{\lambda}(0) = 0$, and $\int_y^1[1/\hat{\lambda}(t)]dt \to +\infty$ as $y \to 0^+$. Moreover, $\hat{U}(\cdot)$ is a shape-preserving estimator in the sense that it is strictly increasing and concave on $[0, \bar{y}]$ regardless of the sample size; note that $\hat{U}'(y) = \hat{U}(y)/\hat{\lambda}(y) > 0$ for $y \in (0, \bar{y})$.

To estimate the conditional density of private values $f_{V|X}(\cdot|\cdot)$, I extend Guerre et al. (2000)'s approach to accommodate for risk aversion. In view of eq. (3), I construct the pseudo private values

$$\hat{V}_{pl} = B_{pl} + \hat{\lambda}^{-1}\left[\frac{\hat{\psi}\,(B_{lp}, I_l, X_l)}{I_l - 1}\right], \tag{16}$$

where $\hat{\psi}(\cdot,\cdot,\cdot)$ is defined in Guerre et al. (2000), eq. (19). Then, the latent density $f_{V|X}(v|x)$ is estimated by $\hat{f}_{V|X}(v|x) = \hat{f}_{VX}(v,x)/\hat{f}_X(x)$, where

$$\hat{f}_{VX}(v,x) = \frac{1}{Lh_f^{D+1}}\sum_{l=1}^{L}\frac{1}{I_l}\sum_{p=1}^{I_l} k\left(\frac{v - \hat{V}_{pl}}{h_f}\right) K\left(\frac{x - X_l}{h_f}\right), \tag{17}$$

$\hat{f}_X(x) = \sum_{i\in\mathcal{I}}\hat{f}_I(i)\hat{f}_{IX}(i,x)$, $\hat{f}_I(i) = (1/L)\sum_{l=1}^{L}\mathbb{1}\{I_l = i\}$, and h_f is a bandwidth.

The next proposition states the asymptotic properties of the proposed estimators.

Proposition 1. *Under Assumptions 1-5, the following statements hold.*

1. $r_L^* \|\hat{\lambda}(\cdot) - \lambda(\cdot)\|_{[0,\bar{y}],\infty} = O_P(1)$ *and* $r_L^* \|\hat{U}(\cdot) - U(\cdot)\|_{[0,\bar{y}],\infty} = O_P(1)$.

2. *If* $h_f \to 0$ *and* $h_f r_L^* \to +\infty$, *then*

$$\|\hat{f}_{V|X}(\cdot|\cdot) - f_{V|X}(\cdot|\cdot)\|_{\mathcal{C},\infty} = O_P\left(h_f^{S+1} + \sqrt{\frac{\log(L)}{Lh_f^D}} + \frac{1}{h_f r_L^*}\right)$$

for any inner compact subset $\mathcal{C} \subset \mathscr{S}_{VX}$.

Proof. See Appendix A.6. □

As a corollary of the first item, we have that $\hat{U}'(\cdot)$ inherits the convergence rate of $\hat{U}(\cdot)$ because $U'(\cdot) = U(\cdot)/\lambda(\cdot)$. Regarding the estimation of the density, as expected, the rate of convergence is affected by the term $\hat{\lambda}^{-1}(\cdot)$. The reason is that the pseudo private values, $\hat{V}_{pl}$, inherit the rate of convergence of $\hat{\lambda}^{-1}(\cdot)$; see Appendix A.6 for a detailed discussion. In particular, if we set

$$h_f = \gamma_f \left[\varphi^{-1}\left(L^{\frac{2S(S+1)}{(2S+D+2)(2S+3)}}\right)\right]^{\frac{-1}{S+2}}, \tag{18}$$

for some constant $\gamma_f > 0$, we obtain $\|\hat{f}_{V|X}(\cdot|\cdot) - f_{V|X}(\cdot|\cdot)\|_{\mathcal{C},\infty} = O_P(h_f^{S+1})$. Furthermore, when $D \in \{0,1\}$, the results of subsection 3.3.1 can be easily applied to Proposition 1 and improve the obtained rates of convergence.

An alternative approach to estimating $f_{V|X}(\cdot|\cdot)$ would be to extend Marmer and Shneyerov (2012)'s procedure to allow for risk-averse bidders. Such an approach is based on the equality $1/f_{V|X}(v|x) = v'[F_{V|X}(v|x)|x]$ being $v'(\alpha|x) = \partial v(\alpha|x)/\partial\alpha$. This equality yields the formula

$$\begin{aligned}\frac{1}{f_{V|X}(v|x)} &= b'[F_{V|X}(v|x)|i,x] \\ &\quad + \lambda^{-1\prime}\left\{\frac{F_{V|X}(v|x)b'[F_{V|X}(v|x)|i,x]}{i-1}\right\}\left(\frac{1}{i-1}\right) \\ &\quad \times\left\{b'[F_{V|X}(v|x)|i,x] - \frac{F_{V|X}(v|x)g'\{b[F_{V|X}(v|x)|i,x]|i,x\}}{[g\{b[F_{V|X}(v|x)|i,x]|i,x\}]^3}\right\},\end{aligned}$$

where $i \in \mathcal{I}$ and $g'(b|i,x) = \partial g(b|i,x)/\partial b$; see eq. (3) in Marmer and Shneyerov (2012). Note that the right hand side depends on $\lambda^{-1\prime}(\cdot)$, the first derivative of $\lambda^{-1}(\cdot)$. The alternative estimator of $f_{V|X}(\cdot|\cdot)$ consists in replacing the unknown functions on the right-hand side with their empirical counterparts. The rate of uniform convergence in probability would be given by the slowest convergent term, i.e., the estimator of $\lambda^{-1\prime}(\cdot)$.

5 Implementation Guide and Simulations

5.1 Implementation Guide

The proposed estimators for $\lambda^{-1}(\cdot)$, $U(\cdot)$, and $f_{V|X}(\cdot|\cdot)$ involve many nonparametric estimation steps. This subsection provides an implementation guide to compute these estimators: $\hat{\lambda}^{-1}(\cdot)$, $\hat{U}(\cdot)$, and $\hat{f}_{V|X}(\cdot|\cdot)$. Suggestions about how to choose bandwidths and tuning parameters are provided. These suggestions are based on computational simplicity and the performance in simulations (subsection 5.2 below).

Pick a fixed value of $x \in \mathcal{X}$ such that the compatibility condition (4) holds. The degree of smoothness is set at $S = 1$. With respect to the kernel $k(\cdot)$, I suggest using the triweight kernel over $[-1,1]$: $k(t) = (35/32)(1-t^2)^3\mathbb{1}\{|t| \le 1\}$. To compute the proposed estimators, follow the next steps that can be implemented using a matrix-oriented software such as MATLAB.

Step 1 For each $i \in \mathcal{I}$, compute $\hat{f}_{IX}(i,x)$ from eq. (6). The suggested bandwidth is $h_X = (h_{X,1}, \ldots, h_{X,D})$, where $h_{X,d} = 1.06\hat{\sigma}_d L^{-1/(3D+3)}$ and $\hat{\sigma}_d$ denotes the sample standard deviation of $\{X_{1,d}, \ldots, X_{L,d}\}$ with $d = 1, \ldots, D$.

Step 2 Set $h_\varepsilon = L^{-2/(D+5)}$ and construct an equally spaced grid $\mathrm{A}^{(L)} = \{\mathrm{a}_1 \equiv h_\varepsilon < \mathrm{a}_2 < \mathrm{a}_3 < \cdots < \mathrm{a}_{\lceil T_\mathrm{A} \rceil} \equiv 1 - h_\varepsilon\}$ of size T_A. I suggest $T_\mathrm{A} = \lceil L^{6/5} \rceil$.

Step 3 For each $i \in \mathcal{I}$ and $t = 1, \ldots, T_\mathrm{A}$, calculate $\mathrm{b}_{i,t} \equiv \hat{b}(\mathrm{a}_t | i, x)$ from eq. (7).

Step 4 To estimate the boundaries $\underline{b}(x)$ and $\bar{b}(i,x)$ for $i \in \mathcal{I}$, I suggest using the estimators proposed by Guerre et al. (2000), eqs. (16)-(17). Such estimators can be computed in two steps.

Step 4A Construct the following hypercube containing $x = (x_1, \ldots, x_D)$:

$$\Pi = [x_1 - h_{\partial,1}, x_1 + h_{\partial,1}] \times \cdots \times [x_D - h_{\partial,D}, x_D + h_{\partial,D}]$$

with $h_{\partial,d} = \hat{\sigma}_d [\log(L)/L]^{1/(D+1)}$.

Step 4B $\underline{b}(x)$ and $\bar{b}(i,x)$ are estimated by

$$\hat{\underline{b}}(x) = \min_{p,l} \{B_{pl} : X_l \in \Pi\}$$

and $\hat{\bar{b}}(i,x) = \max_{p,l}\{B_{pl} : I_l = i, X_l \in \Pi, \}$, respectively.[8]

Step 5 For each $i \in \mathcal{I}$, compute the coefficients associated with $\hat{g}^*(\cdot|i,x)$ following the next steps.

[8] These estimators converges uniformly over $\mathcal{X}$ at a rate of $[\log(L)/L]^{1/(D+1)}$. Alternatively, as in Campo et al. (2011), one can use Korostelev and Tsybakov (1993)'s approach to estimate $\underline{b}(x)$ and $\bar{b}(i,x)$. This approach leads to a faster convergence rate, but is computationally intensive. I remark that both procedures lead to the same estimators when there are no covariates: $\hat{\underline{b}}(x) = \min_{p,l}\{B_{pl}\}$ and $\hat{\bar{b}}(i,x) = \max_{p,l}\{B_{pl} : I_l = i\}$.

Step 5A Construct a set of positive integers whose elements are candidates for J_L; e.g., $\mathrm{J}^{(L)} = \{5, 10, \ldots, \lceil L^{1/4} \rceil\}$.

Step 5B For each $J_L \in \mathrm{J}^{(L)}$, compute the coefficients that solve eqs. (8) and denote them by $\hat{\theta}^{(J_L)}(i,x) = \left(\hat{\theta}_1^{(J_L)}(i,x), \ldots, \hat{\theta}_{J_L}^{(J_L)}(i,x)\right)$. Such coefficients can be obtained using a standard Newton-Raphson algorithm; in MATLAB e.g., use the command 'fsolve' with zeros as initial values. To speed up this procedure, series basis can be employed instead of Legendre polynomials.

Step 5C Let $\hat{g}^{*(J_L)}(\cdot|i,x)$ be the conditional density estimator associated with $\hat{\theta}^{(J_L)}(i,x)$. Choose the value of $J_L \in \mathrm{J}^{(L)}$ that minimizes the expression

$$\int_0^1 \hat{g}^{*(J_L)}(b^*|i,x)^2 db^* \\ - \frac{2}{\hat{f}_{IX}(i,x) i L h_X^D} \sum_{l=1}^{L} \sum_{p=1}^{I_l} \hat{g}^{*(J_L)}\left(\hat{B}_{lp}^*|i,x\right) \mathbb{1}\{I_l = i\} K\left(\frac{x - X_l}{h_X}\right). \tag{19}$$

Denote it by $\tilde{J}_L$ and let $\hat{\theta}^{(\tilde{J}_L)}(i,x)$ be the corresponding vector of coefficients.[9]

Step 6 For each $i \in \mathcal{I}$ and $t = 1, \ldots, T_{\mathrm{A}}$, calculate

$$\hat{g}\left(\mathrm{b}_{i,t}|i,x\right) = \left[\hat{\bar{b}}(i,x) - \hat{\underline{b}}(x)\right]^{-1} \hat{g}^{*(\tilde{J}_L)}\left[\frac{\mathrm{b}_{i,t} - \hat{\underline{b}}(x)}{\hat{\bar{b}}(i,x) - \hat{\underline{b}}(x)}\middle| i,x\right]$$

and $\mathrm{R}_{i,t} \equiv \hat{R}(\mathrm{a}_t|i,x) = \mathrm{a}_t / \left[(i-1)\hat{g}\left(\mathrm{b}_{i,t}|i,x\right)\right]$, using $\hat{\theta}^{(\tilde{J}_L)}(i,x)$ from *Step 5C*.

Step 7 Compute the coefficients associated with $\hat{\lambda}^{-1}(\cdot)$ following the next steps.

[9]Expression (19) is associated with integrated squared difference between $\hat{g}^{*(J_L)}(\cdot|i,x)$ and $g^*(\cdot|i,x)$; see eqs. (19)-(20) in Li and Racine (2007). In the absence of covariates, $D = 0$, J_L can be chosen according to Wu (2010)'s suggestion, which consists in minimizing a Akaike information criterion.

Step 7A Construct a set of positive integers whose elements are potential values for M_L; e.g., $\mathrm{M}^{(L)} = \{5, 10, \ldots, \lceil L^{1/2} \rceil\}$.

Step 7B Pick $\bar{U} = \max\{\mathrm{R}_{i,t} : i \in \mathcal{I}, t = 1, \ldots, T_{\mathrm{A}}\}(1 + 0.01\sqrt[3]{1/L})$ and let $\mathcal{B}^{(M_L)}$ denote the set of coefficients $(\beta_1, \ldots, \beta_{M_L})$ satisfying linear conditions (11).[10] For each $M_L \in \mathrm{M}^{(L)}$, compute the coefficients that solve

$$\min_{\beta \in \mathcal{B}^{(M_L)}} \sum_{i \in \mathcal{I} \setminus \{\underline{i}\}} \max_{t=1,\ldots,T_{\mathrm{A}}} \left| \mathrm{b}_{i,t} - \mathrm{b}_{\underline{i},t} + \sum_{m=1}^{M_L} \beta_m \left[P_{M_L,m}(\mathrm{R}_{i,t}) - P_{M_L,m}\left(\mathrm{R}_{\underline{i},t}\right) \right] \right|$$

and denote them by $\hat{\beta}^{(M_L)} = (\hat{\beta}_1^{(M_L)}, \ldots, \hat{\beta}_{(M_L)}^{(M_L)})$. This minimization problem can be solved using a minimax optimization algorithm such as Stocco, Salcudean, and Sassani (1998). In MATLAB e.g., use the command 'minimax' included in the optimization toolbox. I suggest $\beta_m = m\bar{U}/(2M_L)$ as initial values.

Step 7C From the set $\{\hat{\beta}^{(M_L)} : M_L \in \mathrm{M}^{(L)}\}$, choose the vector of coefficients that minimizes the sum of squared deviations from the compatibility condition:

$$\sum_{i \in \mathcal{I} \setminus \{\underline{i}\}} \sum_{t=1}^{T_{\mathrm{A}}} \left\{ \mathrm{b}_{i,t} - \mathrm{b}_{\underline{i},t} + \sum_{m=1}^{M_L} \hat{\beta}_m^{(M_L)} \left[P_{M_L,m}(\mathrm{R}_{i,t}) - P_{M_L,m}\left(\mathrm{R}_{\underline{i},t}\right) \right] \right\}^2$$

Denote such a vector by $\hat{\beta}^{(\tilde{M}_L)} = (\hat{\beta}_1^{(\tilde{M}_L)}, \ldots, \hat{\beta}_{(\tilde{M}_L)}^{(\tilde{M}_L)})$.

Step 8 The estimator of $\lambda^{-1}(u)$ is given by

$$\hat{\lambda}^{-1}(u) = \sum_{m=1}^{\tilde{M}_L} \hat{\beta}_m^{(\tilde{M}_L)} P_{\tilde{M}_L,m}(u).$$

[10] Since $\bar{C}_{\mathscr{H}} > 0$ can be taken to arbitrarily large, condition (12) can be ignored.

Researchers interested only in the density of private values can skip *Step 9* and move directly to *Step 10*.

Step 9 Computation of $\hat{U}(y)$ can be done in two steps.

Step 9A Pick $\bar{u} = \bar{U}/(1 + 0.02\sqrt[3]{1/L})$, $\underline{y} = \hat{\lambda}^{-1}(\bar{u})/1.01$, and construct an equally spaced grid $\Lambda^{(L)} = \{\mathrm{y}_1 \equiv \underline{y} < \mathrm{y}_2 < \cdots < \mathrm{y}_{T_\Lambda} \equiv \bar{y}\}$ of size T_Λ. I suggest $T_\Lambda = L^2$. Then, calculate $\hat{\lambda}(\mathrm{y}_t)$ for each $t = 1, \ldots, T_\Lambda$.

Step 9B The estimator of $U(y)$ is given by

$$\hat{U}(y) = \exp\left[-\frac{\bar{y} - y}{T_\Lambda - 1}\sum_{t=1}^{T_\Lambda}\frac{1}{\hat{\lambda}(\mathrm{y}_t)}\right].$$

Step 10 Compute $\hat{f}_{V|X}(v|x)$ following the next steps.

Step 10A Use formula (16) to obtain the pseudo private values $\hat{V}_{pl}$. To avoid trimming observations, when $D = 0$, I suggest computing $\hat{V}_{pl}$ as follows:

$$\hat{V}_{pl} = B_{pl} + \hat{\lambda}^{-1}\left[\frac{1}{(I_l - 1)}\frac{\hat{G}(B_{pl}|I_l)}{\hat{g}(B_{pl}|I_l)}\right], \tag{20}$$

being $\hat{G}(\cdot|\cdot)$ and $\hat{g}(\cdot|\cdot)$ the estimators proposed in subsection 3.2. To ensure that pseudo private values are monotone in bids, use the procedure suggested in sec. 4 of Henderson, List, Millimet, Parmeter, and Price (2012).

Step 10B Use formula (17) to obtain $\hat{f}_{V|X}(v|x) = \hat{f}_{VX}(v,x)/\hat{f}_X(x)$. The suggested bandwidth is $h_f = 1.06\hat{\sigma}_{\hat{V}}L^{-1/(3D+6)}$, being $\hat{\sigma}_{\hat{V}}$ the sample standard deviation of the pseudo private values.

Table 1: Sensitivity Analysis of $\hat{\lambda}^{-1}(\cdot)$ based on IMSE

Design $f_V(\cdot) - U(\cdot)$		$L = 450$ $M_L = 3$	6	9	13	M_L^*	$L = 900$ $M_L = 3$	6	9	13	M_L^*
D1 -	U1	4.43	4.19	3.96	3.79	3.81	2.39	2.35	2.24	2.20	2.18
	U2	4.96	4.67	4.36	4.15	4.17	3.04	2.68	2.42	2.30	2.29
	U3	4.70	3.82	3.71	3.69	3.64	3.12	2.19	2.04	1.92	1.89
	U4	2.53	3.09	5.70	11.37	3.91	1.28	1.71	5.51	12.53	1.92
D2 -	U1	5.01	5.07	4.88	4.75	4.75	2.52	2.90	2.69	2.60	2.59
	U2	5.23	5.19	4.92	4.66	4.73	3.39	3.23	2.97	2.70	2.72
	U3	5.74	4.99	4.90	4.81	4.72	3.78	2.93	2.78	2.75	2.65
	U4	3.41	4.47	6.85	12.80	5.59	1.39	1.97	4.93	13.38	2.71

All numbers have been multiplied by 10.

5.2 Monte Carlo Experiments

This subsection presents Monte Carlo simulations to investigate the finite sample performance of the proposed estimators. The evaluation criteria are the integrated mean squared error (IMSE), bias, and mean squared error (MSE). The proposed nonparametric estimators are compared with Campo et al. (2011)'s CARA semiparametric estimator.

The design of the experiment is as follows. It is assumed that there are no covariates ($D = 0$). We consider the next two densities of valuations for $f_V(\cdot)$.

D1: Truncated lognormal density with parameters 0 and 1, truncated at 0.055 and 2.5, and rescaled so that it has support $[0, 10]$.

D2: Truncated exponential density with parameter 1/5 and truncated at 10.

We set $\bar{u} = 4$ and $\bar{y} = 1$. The functional forms of $\lambda^{-1}(\cdot)$ and $U(\cdot)$ are given by $\lambda^{-1}(u) = \log(1 + u/\eta_2)/\eta_1$ and

$$U(y) = \exp\left(\frac{\bar{y} - y}{\eta_2}\right)\left[\frac{1 - \exp(\eta_1 y)}{1 - \exp(\eta_1 \bar{y})}\right]^{\frac{1}{\eta_1 \eta_2}},$$

respectively, being (η_1, η_2) the risk-aversion parameters. The choices of $f_V(\cdot)$ and $U(\cdot)$ are convenient because the corresponding bidding functions have closed-form expres-

Table 2: Performance of $\hat{U}(\cdot)$ based on IMSE, Bias, and MSE

Design / Estimator		$L=450$					$L=900$				
		IMSE	$U(0.33)$		$U(0.67)$		IMSE	$U(0.33)$		$U(0.67)$	
			Bias	MSE	Bias	MSE		Bias	MSE	Bias	MSE
D1 - U1)	Nonparam.	0.111	-0.145	0.156	-0.123	0.038	0.077	-0.121	0.110	-0.090	0.025
	CARA	0.241	-2.068	0.437	-1.206	0.153	0.230	-2.024	0.417	-1.163	0.142
U2)	Nonparam.	0.140	-0.387	0.180	-0.184	0.035	0.090	-0.306	0.108	-0.131	0.019
	CARA	0.659	-3.340	1.121	-1.817	0.335	0.659	-3.343	1.122	-1.819	0.334
U3)	Nonparam.	0.069	0.451	0.111	0.184	0.033	0.037	0.315	0.063	0.130	0.020
	CARA	0.010	-0.239	0.016	-0.228	0.014	0.008	-0.171	0.012	-0.164	0.011
U4)	Nonparam.	0.068	0.422	0.102	0.229	0.028	0.038	0.277	0.058	0.150	0.016
	CARA	0.002	0.090	0.003	0.086	0.003	0.002	0.081	0.002	0.079	0.002
D2 - U1)	Nonparam.	0.136	-0.110	0.190	-0.121	0.047	0.089	-0.059	0.123	-0.073	0.028
	CARA	0.239	-2.064	0.434	-1.201	0.151	0.228	-2.015	0.413	-1.154	0.139
U2)	Nonparam.	0.159	-0.420	0.208	-0.209	0.042	0.109	-0.342	0.136	-0.159	0.024
	CARA	0.660	-3.344	1.123	-1.820	0.335	0.662	-3.351	1.126	-1.827	0.337
U3)	Nonparam.	0.099	0.565	0.158	0.229	0.045	0.057	0.423	0.095	0.177	0.029
	CARA	0.010	-0.225	0.015	-0.215	0.013	0.007	-0.151	0.012	-0.146	0.010
U4)	Nonparam.	0.095	0.555	0.143	0.298	0.039	0.055	0.367	0.083	0.196	0.022
	CARA	0.002	0.093	0.003	0.090	0.003	0.001	0.077	0.002	0.075	0.002

All numbers have been multiplied by 10.

sions. Observe that the case $\eta_2 = 1/\eta_1$ corresponds to a CARA utility function with parameter η_1. The next four values of (η_1, η_2) are considered.[11]

U1: $\eta_1, = 0.3$ and $\eta_2 = 6$.

U2: $\eta_1 = -0.1$ and $\eta_2 = -30$.

U3: $\eta_1 = 0.5$ and $\eta_2 = 2$, i.e., CARA case with $\lambda^{-1}(u) = \log(1 + u/2)/0.5$ and $U(y) = [1 - \exp(0.5y)]/[1 - \exp(0.5)]$.

U4: $\eta_1 = 0$ and $\eta_2 = +\infty$, i.e., risk-neutral case with $\lambda^{-1}(u) = u$ and $U(y) = y$.

For each combination of density and risk-aversion parameters, we study two sample sizes. A small sample of $L = 450$ auctions: 300 auctions with 2 bidders and 150 auctions with 4 bidders, giving a total of 1,200 bids. A large sample of $L = 900$ auctions: 600

[11]The values of $\bar{u}$ and $\bar{y}$ have been chosen so that $\bar{u} < \bar{R}$ and $\bar{y} < \lambda^{-1}(\bar{u})$ hold for each case.

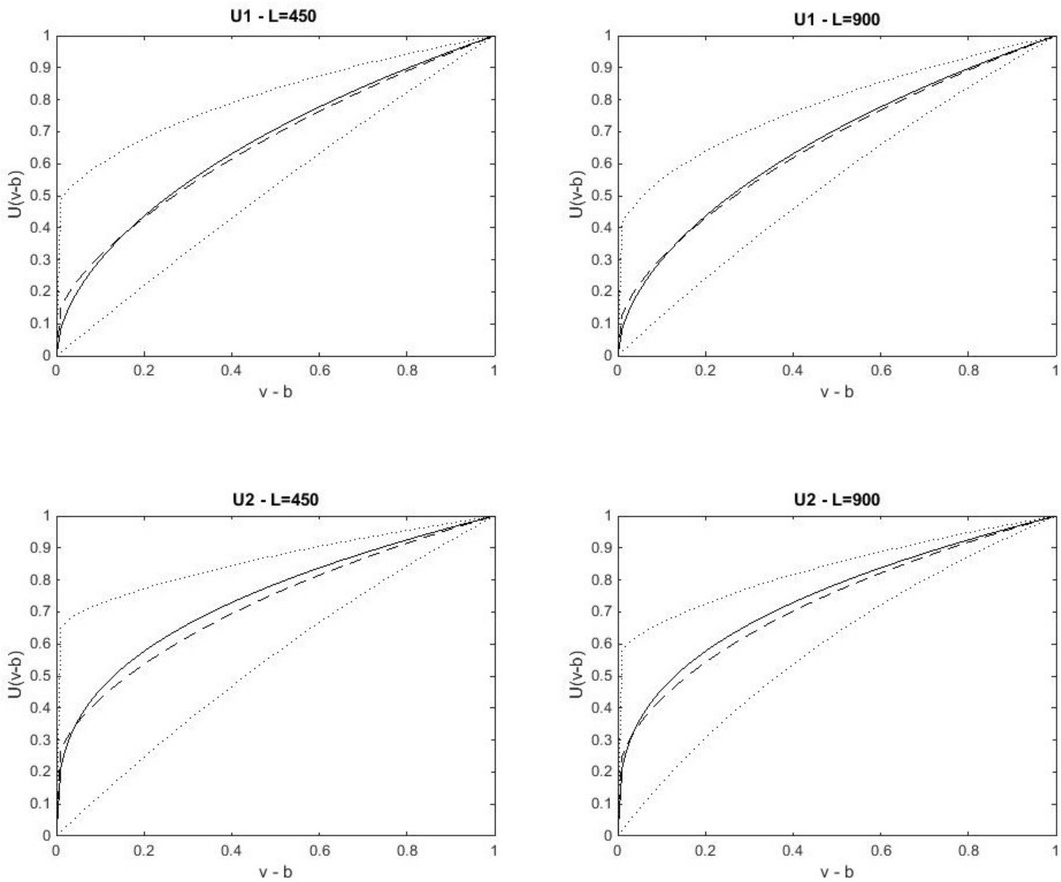

Figure 1: Estimating $U(\cdot)$ –cases $U1$ and $U2$ under density $D1$–

auctions with 2 bidders and 300 auctions with 4 bidders, giving a total of $2,400$ bids. Each experiment employs 2,000 replications.

In each replication, first, we generate independent valuations from either $D1$ or $D2$. Second, we compute the associated bids according to the values of (η_1, η_2) and using the equilibrium bidding function (2). Third, we compute $\hat{\lambda}^{-1}(\cdot)$ using the generated bids and following *Steps 1-8* of previous subsection. Fourth, we compute $\hat{U}(\cdot)$ and $\hat{f}_V(\cdot)$ following *Steps 9-10*. Several remarks are noteworthy. In *Step 5*, we consider $\mathrm{J}^{(L)} = \{3, 5, 7\}$ and series basis are employed instead of Legendre polynomials. In *Step 7*, we consider $\mathrm{M}^{(L)} = \{3, 6, 9, 13\}$ as potential candidates for M_L. In *Step 10*, pseudo

Table 3: Performance of $\hat{f}_V(\cdot)$ based on IMSE, Bias, and MSE

Design / Estimator		$L = 450$ IMSE	$f_V(3)$ Bias	$f_V(3)$ MSE	$f_V(7)$ Bias	$f_V(7)$ MSE	$L = 900$ IMSE	$f_V(3)$ Bias	$f_V(3)$ MSE	$f_V(7)$ Bias	$f_V(7)$ MSE
D1 - U1)	Nonparam.	0.020	-0.002	0.002	0.000	0.002	0.010	0.004	0.001	0.014	0.001
	CARA	0.019	-0.097	0.003	0.041	0.002	0.015	-0.081	0.002	0.052	0.001
	Neutral	0.028	-0.228	0.006	-0.051	0.001	0.025	-0.233	0.006	-0.057	0.001
U2)	Nonparam.	0.024	0.001	0.003	-0.001	0.003	0.011	0.004	0.001	0.012	0.001
	CARA	0.034	-0.182	0.005	0.097	0.003	0.029	-0.182	0.005	0.096	0.002
	Neutral	0.037	-0.263	0.008	0.013	0.001	0.033	-0.265	0.008	0.004	0.000
U3)	Nonparam.	0.020	-0.029	0.002	-0.029	0.003	0.009	-0.010	0.001	0.008	0.001
	CARA	0.016	-0.090	0.003	-0.041	0.001	0.011	-0.059	0.002	-0.028	0.001
	Neutral	0.025	-0.211	0.005	-0.111	0.002	0.023	-0.209	0.005	-0.116	0.002
U4)	Nonparam.	0.032	-0.013	0.003	-0.049	0.006	0.016	-0.008	0.002	0.004	0.003
	CARA	0.030	0.039	0.004	-0.082	0.005	0.020	0.043	0.002	-0.053	0.004
	Neutral	0.021	-0.010	0.002	-0.024	0.004	0.012	-0.002	0.001	-0.004	0.002
D2 - U1)	Nonparam.	0.019	-0.009	0.002	-0.020	0.003	0.009	-0.009	0.001	-0.001	0.001
	CARA	0.016	-0.083	0.003	0.034	0.002	0.011	-0.068	0.002	0.040	0.001
	Neutral	0.025	-0.213	0.005	-0.058	0.001	0.022	-0.220	0.005	-0.060	0.001
U2)	Nonparam.	0.022	0.003	0.002	-0.004	0.003	0.011	0.000	0.001	0.001	0.001
	CARA	0.027	-0.150	0.004	0.094	0.003	0.022	-0.154	0.003	0.087	0.002
	Neutral	0.029	-0.230	0.006	0.008	0.001	0.025	-0.234	0.006	0.002	0.001
U3)	Nonparam.	0.020	-0.029	0.002	-0.037	0.003	0.009	-0.024	0.001	-0.002	0.001
	CARA	0.015	-0.081	0.003	-0.041	0.001	0.010	-0.059	0.002	-0.033	0.001
	Neutral	0.026	-0.206	0.005	-0.115	0.002	0.023	-0.212	0.005	-0.114	0.002
U4)	Nonparam.	0.036	-0.018	0.003	-0.066	0.006	0.018	-0.016	0.002	-0.011	0.003
	CARA	0.034	0.039	0.004	-0.091	0.006	0.021	0.036	0.002	-0.061	0.004
	Neutral	0.024	-0.013	0.002	-0.019	0.004	0.013	-0.010	0.001	-0.019	0.002

All numbers have been multiplied by 10.

private values are constructed using formula (20).

Table 1 reports the IMSE of $\hat{\lambda}^{-1}(\cdot)$ over $[0,4]$ and explores the sensitivity of the results to different values of M_L. The column M_L^* corresponds to the choice of M_L based on *Step 7C* of previous subsection. Increasing M_L reduces the IMSE except for the risk-neutral case $U4$. The intuition behind this exception is that the identity function is a (series) polynomial of degree 1, so increasing the degree of the polynomial rises the IMSE due to overfitting. I highlight that choosing M_L^* as the degree of the Bernstein polynomial works as a solution to this problem.

Table 2 provides the IMSE of $\hat{U}(\cdot)$ over $[0,1]$ together with the bias and MSE at certain points. It also includes Campo et al. (2011)'s CARA semiparametric estimator

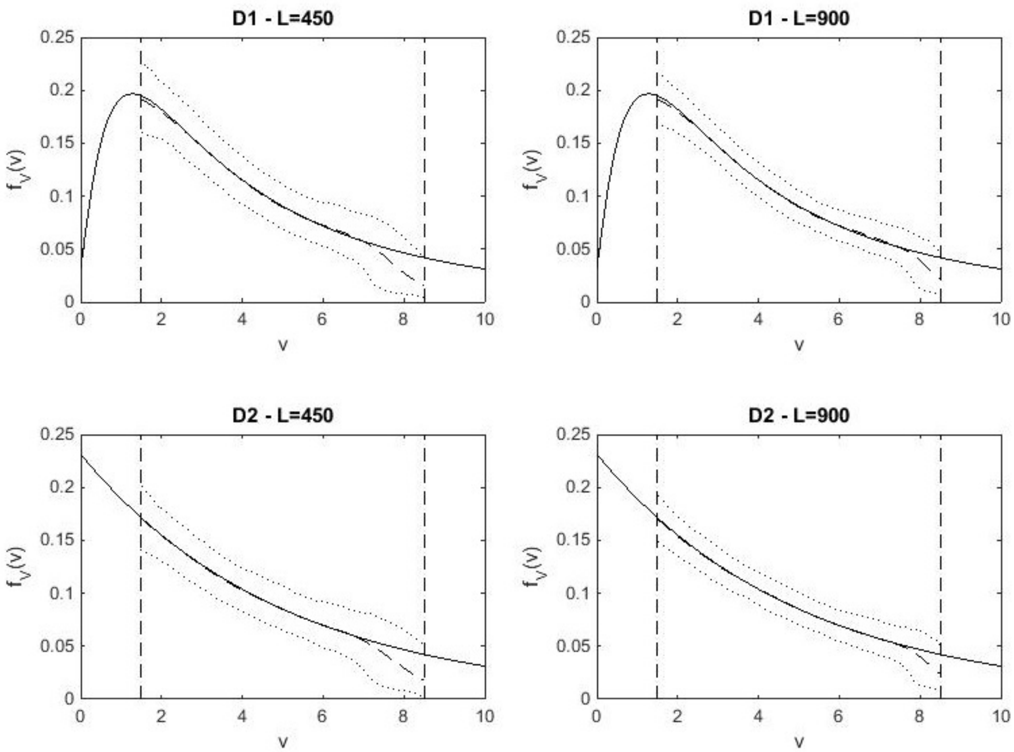

Figure 2: Estimating $f_V(\cdot)$ –cases $D1$ and $D2$ under utility function $U1$–

(row 'CARA') for comparison purposes. As expected, the nonparametric estimator has the best performance IMSE under $U1$ and $U2$, whereas the semiparametric competitor performs better under $U3$ and $U4$ (cases in which the CARA assumption holds). For the cases $U1$ and $U2$ under $D1$, Figure 1 presents the utility function (solid line) and the pointwise mean of $\hat{U}(\cdot)$ (dashed line) along with the $5^{\text{th}}/95^{\text{th}}$ percentiles (dotted lines). As can be noted, $\hat{U}(\cdot)$ has a good finite-sample performance in terms of bias.

Table 3 analyses the performance of $\hat{f}_V(\cdot)$ and studies the consequences of misspecifying the shape of $\lambda^{-1}(\cdot)$. The row 'Nonparam.' reports the results of estimating $f_V(\cdot)$ from formula (17) using the pseudo private values (20). The row 'CARA' repeats the same exercise but, when generating the pseudo private values, $\hat{\lambda}^{-1}(\cdot)$ is replaced with the CARA semiparametric estimator. Similarly, in the row 'Neutral', $\hat{\lambda}^{-1}(\cdot)$ is replaced with the identity function. The nonparametric estimator performs very well in terms of bias and has the lowest IMSE under $U2$ when $L = 450$, as well as, under $U1$ and $U2$ when $L = 900$. For the cases $D1$ and $D2$ under $U1$, Figure 2 shows the private value density (solid line) and the pointwise mean of the nonparametric estimator (dashed line) along with the $5^{\text{th}}/95^{\text{th}}$ percentiles (dotted lines).

6 Concluding Remarks

This paper has studied a first-price sealed-bid auction with risk-averse bidders, independent private values, and a non-binding reserve price. In this context, I have proposed nonparametric estimators for the bidders' utility function and the density of private values. The key idea has been to characterize both functions by an argument that minimizes certain criterion function. After estimating this minimizer by its empirical counterpart, estimators for the bidders' utility function and the density of private values have been constructed from the estimator of such a minimizer.

The method proposed in this paper allows us to estimate bidders' risk preferences without placing any parametric restrictions –such as CARA or CRRA– on the utility function. In this way, this paper extends the literature on structural econometrics of first-price auctions by developing an estimator for the bidders' utility function that can incorporate any type of risk preference. The relevance of this contribution relies on the fact that evidence suggests that risk aversion is an essential component of bidders' behavior, but there is no consensus on which concept of risk aversion is the most appropriate to describe such behavior.

There are many possible directions for further research. The first is to establish the optimal convergence rate for the parameters of the model, i.e., the fastest rate at which $U(\cdot)$ and $f_{V|X}(\cdot|\cdot)$ can be estimated nonparametrically. Second, the asymptotic distribution of the proposed estimators needs to be determined. The difficulty relies on the fact that the criterion function is a nonsmooth functional based on the sup-norm. Third, since the density of private values is over-identified when $\#(\mathcal{I}) \geq 3$, it would be possible to construct a test to verify whether or not bidders' participation is exogenous. In a recent working paper, Liu and Luo (2014) develop such a test under risk-neutral bidders.

Independent private values is a maintained assumption in this paper. However, the obtained results can be extended to allow affiliated private values. To do so, we should construct a criterion function based on eq. (9) of Guerre et al. (2000) and extend the results of Section 3. Considering a more general setting with correlated private values, it would be interesting to apply Aradillas-Lopéz, Gandhi, and Quint (2013)'s approach to first-price auctions with risk-averse bidders and study whether economic measures of interest –such as profits and optimal reserve price– could be partially identified.

In view of future empirical applications, the proposed estimators can be employed to recover the set of optimal reserve prices. This set depends on both the bidders' risk aversion and the distribution of valuations. So far, the optimal reserve price has been obtained only under the assumption that bidders are risk-neutral. For instance, Li et al. (2003) have considered a first-price auction with affiliated private values, but assuming that bidders are risk-neutral. Their approach may be extended by allowing bidders to be risk-averse.

A Appendix: Proofs

This appendix provides the proofs of all lemmas, propositions, and theorems stated in the body of the text. Proofs of the auxiliary lemmas are given in Appendix B.

A.1 Proof of Lemma 1

I start the proof with an auxiliary lemma.

Lemma A.1. *Suppose Assumptions 1-2 hold and pick any* $(i,x) \in \mathcal{I}\times\mathcal{X}$. *Then,* $b(\cdot|i,\cdot)$ *admits* $S+1$ *continuous partial derivatives on* $[0,1]\times\mathcal{X}$, $b^{(S+2)}(\cdot|i,x)$ *is continuous on* $(0,1]$, *and* $b'(\cdot|i,x)$ *is bounded away from 0 uniformly on* $x \in \mathcal{X}$. *Furthermore,* $b'(0|i_1,x)/(i_1-1) > b'(0|i_2,x)/(i_2-1)$ *for every* $(i_1,i_2)\in\mathcal{I}^*$ *and* $x\in\mathcal{X}$.

Proof. See Appendix B.1. □

1. Pick any $i\in\mathcal{I}$. Note that

$$R'(\alpha|i,x) \equiv \frac{\partial R(\alpha|i,x)}{\partial\alpha} = \frac{1}{i-1}[b'(\alpha|i,x) - \alpha b''(\alpha|i,x)] \tag{A.1}$$

by definition of $R(\alpha|i,x)$, stated in Observation 1.3.(a). By Lemma A.1, $b'(\cdot|i,x)$ is bounded away from 0 uniformly on $x\in\mathcal{X}$ and also $|b''(\cdot|i,x)|$ is bounded above uniformly on $x\in\mathcal{X}$ because $b(\cdot|i,\cdot)$ admits (at least) 2 continuous partial derivatives on $[0,1]\times\mathcal{X}$ and this set is compact. Then, we pick

$$\tilde{\alpha} = \frac{\min\{b'(\alpha|i,x) : (\alpha,x)\in[0,1]\times\mathcal{X}\}}{2\max\{|b''(\alpha|i,x)| : (\alpha,x)\in[0,1]\times\mathcal{X}\}};$$

if the numerator is bigger than the denominator, we choose $\tilde{\alpha} = 1$. From expression (A.1), for any $(\alpha, x) \in [0, \tilde{\alpha}] \times \mathcal{X}$, $R'(\alpha|i, x)$ can be bounded below by

$$\begin{aligned}&\frac{1}{i-1}[\min\{b'(\alpha|i,x) : (\alpha,x) \in [0,1] \times \mathcal{X}\} - \tilde{\alpha} \max\{|b''(\alpha|i,x)| : (\alpha,x) \in [0,1] \times \mathcal{X}\}] \\ &\geq \frac{1}{2(i-1)} \min\{b'(\alpha|i,x) : (\alpha,x) \in [0,1] \times \mathcal{X}\},\end{aligned}$$

so $\underline{c}_R > 0$ can be taken to be the right hand side. This inequality follows by construction of $\tilde{\alpha}$ and becomes an equality when $\tilde{\alpha} < 1$. Recall that $\mathcal{I}$ is finite, thus the dependence of $\underline{c}_R$ on i is irrelevant.

2. Pick any $(i_1, i_2) \in \mathcal{I}^*$. For $(\alpha, x) \in [0,1] \times \mathcal{X}$, define the nonnegative function $\delta(\alpha, x) = R(\alpha|i_1, x) - R(\alpha|i_2, x) = \tilde{\delta}(\alpha, x)\alpha$, being

$$\tilde{\delta}(\alpha, x) \equiv \frac{b'(\alpha|i_1, x)}{i_1 - 1} - \frac{b'(\alpha|i_2, x)}{i_2 - 1} > 0. \tag{A.2}$$

Then, we can write $R'(\alpha|i_1, x) - R'(\alpha|i_2, x) = \tilde{\delta}'(\alpha, x)\alpha + \tilde{\delta}(\alpha, x)$, where $\tilde{\delta}'(\cdot, \cdot)$ stands for the partial derivative of $\tilde{\delta}(\cdot, \cdot)$ with respect to its first argument. This difference can be bounded as follows:

$$R'(\alpha|i_1, x) - R'(\alpha|i_2, x) = \tilde{\delta}'(\alpha, x)\alpha + \tilde{\delta}(\alpha, x) \geq -\bar{C}_\delta \alpha + \Delta(\alpha),$$

where $\bar{C}_\delta \equiv \max\{|\tilde{\delta}'(\alpha, x)| : (\alpha, x) \in [0,1] \times \mathcal{X}\} < +\infty$ by Lemma A.1 and $\Delta(\alpha) \equiv \min_{x \in \mathcal{X}} \tilde{\delta}(\alpha, x)$. Considering the second term on the right-hand side, note that $\Delta(\cdot)$ is continuous on $[0,1]$; see e.g. Theorem 3.6 in Stokey, Lucas, and Prescott (1989). It also satisfies $\Delta(0) = \min_{x \in \mathcal{X}} \tilde{\delta}(0, x) > 0$ because $\tilde{\delta}(0, \cdot)$ is strictly positive (Lemma A.1) and continuous, while $\mathcal{X}$ is compact. By continuity of $\Delta(\cdot)$, there is $\tilde{\alpha}'_\Delta \in (0, \tilde{\alpha}]$ such that $\Delta(\alpha) \geq \Delta(0)/2 > 0$ for all $\alpha \in [0, \tilde{\alpha}'_\Delta]$. Now pick $\tilde{\alpha}' = \min\{\Delta(0)/(4\bar{C}_\delta), \tilde{\alpha}'_\Delta\} \in (0, \tilde{\alpha}]$, so for every $\alpha \in [0, \tilde{\alpha}']$ we have

$$R'(\alpha|i_1, x) - R'(\alpha|i_2, x) \geq -\bar{C}_\delta \alpha + \Delta(\alpha) \geq -\bar{C}_\delta \tilde{\alpha}' + \frac{\Delta(0)}{2} \geq -\bar{C}_\delta \frac{\Delta(0)}{4\bar{C}_\delta} + \frac{\Delta(0)}{2} = \frac{\Delta(0)}{4}.$$

To complete the proof, choose $\underline{c}'_R = \Delta(0)/4$ and note that it is independent of x.

3. First, define the function

$$\tilde{\Delta}_1(\alpha) = \min_{x \in \mathcal{X}} [R'(\alpha|i_1, x) - R'(0|i_2, x)]$$

for $\alpha \in [0, \tilde{\alpha}']$, which satisfies $\tilde{\Delta}_1(0) \geq \underline{c}'_R$ due to the second item. By continuity of $\tilde{\Delta}_1(\cdot)$, there is $\tilde{\alpha}''_1 \in (0, \tilde{\alpha}']$ such that $\tilde{\Delta}_1(\alpha) > \underline{c}'_R/2$ for all $\alpha \in [0, \tilde{\alpha}''_1]$. By construction of $\tilde{\Delta}_1(\cdot)$ and $\tilde{\alpha}''_1$, we have $R'(\alpha|i_1, x) - R'(0|i_2, x) \geq \tilde{\Delta}_1(\alpha) > \underline{c}'_R/2$ for every $(\alpha, x) \in [0, \tilde{\alpha}''_1] \times \mathcal{X}$ and therefore

$$R'(\alpha|i_1, x) > \frac{\underline{c}'_R}{2} + R'(0|i_2, x). \tag{A.3}$$

Second, proceeding in a similar manner, define the function

$$\tilde{\Delta}_2(\alpha) = \max_{x \in \mathcal{X}} [R'(\alpha|i_2, x) - R'(0|i_2, x)]$$

for $\alpha \in [0, \tilde{\alpha}']$, which clearly satisfies $\tilde{\Delta}_2(0) = 0$. By continuity of $\tilde{\Delta}_2(\cdot)$, there is $\tilde{\alpha}''_2 \in (0, \tilde{\alpha}']$ such that $\tilde{\Delta}_2(\alpha) < \underline{c}'_R/2$ for all $\alpha \in [0, \tilde{\alpha}''_2]$. Hence, we have $R'(\alpha|i_2, x) - R'(0|i_2, x) \leq \tilde{\Delta}_2(\alpha) < \underline{c}'_R/2$ for every $(\alpha, x) \in [0, \tilde{\alpha}''_2] \times \mathcal{X}$ and, as a result,

$$R'(\alpha|i_2, x) < \frac{\underline{c}'_R}{2} + R'(0|i_2, x). \tag{A.4}$$

After taking $\tilde{\alpha}'' = \min\{\tilde{\alpha}_1'', \tilde{\alpha}_2''\} \in (0, \tilde{\alpha}']$, it follows that

$$\underline{c}_R \leq \max_{\alpha \in [0,\tilde{\alpha}'']} R'(\alpha|i_2, x) < \frac{\underline{c}_R'}{2} + R'(0|i_2, x) < \min_{\alpha \in [0,\tilde{\alpha}'']} R'(\alpha|i_1, x).$$

for every $x \in \mathcal{X}$, where $\underline{c}_R > 0$ was obtained in item 1 and $\tilde{\alpha}''$ is clearly independent of x. The first inequality follows by Lemma 1.1, the second by eq. (A.4), and the third by eq. (A.3). The desired result finally emerges by $\min\{R'(\alpha|i_1, x) : \alpha \in [0, \tilde{\alpha}'']\} > 0$, compactness of $\mathcal{X}$, and continuity of

$$\frac{\max\{R'(\alpha|i_2, \cdot) : \alpha \in [0, \tilde{\alpha}'']\}}{\min\{R'(\alpha|i_1, \cdot) : \alpha \in [0, \tilde{\alpha}'']\}} \in (0, 1).$$

A.2 Proof of Lemma 2

Pick any $i \in \mathcal{I} \backslash \{\underline{i}\}$ and $x \in \text{interior}(\mathcal{X})$. This proof starts with an auxiliary lemma. Consider $\tilde{\alpha}'' > 0$ obtained in Lemma 1.3 and denote the ceiling function by $\lceil \cdot \rceil$.

Lemma A.2. *For each $u \in [0, \bar{R}]$, define recursively the following sequence: $\alpha_0(u) = \min\{\alpha \in [0, 1] : R(\alpha|\underline{i}, x) = u\}$ and*

$$\alpha_t(u) = \min\{\alpha \in [0, 1] : R(\alpha|\underline{i}, x) = R[\alpha_{t-1}(u)|i, x]\} \tag{A.5}$$

for $t \in \mathbb{N}$. Under Assumptions 1-2, the following statements hold.

1. *There exists a finite $\tilde{T} \in \mathbb{N}$, which is independent of (u, x), such that $\alpha_t(u) \leq \tilde{\alpha}''$ for all $t \geq \tilde{T}$ and $u \in [0, \bar{R}]$.*

2. *Using Lemma 1.3, define*

$$\bar{\kappa} = \max_{x \in \mathcal{X}} \left[\frac{\max\{R'(\alpha|i, x) : \alpha \in [0, \tilde{\alpha}'']\}}{\min\{R'(\alpha|\underline{i}, x) : \alpha \in [0, \tilde{\alpha}'']\}} \right] \in (0, 1)$$

and $T_\varepsilon = 4\lceil \log(\varepsilon^{-1})/\log(\bar{\kappa}^{-1})\rceil$ *with* $\varepsilon > 0$. *For any* $\varepsilon > 0$ *sufficiently small, we have* $\alpha_{T_\varepsilon}(u) \leq \varepsilon$ *for all* $u \in [0, \bar{R}]$.

Proof. See Appendix B.2. □

Choose $\varepsilon \in (0,1)$ sufficiently small (according to Lemma A.2.2) and pick any $\phi \in \mathscr{H}_S$ such that $\|\phi - \lambda^{-1}\|_{[0,\bar{u}],\infty} \geq \varepsilon$. Consider any $i \in \mathcal{I}\backslash\{\underline{i}\}$. By the compatibility condition, we have

$$b(\alpha|i,x) - b(\alpha|\underline{i},x) = \lambda^{-1}[R(\alpha|\underline{i},x)] - \lambda^{-1}[R(\alpha|i,x)]$$

for every $\alpha \in [0,1]$, so we can write

$$Q_\varepsilon(\phi|i) = \max_{\alpha\in[\varepsilon,1-\varepsilon]} \left|\tilde{\phi}[R(\alpha|\underline{i},x)] - \tilde{\phi}[R(\alpha|i,x)]\right| = \left\|\tilde{\phi}[R[(\cdot|\underline{i},x)] - \tilde{\phi}[R(\cdot|i,x)]\right\|_{[\varepsilon,1-\varepsilon],\infty}$$

with $\tilde{\phi}(u) \equiv \phi(u) - \lambda^{-1}(u)$. Note that $\|\tilde{\phi}\|_{[0,\bar{u}],\infty} \geq \varepsilon$. In what follows, I show that

$$Q_\varepsilon(\phi) \geq \left\|\tilde{\phi}[R[(\cdot|\underline{i},x)] - \tilde{\phi}[R(\cdot|i,x)]\right\|_{[\varepsilon,1-\varepsilon],\infty} \geq \frac{\underline{c}_Q\varepsilon}{\log(\varepsilon^{-1})}, \tag{A.6}$$

where the constant $\underline{c}_Q > 0$ can be taken to be $\underline{c}_Q = \min\{\underline{c}'_R, \log(\bar{\kappa}^{-1})/16\}$. The rests of the proof is divided into two cases: $|\tilde{\phi}'(0)| \geq \varepsilon^{1/3}$ and $|\tilde{\phi}'(0)| < \varepsilon^{1/3}$. I show next that in either case (A.6) holds.

Case 1: Suppose that $|\tilde{\phi}'(0)| \geq \varepsilon^{1/3}$. Then,

$$\begin{aligned}
|\tilde{\phi}[R(\varepsilon^{1/2}|\underline{i},x)] - \tilde{\phi}[R(\varepsilon^{1/2}|i,x)]| &= |\tilde{\phi}'(u^*_\varepsilon)|[R(\varepsilon^{1/2}|\underline{i},x) - R(\varepsilon^{1/2}|i,x)] \\
&= |\tilde{\phi}'(u^*_\varepsilon)|[R'(\varepsilon^*|\underline{i},x) - R'(\varepsilon^*|i,x)]\varepsilon^{1/2} \\
&\geq |\tilde{\phi}'(u^*_\varepsilon)|\underline{c}'_R\varepsilon^{1/2}
\end{aligned}$$

for some $u_\varepsilon^* \in [R(\varepsilon^{1/2}|i,x), R(\varepsilon^{1/2}|\underline{i},x)]$ and $\varepsilon^* \in [0,\varepsilon^{1/2}]$, where $\underline{c}_R' > 0$ was obtained in Lemma 1.2. Observe that

$$0 < u_\varepsilon^* \leq R(\varepsilon^{1/2}|\underline{i},x) \leq \bar{R}'\varepsilon^{1/2}$$

where $\bar{R}' \equiv \max\{R'(\alpha|i,x) : (\alpha,i,x)) \in [0,1] \times \mathcal{I} \times \mathcal{X}\}$ is strictly positive, finite, and independent of ε. By Lipschitz continuity of $\tilde{\phi}'(\cdot)$ with constant $2\bar{C}_{\mathscr{H}}$ from Definition 2, it follows that

$$|\tilde{\phi}'(0)| - |\tilde{\phi}'(u_\varepsilon^*)| \leq |\tilde{\phi}'(u_\varepsilon^*) - \tilde{\phi}'(0)| \leq 2\bar{C}_{\mathscr{H}}|u_\varepsilon^*| \leq 2\bar{C}_{\mathscr{H}}\bar{R}'\varepsilon^{1/2}.$$

As $|\tilde{\phi}'(0)| \geq \varepsilon^{1/3}$, these inequalities imply $\varepsilon^{1/3} - 2\bar{C}_{\mathscr{H}}\bar{R}'\varepsilon^{1/2} \leq |\tilde{\phi}'(u_\varepsilon^*)|$ and therefore

$$\left\|\tilde{\phi}[R[(\cdot|\underline{i},x)] - \tilde{\phi}[R(\cdot|i,x)]\right\|_{[\varepsilon,1-\varepsilon],\infty} \geq |\tilde{\phi}'(u_\varepsilon^*)|\underline{c}_R'\varepsilon^{1/2} \geq (\varepsilon^{1/3} - 2\bar{C}_{\mathscr{H}}\bar{R}'\varepsilon^{1/2})\underline{c}_R'\varepsilon^{1/2} \geq \underline{c}_R'\varepsilon.$$

Last inequality follows by taking $\varepsilon > 0$ sufficiently small so that $\varepsilon^{1/3} - 2\bar{C}_{\mathscr{H}}\bar{R}'\varepsilon^{1/2} > \varepsilon^{1/2}$.

Case 2: Suppose $|\tilde{\phi}'(0)| < \varepsilon^{1/3}$. Consider $\varepsilon > 0$ sufficiently small so that

$$\max\{1,\bar{R}'\}\varepsilon^{4/3} + \bar{C}_{\mathscr{H}}\varepsilon^2 < \varepsilon/2$$

and $\|R(\cdot|\underline{i},x)\|_{[0,1-\varepsilon],\infty} > \bar{u}$. Observe that there is $u^* \in (0,\bar{u}]$ such that $|\tilde{\phi}(u^*)| = \varepsilon$ due to continuity of $\tilde{\phi}(\cdot)$, $\tilde{\phi}(0) = 0$, and $\|\tilde{\phi}\|_{[0,\bar{u}],\infty} \geq \varepsilon$. Since $\tilde{\phi}(0) = 0$ and $|\tilde{\phi}'(0)| < \varepsilon^{1/3}$, we have $|\tilde{\phi}(u) - \tilde{\phi}'(0)u| \leq u^2\bar{C}_{\mathscr{H}}$ (error of Taylor series) and $|\tilde{\phi}(u)| < |\tilde{\phi}'(0)|\varepsilon + \bar{C}_{\mathscr{H}}\varepsilon^2 \leq \varepsilon^{4/3} + \bar{C}_{\mathscr{H}}\varepsilon^2 < \varepsilon/2$ for any $u \in [0,\varepsilon)$. So we must have that $u^* > \varepsilon$.

Now consider the sequence $\{\alpha_t(u^*) : t \in \mathbb{N}\}$ defined by (A.5) and denote $\alpha_t \equiv \alpha_t(u^*)$ for the rest of the proof. Define $T_\varepsilon^* = \min\{t \in \mathbb{N} : \alpha_t \leq \varepsilon\}$ and observe that $T_\varepsilon^* \leq T_\varepsilon = 4\lceil\log(\varepsilon^{-1})/\log(\bar{\kappa}^{-1})\rceil$ for any $u^* \in (0,\bar{u}]$ by Lemma A.2.2. Since $R(\alpha_0|\underline{i},x) = u^*$ and

$R(\alpha_t|\underline{i},x) = R(\alpha_{t-1}|i,x)$ for every $t \in \mathbb{N}$, by repeated triangular inequalities, it follows

$$\begin{aligned}\varepsilon < u^* = |\tilde{\phi}[R(\alpha_0|\underline{i},x)]| &\leq |\tilde{\phi}[R(\alpha_0|\underline{i},x)] - \tilde{\phi}[R(\alpha_1|\underline{i},x)]| + |\tilde{\phi}[R(\alpha_1|\underline{i},x)]| \\ &= |\tilde{\phi}[R(\alpha_0|\underline{i},x)] - \tilde{\phi}[R(\alpha_0|i,x)]| + |\tilde{\phi}[R(\alpha_1|\underline{i},x)]| \\ &\leq \sum_{t=0}^{T_\varepsilon^*-1} |\tilde{\phi}[R(\alpha_t|\underline{i},x)] - \tilde{\phi}[R(\alpha_t|i,x)]| + |\tilde{\phi}[R(\alpha_{T_\varepsilon^*}|\underline{i},x)]|.\end{aligned}$$

By construction of T_ε^*, $\alpha_t \in (\varepsilon, 1-\varepsilon]$ for all $t \leq T_\varepsilon^* - 1$ and therefore

$$\varepsilon - |\tilde{\phi}[R(\alpha_{T_\varepsilon^*}|\underline{i},x)]| \leq T_\varepsilon^* \left\| \tilde{\phi}[R(\cdot|\underline{i},x)] - \tilde{\phi}[R(\cdot|i,x)] \right\|_{[\varepsilon,1-\varepsilon],\infty}. \tag{A.7}$$

We have $T_\varepsilon^* \leq T_\varepsilon = 4\lceil \log(\varepsilon^{-1})/\log(\bar{\kappa}^{-1}) \rceil$. By construction of T_ε^*, we also have that $\alpha_{T_\varepsilon^*} \leq \varepsilon$ and consequently $R(\alpha_{T_\varepsilon^*}|\underline{i},x) < \bar{R}'\varepsilon$. Since $|\tilde{\phi}'(0)| < \varepsilon^{1/3}$ and $\varepsilon > 0$ is sufficiently small, we have that $|\tilde{\phi}[R(\alpha_{T_\varepsilon^*}|\underline{i},x)]| < \varepsilon/2$. From eq. (A.7) and the fact that $T_\varepsilon^* \leq T_\varepsilon \leq 8\log(\varepsilon^{-1})/\log(\bar{\kappa}^{-1})$, we obtain

$$\frac{\varepsilon}{2} \leq 8 \frac{\log(\varepsilon^{-1})}{\log(\bar{\kappa}^{-1})} \left\| \tilde{\phi}[R[(\cdot|\underline{i},x)] - \tilde{\phi}[R(\cdot|i,x)] \right\|_{[\varepsilon,1-\varepsilon],\infty},$$

which implies $[\log(\bar{\kappa}^{-1})\varepsilon]/[16\log(\varepsilon^{-1})] \leq \left\| \tilde{\phi}[R[(\cdot|\underline{i},x)] - \tilde{\phi}[R(\cdot|i,x)] \right\|_{[\varepsilon,1-\varepsilon],\infty}$.

A.3 Proof of Lemma 3

1. The proof of this item is mainly based on Lemma 1 of Marmer and Shneyerov (2012).[12] From part (c) of this lemma, we obtain

$$\|\hat{G}(\cdot|i,x) - G(\cdot|i,x)\|_{[\underline{b}(x),\bar{b}(i,x)],\infty} = O_P\left(\left[\frac{\log(L)}{Lh_X^D}\right]^{1/2} + h_X^{S+1}\right) = O_P(h_X^{S+1}); \tag{A.8}$$

[12] I emphasize that Marmer and Shneyerov (2012) uses a different notation: the smoothness of the valuation density is $R-1$, being R an integer greater than 1, while the density of covariates has smoothness R. In this paper, the densities of private values and covariates have smoothness S and $S+1$, respectively.

the last equality is due to the form of the bandwidth h_X; specifically, $[\log(L)/(Lh_X^D)]^{1/2}$ and h_X^{S+1} are of the same order by Assumption 4. As a result,

$$P[\hat{b}(h_\varepsilon|i,x) \leq \underline{b}(x)] = P\left[\inf_{b \in \mathbb{R}_{\geq 0}} \{\hat{G}(b|i,x) \geq h_\varepsilon\} \leq \underline{b}(x)\right] \leq P[\hat{G}[\underline{b}(x)|i,x] \geq h_\varepsilon] = o(1);$$

the last equality follows from eq. (A.8) and the fact $h_\varepsilon/h_X^{S+1} \to +\infty$ as $L \to +\infty$. In words, $\hat{G}[\underline{b}(x)|i,x]$ converges in probability to zero faster than h_ε. Symmetrically, it can be shown that $P[\hat{b}(1-h_\varepsilon|i,x) \geq \bar{b}(i,x)] = o(1)$. Hence, $\underline{b}(x) < \hat{b}(h_\varepsilon|i,x) \leq \hat{b}(1-h_\varepsilon|i,x) < \bar{b}(i,x)$ w.p.a.1. The rest of the proof follows exactly by the same arguments of eqs. (40)-(48) in the appendix of Marmer and Shneyerov (2012).

2. Pick any $(i,x) \in \mathcal{I} \times \text{interior}(\mathcal{X})$. Before proceeding, we state an auxiliary lemma. Write $(g^*)^{(S+1)}(b^*|i,x) = \partial^{S+1} g^*(b^*|i,x)/\partial^{S+1} b^*$ for $b^* \in (0,1]$.

Lemma A.3. *Under Assumptions 1-2,* $\int_0^1 [(g^*)^{(S+1)}(y|i,x)]^2 [y(1-y)]^{S+1} dy < +\infty$.

Proof. See Appendix B.3. □

To prove existence (w.p.a.1) of the conditional exponential series estimator $\hat{g}^*(\cdot|i,x)$, consider the information projection of $g^*(\cdot|i,x)$:

$$\tilde{g}^*(b^*|i,x) \equiv \frac{\exp\left[\sum_{1 \leq j \leq J_L} \tilde{\theta}_j(i,x)\pi_j(b^*)\right]}{\int_0^1 \exp\left[\sum_{1 \leq j \leq J_L} \tilde{\theta}_j(i,x)\pi_j(y)\right] dy},$$

where $b^* \in [0,1]$ and the coefficients $\{\tilde{\theta}_j(i,x) : 1 \leq j \leq J_L\}$ are obtained by solving

$$\frac{\int_0^1 \pi_j(y) \exp\left[\sum_{1 \leq j \leq J_L} \tilde{\theta}_j(i,x)\pi_j(y)\right] dy}{\int_0^1 \exp\left[\sum_{1 \leq j \leq J_L} \tilde{\theta}_j(i,x)\pi_j(y)\right] dy} = \mu_j(i,x) \tag{A.9}$$

for $j = 1, 2, \ldots, J_L$. In words, $\tilde{g}^*(\cdot|i,x)$ is characterized as the unique density in the exponential family that satisfies $\int_0^1 \pi_j(y)\tilde{g}^*(y|i,x)dy = \mu_j(i,x)$ for all $j = 1, 2, \ldots, J_L$.

Combining Lemma A.3 with Barron and Sheu (1991), Theorem 3 and eqs. (7.4)-(7.5), we obtain that there exists a unique solution to (A.9) when L is sufficiently large. Moreover, the information projection satisfies

$$\| \log[\tilde{g}^*(\cdot|i,x)] - \log[g^*(\cdot|i,x)]\|_{[0,1],\infty} = O\left(J_L^{-S}\right). \tag{A.10}$$

By Lemma 5 of Barron and Sheu (1991) and the next lemma., we can now affirm that there exists a unique solution $\{\hat{\theta}_j(i,x) : 1 \leq j \leq J_L\}$ to eqs. (8) w.p.a.1.

Lemma A.4. *Under Assumptions 1-4,*

$$\left\{\sum_{j=1}^{J_L} [\hat{\mu}_j(i,x) - \mu_j(i,x)]^2\right\}^{1/2} = O_P\left(J_L^{1/2} L^{-\frac{S+1}{2S+D+2}}\right).$$

I remark that the sequences $J_L^{1/2} L^{-\frac{S+1}{2S+D+2}}$ *and* $J_L^{-(S+1)}$ *are of the same order.*

Proof. See Appendix B.4. □

To obtain the desired rate of convergence, since $g^*(\cdot|i,x)$ is bounded away from 0, it suffices to show that

$$\| \log[\hat{g}^*(\cdot|i,x)] - \log[g^*(\cdot|i,x)]\|_{[0,1],\infty} = O_P\left(L^{-\frac{2S(S+1)}{(2S+3)(2S+D+2)}}\right).$$

By the triangle inequality, $\| \log[\hat{g}^*(\cdot|i,x)] - \log[g^*(\cdot|i,x)]\|_{[0,1],\infty}$ is bounded above by

$$\| \log[\hat{g}^*(\cdot|i,x)] - \log[\tilde{g}^*(\cdot|i,x)]\|_{[0,1],\infty} + \| \log[\tilde{g}^*(\cdot|i,x)] - \log[g^*(\cdot|i,x)]\|_{[0,1],\infty}.$$

The limiting behavior of the second term was already addressed in eq. (A.10). Regarding the first term, we have

$$\|\log[\hat{g}^*(\cdot|i,x) - \log[\tilde{g}^*(\cdot|i,x)]\|_{[0,1],\infty} \leq \bar{C}_{BS} J_L \left\{ \sum_{j=1}^{J_L} [\hat{\mu}_j(i,x) - \mu_j(i,x)]^2 \right\}^{1/2} \quad \text{(A.11)}$$

for some finite constant $\bar{C}_{BS} > 0$; see eq. (5.7) in Barron and Sheu (1991). The desired result then emerges from Lemma A.4 and the form of J_L.

3. By construction of $\hat{R}(\cdot|\cdot,\cdot)$, we have

$$|\hat{R}(\alpha|i,x) - R(\alpha|i,x)| = \frac{\alpha}{i-1} \left| \frac{1}{\hat{g}[\hat{b}(\alpha|i,x)|i,x]} - \frac{1}{g[b(\alpha|i,x)|i,x]} \right| \quad \text{(A.12)}$$

for any (α, i, x). Since $g(\cdot|i,x)$ is bounded away from zero, $\underline{b}(x) < \hat{b}(h_\varepsilon|i,x) \leq \hat{b}(1-h_\varepsilon|i,x) < \bar{b}(i,x)$ w.p.a.1, and $\hat{g}(\cdot|i,x)$ is uniformly consistent on $[\underline{b}(x), \bar{b}(i,x)]$, it follows that $\inf\{\hat{g}[\hat{b}(\alpha|i,x)|i,x] : \alpha \in [h_\varepsilon, 1-h_\varepsilon]\} \geq \underline{c}_g/2$ w.p.a.1; recall that $\underline{c}_g > 0$ has been obtained in subsection 2.2 and satisfies $g(\cdot|\cdot,\cdot) \geq \underline{c}_g$. Consequently, for any $\alpha \in [h_\varepsilon, 1-h_\varepsilon]$, the difference $\hat{R}(\alpha|i,x) - R(\alpha|i,x)$ can be bounded w.p.a.1 as follows:

$$\begin{aligned} |\hat{R}(\alpha|i,x) - R(\alpha|i,x)| &\leq \left| \frac{1}{\hat{g}[\hat{b}(\alpha|i,x)|i,x]} - \frac{1}{g[b(\alpha|i,x)|i,x]} \right| \\ &= \frac{1}{\bar{g}^2} |\hat{g}[\hat{b}(\alpha|i,x)|i,x] - g[b(\alpha|i,x)|i,x]| \\ &\leq \frac{4}{\underline{c}_g^2} |\hat{g}[\hat{b}(\alpha|i,x)|i,x] - g[b(\alpha|i,x)|i,x]| \end{aligned} \quad \text{(A.13)}$$

for some $\bar{g}$ that lies between $\hat{g}[\hat{b}(\alpha|i,x)|i,x]$ and $g[b(\alpha|i,x)|i,x]$, so it satisfies $\bar{g} \geq \underline{c}_g/2$ w.p.a.1. The first inequality is a consequence of eq. (A.12). The second equality follows after applying the mean value theorem to the function $f(t) = 1/t$: for any $\hat{t}, t > 0$, we have $(1/\hat{t}) - (1/t) = (-1/\bar{t}^2)(\hat{t} - t)$ for some $\bar{t}$ that lies between $\hat{t}$ and t. The third equality is due to $\bar{g}^2 \geq \underline{c}_g^2/4$.

To complete the proof, observe that

$$\begin{aligned}
|\hat{g}[\hat{b}(\alpha|i,x)|i,x] - g[b(\alpha|i,x)|i,x]| &\le |\hat{g}[\hat{b}(\alpha|i,x)|i,x] - g[\hat{b}(\alpha|i,x)|i,x]| \\
&\quad + |g[\hat{b}(\alpha|i,x)|i,x] - g[b(\alpha|i,x)|i,x]| \\
&\le \|\hat{g}(\cdot|i,x) - g(\cdot|i,x)\|_{[\underline{b}(x),\bar{b}(i,x)],\infty} \\
&\quad + \bar{C}_{g\prime}\left|\hat{b}(\alpha|i,x) - b(\alpha|i,x)\right| \\
&\le \|\hat{g}(\cdot|i,x) - g(\cdot|i,x)\|_{[\underline{b}(x),\bar{b}(i,x)],\infty} \qquad \text{(A.14)} \\
&\quad + \bar{C}_{g\prime}\left\|\hat{b}(\cdot|i,x) - b(\cdot|i,x)\right\|_{[h_\varepsilon,1-h_\varepsilon]}
\end{aligned}$$

where $\alpha \in [h_\varepsilon, 1-h_\varepsilon]$, $\bar{C}_{g\prime} \equiv \max\{|g'(b|i,x)| : i \in \mathcal{I}, (b,x) \in \mathscr{S}_{BX}(i)\} < +\infty$ by Observation 1.1.(a), and $g'(b|i,x) = \partial g(b|i,x)/\partial b$. After combining together inequalities (A.13) and (A.14) with items 1 and 2 of this lemma, we obtain the desired result. Since $\hat{g}(\cdot|i,x)$ converges slower than $\hat{b}(\cdot|i,x)$, we write

$$\|\hat{R}(\cdot|i,x) - R(\cdot|i,x)\|_{[h_\varepsilon,1-h_\varepsilon],\infty} = O_P\left(\|\hat{g}(\cdot|i,x) - g(\cdot|i,x)\|_{[\underline{b}(x),\bar{b}(i,x)],\infty}\right).$$

A.4 Proof of Lemma 4

Choose any $\phi \in \mathscr{H}_S$. Write

$$\hat{Q}(\phi) = \sum_{i\in\mathcal{I}\setminus\{\underline{i}\}} \left\|\hat{b}(\cdot|i,x) - \hat{b}(\cdot|\underline{i},x) + \phi[\hat{R}(\cdot|i,x)] - \phi[\hat{R}(\cdot|\underline{i},x)]\right\|_{[h_\varepsilon,1-h_\varepsilon],\infty}$$

and $Q_{h_\varepsilon}(\phi) = \sum_{i \in \mathcal{I} \setminus \{\underline{i}\}} \| b(\cdot|i,x) - b(\cdot|\underline{i},x) + \phi[R(\cdot|i,x)] - \phi[R(\cdot|\underline{i},x)] \|_{[h_\varepsilon, 1-h_\varepsilon],\infty}$. By the triangle inequality, the difference $\hat{Q}(\phi) - Q_{h_\varepsilon}(\phi)$ can be bounded by

$$\left|\hat{Q}(\phi) - Q_{h_\varepsilon}(\phi)\right| \le \sum_{i \in \mathcal{I} \setminus \{\underline{i}\}} \Big| \left\| \hat{b}(\cdot|i,x) - \hat{b}(\cdot|\underline{i},x) + \phi[\hat{R}(\cdot|i,x)] - \phi[\hat{R}(\cdot|\underline{i},x)] \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty} - \| b(\cdot|i,x) - b(\cdot|\underline{i},x) + \phi[R(\cdot|i,x)] - \phi[R(\cdot|\underline{i},x)] \|_{[h_\varepsilon, 1-h_\varepsilon],\infty} \Big| \tag{A.15}$$

As $\|\cdot\|_{[h_\varepsilon, 1-h_\varepsilon],\infty}$ is a norm, each summand on the right-hand side satisfies the inequalities

$$\begin{aligned}
&\Big| \left\| \hat{b}(\cdot|i,x) - \hat{b}(\cdot|\underline{i},x) + \phi[\hat{R}(\cdot|i,x)] - \phi[\hat{R}(\cdot|\underline{i},x)] \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty} \\
&\qquad - \| b(\cdot|i,x) - b(\cdot|\underline{i},x) + \phi[R(\cdot|i,x)] - \phi[R(\cdot|\underline{i},x)] \|_{[h_\varepsilon, 1-h_\varepsilon],\infty} \Big| \\
&\le \left\| \left\{ \hat{b}(\cdot|i,x) - \hat{b}(\cdot|\underline{i},x) + \phi[\hat{R}(\cdot|i,x)] - \phi[\hat{R}(\cdot|\underline{i},x)] \right\} \right. \\
&\qquad \left. - \{ b(\cdot|i,x) - b(\cdot|\underline{i},x) + \phi[R(\cdot|i,x)] - \phi[R(\cdot|\underline{i},x)] \} \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty} \\
&\le \left\| \hat{b}(\cdot|i,x) - b(\cdot|i,x) \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty} + \left\| \hat{b}(\cdot|\underline{i},x) - b(\cdot|\underline{i},x) \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty} \qquad \text{(A.16)} \\
&\qquad + \left\| \phi[\hat{R}(\cdot|i,x)] - \phi[R(\cdot|i,x)] \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty} + \left\| \phi[\hat{R}(\cdot|\underline{i},x)] - \phi[R(\cdot|\underline{i},x)] \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty}.
\end{aligned}$$

Observe that $\left|\phi[\hat{R}(\alpha|i,x)] - \phi[R(\alpha|i,x)]\right| = \phi'\left(\tilde{R}\right) \left|\hat{R}(\alpha|i,x) - R(\alpha|i,x)\right|$ for some $\tilde{R} > 0$ between $\hat{R}(\alpha|i,x)$ and $R(\alpha|i,x)$, and since $0 \le \phi'(\cdot) \le 1$, it follows

$$\left\| \phi[\hat{R}(\cdot|i,x)] - \phi[R(\cdot|i,x)] \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty} \le \left\| \hat{R}(\cdot|i,x) - R(\cdot|i,x) \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty}. \tag{A.17}$$

Combining together inequalities (A.15)-(A.17), $\left|\hat{Q}(\phi) - Q_{h_\varepsilon}(\phi)\right|$ can be bounded by

$$\begin{aligned}
&\sum_{i \in \mathcal{I} \setminus \{\underline{i}\}} \left\| \hat{b}(\cdot|i,x) - b(\cdot|i,x) \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty} + \left\| \hat{b}(\cdot|\underline{i},x) - b(\cdot|\underline{i},x) \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty} \\
&\quad + \sum_{i \in \mathcal{I} \setminus \{\underline{i}\}} \left\| \hat{R}(\cdot|i,x) - R(\cdot|i,x) \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty} + \left\| \hat{R}(\cdot|\underline{i},x) - R(\cdot|\underline{i},x) \right\|_{[h_\varepsilon, 1-h_\varepsilon],\infty}.
\end{aligned}$$

This upper bound does not depend on ϕ, hence the desired conclusion emerges from items 1 and 3 of Lemma 3.

A.5 Proof of Theorem 1

To simplify the exposition, I set $\gamma_\varepsilon = 1$. Pick any constant $\bar{C} > 0$ and observe

$$\begin{aligned} P\left\{r_L^* \left\|\hat{\lambda}^{-1} - \lambda^{-1}\right\|_{[0,\bar{u}],\infty} \geq \bar{C}\right\} &= P\left\{\left\|\hat{\lambda}^{-1} - \lambda^{-1}\right\|_{[0,\bar{u}],\infty} \geq \frac{\bar{C}}{r_L^*}\right\} \\ &\leq P\left\{Q_{1/r_L^*}(\hat{\lambda}^{-1}) \geq \frac{\underline{c}_Q \bar{C}}{r_L^* \log(r_L^*)}\right\} \\ &\leq P\left\{Q_{h_\varepsilon}(\hat{\lambda}^{-1}) \geq \frac{\underline{c}_Q \bar{C}}{r_L^* \log(r_L^*)}\right\} \\ &= P\left\{Q_{h_\varepsilon}(\hat{\lambda}^{-1}) \geq \underline{c}_Q \bar{C} L^{-\frac{2S(S+1)}{(2S+3)(2S+D+2)}}\right\}. \end{aligned} \tag{A.18}$$

The first equality is trivial. The second inequality employs Lemma 2 with $\phi(\cdot) = \hat{\lambda}^{-1}(\cdot)$ and $\varepsilon = 1/r_L^*$. The third one follows by $h_\varepsilon \leq 1/r_L^*$ (Assumption 4), which implies $Q_{h_\varepsilon} \geq Q_{1/r_L^*}$. The last equality combines together the definition r_L^* and the fact that $\varphi^{-1}(\cdot)\log[\varphi^{-1}(\cdot)]$ is equal to the identity function.

To complete the proof, since $\underline{c}_Q > 0$ and $\bar{C}$ can be taken to arbitrarily large, it suffices to show that $Q_{h_\varepsilon}(\hat{\lambda}^{-1}) = O_P\left(L^{-\frac{2S(S+1)}{(2S+3)(2S+D+2)}}\right)$. For this purpose, observe that there exists a sequence of functions $\{\lambda_{(L)}^{-1} \in \mathscr{H}^{(L)} : L \in \mathbb{N}\}$ satisfying

$$\|\lambda_{(L)}^{-1} - \lambda^{-1}\|_{[0,\bar{R}],\infty} = O\left(L^{-\frac{2S(S+1)}{(2S+3)(2S+D+2)}}\right) \tag{A.19}$$

by Assumption 5. Then, consider the inequalities

$$\begin{aligned}
Q_{h_\varepsilon}(\hat{\lambda}^{-1}) &\leq \hat{Q}(\hat{\lambda}^{-1}) + |\hat{Q}(\hat{\lambda}^{-1}) - Q_{h_\varepsilon}(\hat{\lambda}^{-1})| \\
&\leq \hat{Q}(\lambda_{(L)}^{-1}) + \sup_{\phi \in \mathscr{H}_S} \left|\hat{Q}(\phi) - Q_{h_\varepsilon}(\phi)\right| \\
&\leq Q_{h_\varepsilon}(\lambda_{(L)}^{-1}) + 2 \sup_{\phi \in \mathscr{H}_S} \left|\hat{Q}(\phi) - Q_{h_\varepsilon}(\phi)\right|.
\end{aligned}$$

The first inequality follows by the triangle inequality. The second by construction of $\hat{\lambda}^{-1}(\cdot)$, which minimizes $\hat{Q}(\cdot)$, and definition of supremum. The third is due to

$$\hat{Q}(\lambda_{(L)}^{-1}) \leq \left|\hat{Q}(\lambda_{(L)}^{-1}) - Q_{h_\varepsilon}(\lambda_{(L)}^{-1})\right| + Q_{h_\varepsilon}(\lambda_{(L)}^{-1}) \leq \sup_{\phi \in \mathscr{H}_S} |\hat{Q}(\phi) - Q_{h_\varepsilon}(\phi)| + Q_{h_\varepsilon}(\lambda_{(L)}^{-1}).$$

We already know that $\sup_{\phi \in \mathscr{H}_S} |\hat{Q}(\phi) - Q_{h_\varepsilon}(\phi)| = O_P\left(L^{-\frac{2S(S+1)}{(2S+3)(2S+D+2)}}\right)$ by Lemma 4, so to complete the proof, we have to show that $Q_{h_\varepsilon}(\lambda_{(L)}^{-1}) = O\left(L^{-\frac{2S(S+1)}{(2S+3)(2S+D+2)}}\right)$. Following the notation of Appendix A.4 and using the compatibility condition, $b(\alpha|i,x) - b(\alpha|\underline{i},x) = \lambda^{-1}[R(\alpha|\underline{i},x)] - \lambda^{-1}[R(\alpha|i,x)]$, we write

$$\begin{aligned}
Q_{h_\varepsilon}\left(\lambda_{(L)}^{-1}\right) &= \sum_{i \in \mathcal{I} \setminus \{\underline{i}\}} \| b(\cdot|i,x) - b(\cdot|\underline{i},x) \\
&\qquad + \lambda_{(L)}^{-1}[R(\cdot|i,x)] - \lambda_{(L)}^{-1}[R(\cdot|\underline{i},x)] \Big\|_{[h_\varepsilon, 1-h_\varepsilon], \infty} \\
&= \sum_{i \in \mathcal{I} \setminus \{\underline{i}\}} \Big\| \lambda^{-1}[R(\cdot|\underline{i},x)] - \lambda^{-1}[R(\cdot|i,x)] \\
&\qquad + \lambda_{(L)}^{-1}[R(\cdot|i,x)] - \lambda_{(L)}^{-1}[R(\cdot|\underline{i},x)] \Big\|_{[h_\varepsilon, 1-h_\varepsilon], \infty}.
\end{aligned}$$

Then, the triangle inequality and condition (A.19) yields the desired result:

$$\begin{aligned} Q_{h_\varepsilon}\left(\lambda_{(L)}^{-1}\right) &\leq \sum_{i\in\mathcal{I}\backslash\{\underline{i}\}} \left\| \lambda^{-1}[R(\cdot|\underline{i},x)] - \lambda_{(L)}^{-1}[R(\cdot|\underline{i},x)] \right\|_{[h_\varepsilon,1-h_\varepsilon],\infty} \\ &\quad + \sum_{i\in\mathcal{I}\backslash\{\underline{i}\}} \left\| \lambda^{-1}[R(\cdot|i,x)] - \lambda_{(L)}^{-1}[R(\cdot|i,x)] \right\|_{[h_\varepsilon,1-h_\varepsilon],\infty} \\ &\leq 2[\#(\mathcal{I})-1] \left\| \lambda^{-1}(\cdot) - \lambda_{(L)}^{-1}(\cdot) \right\|_{[0,\bar{R}],\infty} = O\left(L^{-\frac{2S(S+1)}{(2S+3)(2S+D+2)}} \right). \end{aligned}$$

A.6 Proof of Proposition 1

1. The next auxiliary lemma establishes the uniform consistency of $\hat{\lambda}(\cdot)$.

Lemma A.5. *Under Assumptions 1-5,* $r_L^*\|\hat{\lambda}(\cdot) - \lambda(\cdot)\|_{[0,\bar{y}],\infty} = O_P(1)$.

Proof. See Appendix B.5. □

To prove the second statement, just note that there is constant $\bar{C}_U > 0$ such that $\|\hat{U}(\cdot) - U(\cdot)\|_{[0,\bar{y}],\infty} \leq \bar{C}_U \|\hat{\lambda}(\cdot) - \lambda(\cdot)\|_{[0,\bar{y}],\infty}$.

2. This part follows closely the proof of Theorem 3 in Guerre et al. (2000). Pick any any inner compact subset $\mathcal{C} \subset \mathscr{S}_{VX}$ and consider the unfeasible kernel estimator

$$\tilde{f}_{VX}(v,x) = \frac{1}{Lh_f^{D+1}} \sum_{l=1}^{L} \frac{1}{I_l} \sum_{p=1}^{I_l} k\left(\frac{v - V_{pl}}{h_f}\right) K\left(\frac{x - X_l}{h_f}\right).$$

By standard arguments, we know that

$$\|\tilde{f}_{VX}(v,x) - f_{VX}(v,x)\|_{\mathcal{C},\infty} = O_P\left(h_f^{S+1} + \sqrt{\frac{\log(L)}{Lh_f^D}} \right),$$

so it suffices to prove

$$\|\hat{f}_{VX}(v,x) - \tilde{f}_{VX}(v,x)\|_{\mathcal{C},\infty} = O_P\left(\frac{1}{r_L^* h_f}\right).$$

We observe that $\hat{V}_{pl}$ inherits the rate of $\hat{\lambda}^{-1}(\cdot)$; more specifically,

$$\sup_{p,l} \mathbb{1}\left\{(V_{pl}, X_l) \in \mathcal{C}'\right\} |\hat{V}_{pl} - V_{pl}| = O_P\left(\frac{1}{r_L^*}\right)$$

for any inner compact subset $\mathcal{C}' \subset \mathscr{S}_{VX}$. This result is based on arguments similar to that of Proposition 3.(ii) in Guerre et al. (2000). Finally, the desired result emerges by implementing the inequalities of eq. (23) in this reference.

B Appendix: Proofs of Auxiliary Lemmas

B.1 Proof of Lemma A.1

As a starting point, recall from Subsection 2.2 that $g(\cdot|i,\cdot) \geq \underline{c}_g > 0$ on its support $\mathscr{S}_{BX}(i)$. Moreover, we know that $g(\cdot|i,\cdot)$ admits S continuous partial derivatives on $\mathscr{S}_{BX}(i)$ by Observation 1.1.(a). So first, as $b'(\alpha|i,x) = 1/g[b(\alpha|i,x)|i,x]$, we have that $b(\cdot|i,\cdot)$ admits $S+1$ continuous partial derivatives on $[0,1] \times \mathcal{X}$. Second, since $g^{(S+1)}(\cdot|i,x)$ is continuous on $(\underline{b}(x), \bar{b}(i,x)]$ by Observation 1.1.(b), $b^{(S+2)}(\cdot|i,x)$ is continuous on $(0,1]$. Third, as $g(\cdot|i,\cdot)$ is continuous on $\mathscr{S}_{BX}(i)$ and this set is compact, there is a finite constant $\bar{C}_g > 0$ such that $g(\cdot|i,\cdot) \leq \bar{C}_g$ on $\mathscr{S}_{BX}(i)$. Then,

$$0 < \frac{1}{\bar{C}_g} \leq b'(\alpha|i,x) = \frac{1}{g[b(\alpha|i,x)|i,x]} \leq \frac{1}{\underline{c}_g} < +\infty$$

for all $(\alpha, x) \in [0,1] \times \mathcal{X}$.

The last statement, $b'(0|i_1,x)/(i_1-1) > b'(0|i_2,x)/(i_2-1)$ for all $(i_1,i_2) \in \mathcal{I}^*$ and $x \in \mathcal{X}$, is derived from Observation 1.3.(b). Pick any $(i_1,i_2) \in \mathcal{I}^*$ and $x \in \mathcal{X}$. Observe that $\upsilon(\alpha|x) \equiv \xi_{i_1}[b(\alpha|i_1,x),x] = \xi_{i_2}[b(\alpha|i_2,x),x]$ for all $\alpha \in [0,1]$. Moreover, $\upsilon'(0|x) = \xi'_{i_j}[\underline{\upsilon}(x),x] \times b'(0|i_j,x) > 0$ because $b'(\cdot|\cdot,\cdot)$ is bounded away from 0 and $\xi'_{i_j}[\underline{\upsilon}(x),x] =$

$1 + 1/[\lambda'(0)(i_j - 1)] > 0$, where $j = 1, 2$ and $\xi'_{i_j}(v, x) = \partial\xi_{i_j}(v, x)/\partial v$. Since $\xi_{i_j}(\cdot, x)$ is strictly increasing, it follows that $b(\alpha|i_j, x) = \xi^{-1}_{i_j}[v(\alpha|x), x]$, where $\xi^{-1}_{i_j}(\cdot, x)$ stands for the inverse of $\xi_{i_j}(\cdot, x)$. After taking the derivative with respect to α and evaluating at $\alpha = 0$, we obtain

$$b'(0|i_j, x) = (\xi^{-1}_{i_j})'[\underline{v}(x), x] \times v'(0|x). \tag{B.20}$$

At the same time, we have that

$$(\xi^{-1}_{i_j})'[\underline{v}(x), x] = \frac{1}{\xi'_{i_j}[\underline{v}(x), x]} = \frac{\lambda'(0)(i_j - 1)}{\lambda'(0)(i_j - 1) + 1}, \tag{B.21}$$

so combining together (B.20) and (B.21) yields

$$\frac{b'(0|i_1, x)}{i_1 - 1} = \frac{\lambda'(0)v'(0|x)}{\lambda'(0)(i_1 - 1) + 1} > \frac{\lambda'(0)v'(0|x)}{\lambda'(0)(i_2 - 1) + 1} = \frac{b'(0|i_2, x)}{i_2 - 1}.$$

The strict inequality is due to $i_2 > i_1$, $\lambda'(0) > 0$, and $v'(0|x) > 0$.

B.2 Proof of Lemma A.2

Before proceeding, note that the sequence $\{\alpha_t(\cdot) : t \in \mathbb{N}\}$ is well-defined because $R(0|\underline{i}, x) = 0$, $R(\cdot|\underline{i}, x)$ is continuous, and by definition of $\bar{R}$. Observe also that parts 1 and 2 of Lemma 1 hold for $\alpha \in [0, \tilde{\alpha}'']$ as $\tilde{\alpha}'' \le \tilde{\alpha}' \le \tilde{\alpha}$.

1. Define $\underline{\Delta R} = \min\{R(\alpha|\underline{i}, x) - R(\alpha|i, x) : (\alpha, i, x) \in [\tilde{\alpha}'', 1] \times \mathcal{I}\backslash\{\underline{i}\} \times \mathcal{X}\} > 0$ and $\tilde{T} = \lceil \bar{R}/\underline{\Delta R} \rceil + 1$. Note that $\tilde{T}$ is independent of (u, x). Pick any $u \in (0, \bar{R}]$ and write $\alpha_t \equiv \alpha_t(u)$ to simplify the notation. By construction and since $R(\cdot|\underline{i}, x) > R(\cdot|i, x)$ on $(0, 1]$, we have that $\alpha_{t+1} < \alpha_t$ for all $t \in \mathbb{N}$. On the one hand, if $\alpha_{\tilde{T}} \le \tilde{\alpha}''$, the desired result follows immediately. On the other hand, if $\alpha_{\tilde{T}} > \tilde{\alpha}''$, we have $\alpha_t > \tilde{\alpha}''$ for all $t \le \tilde{T}$

because $(\alpha_t)_t$ is strictly decreasing. As a result,

$$\begin{aligned}
\bar{R} &\geq R(\alpha_0|\underline{i},x) - R(\alpha_{\tilde{T}}|i,x) \\
&= [R(\alpha_0|\underline{i},x) - R(\alpha_0|i,x)] + R(\alpha_1|\underline{i},x) - R(\alpha_{\tilde{T}}|i,x) \\
&= \sum_{t=0}^{\tilde{T}} [R(\alpha_t|\underline{i},x) - R(\alpha_t|i,x)] \\
&\geq (\tilde{T}+1) \times \underline{\Delta R},
\end{aligned}$$

which is a contradiction by construction of $\tilde{T}$. The first line follows immediately by definition of $\bar{R}$ and $R(\alpha_{\tilde{T}}|i,x) \geq 0$. The second is due to $R(\alpha_1|\underline{i},x) = R(\alpha_0|i,x)$. The third employs an inductive argument; specifically, $R(\alpha_t|\underline{i},x) = R(\alpha_{t-1}|i,x)$ for all t. The fourth line follows by definition of $\underline{\Delta R}$ and $\alpha_t > \tilde{\alpha}''$ for all $t \leq \tilde{T}$.

2. Pick any arbitrary $u \in (0,\bar{R}]$ and $\varepsilon \in (0,1)$ small enough so that $T_\varepsilon > 2\tilde{T}$, where $\tilde{T}$ was obtained in the previous item. Write $\alpha_t \equiv \alpha_t(u)$ to simplify the notation. I will next show that $\alpha_{T_\varepsilon} \leq \varepsilon$. By item 1 of this lemma and the way $\varepsilon > 0$ was chosen above, we have that $\alpha_t \leq \tilde{\alpha}''$ for all $t \geq T_\varepsilon/2 > \tilde{T}$. Since $R(\cdot|\underline{i},x)$ is strictly increasing on $[0,\tilde{\alpha}'']$ (Lemma 1.1), α_t can be defined as $\alpha_t = R^{-1}[R(\alpha_{t-1}|i,x)|\underline{i},x]$ whenever $t \geq T_\varepsilon/2$. As a result,

$$\begin{aligned}
\alpha_t &= R^{-1}[R(\alpha_{t-1}|i,x)|\underline{i},x] \\
&\leq R^{-1}[\alpha_{t-1}\max\{R'(\alpha|i,x) : \alpha \in [0,\tilde{\alpha}'']\}|\underline{i},x] \\
&\leq \alpha_{t-1}\frac{\max\{R'(\alpha|i,x) : \alpha \in [0,\tilde{\alpha}'']\}}{\min\{R'(\alpha|\underline{i},x) : \alpha \in [0,\tilde{\alpha}'']\}} \leq \alpha_{t-1}\bar{\kappa}.
\end{aligned}$$

The second line follows by $|R(\alpha|i,x)| \leq \alpha\|R'(\cdot|i,x)\|_{[0,\tilde{\alpha}''],\infty}$ for any $\alpha \in [0,\tilde{\alpha}'']$ and the third employs a similar argument. Since $\alpha_{T_\varepsilon/2} \leq \tilde{\alpha}''$ and $\alpha_t \leq \bar{\kappa}\alpha_{t-1}$ for all $t \geq T_\varepsilon/2$, it follows by induction that $\alpha_\tau \leq \bar{\kappa}^{\tau - T_\varepsilon/2}\tilde{\alpha}''$ for any integer $\tau \geq T_\varepsilon/2$. So taking $\tau = T_\varepsilon$, we obtain

$$\alpha_{T_\varepsilon} \leq \bar{\kappa}^{T_\varepsilon - T_\varepsilon/2}\tilde{\alpha}'' = \bar{\kappa}^{T_\varepsilon/2}\tilde{\alpha}'' \leq \bar{\kappa}^{2\log(\varepsilon^{-1})/\log(\bar{\kappa}^{-1})}\tilde{\alpha}'' \leq \varepsilon.$$

The last inequality follows by construction of T_ε. More specifically,

$$\begin{aligned}
\bar{\kappa}^{2\log(\varepsilon^{-1})/\log(\bar{\kappa}^{-1})}\tilde{\alpha}'' \leq \varepsilon \quad &\leftrightarrow \quad \frac{2\log(\varepsilon^{-1})}{\log(\bar{\kappa}^{-1})}\log(\bar{\kappa}) + \log(\tilde{\alpha}'') \leq \log(\varepsilon) \\
&\leftrightarrow \quad \frac{2\log(\varepsilon)}{\log(\bar{\kappa})}\log(\bar{\kappa}) + \log(\tilde{\alpha}'') \leq \log(\varepsilon) \\
&\leftrightarrow \quad 2\log(\varepsilon) + \log(\tilde{\alpha}'') \leq \log(\varepsilon) \\
&\leftrightarrow \quad (2-1)\log(\varepsilon) + \log(\tilde{\alpha}'') \leq 0;
\end{aligned}$$

the last inequality holds because $0 < \varepsilon, \tilde{\alpha}'' \leq 1$.

B.3 Proof of Lemma A.3

We prove the result for $S = 1$, the general case $S \geq 1$ follows by the same arguments. We remark that the results of Observation 1.1 also hold for $g^*(\cdot|i,x)$. Let $G^*(\cdot|i,x)$ be the conditional c.d.f. associated with $g^*(\cdot|i,x)$. We have that

$$\frac{\partial\left[\frac{G^*(y|i,x)}{g^*(y|i,x)}\right]}{\partial y} = 1 - \frac{G^*(y|i,x)}{g^*(y|i,x)^2}g^{*\prime}(y|i,x)$$

with $y \in [0,1]$ and

$$\frac{\partial^2\left[\frac{G^*(y|i,x)}{g^*(y|i,x)}\right]}{\partial^2 y} = -\frac{g^{*\prime}(y|i,x)}{g^*(y|i,x)} + 2\frac{G^*(y|i,x)}{g^*(y|i,x)^3}g^{*\prime}(y|i,x)^2 - \frac{G^*(y|i,x)}{g^*(y|i,x)^2}(g^*)^{(2)}(y|i,x),$$

so we can write

$$(g^*)^{(2)}(y|i,x) \times G^*(y|i,x) = -g^*(y|i,x)^2 \frac{\partial^2 \left[\frac{G^*(y|i,x)}{g^*(y|i,x)}\right]}{\partial^2 y} - g^{*\prime}(y|i,x) g^*(y|i,x) + 2G^*(y|i,x) \frac{g^{*\prime}(y|i,x)^2}{g^*(y|i,x)}. \tag{B.22}$$

Since $g^*(\cdot|i,x)$ is bounded away from 0 on $[0,1]$, there is a finite constant $\underline{c}_G^* > 0$ such $\underline{c}_G^* y \le G^*(y|i,x)$ for every $y \in [0,1]$. Thus,

$$[(g^*)^{(2)}(y|i,x)]^2 [y(1-y)]^2 \le \frac{1}{\underline{c}_G^{*2}} [(g^*)^{(2)}(y|i,x) \times G^*(y|i,x)]^2$$

for any $y \in (0,1)$. The desired result follows by noting that the right-hand side is integrable over the interval (0,1) because expression (B.22) is bounded in absolute value; $\lim_{y \downarrow 0} \partial^2 [G^*(y|i,x)/g^*(y|i,x)]/\partial^2 y$ exists and is finite by Observation 1.1.(c).

B.4 Proof of Lemma A.4

Along the lines of Li and Racine (2007), pp. 61-63, define

$$\hat{m}_j(i,x) = \hat{f}_{IX,h_\mu}(i,x) \left[\hat{\mu}_j(i,x) - \mu_j(i,x)\right]$$

for $j \in \mathbb{N}$, where $\hat{f}_{IX,h_\mu}(i,x)$ is a kernel estimator of $f_{IX}(i,x)$ using bandwidth h_μ:

$$\hat{f}_{IX,h_\mu}(i,x) = \frac{1}{L h_\mu^D} \sum_{l=1}^{L} \mathbb{1}\{I_l = i\} K\left(\frac{x - X_l}{h_\mu}\right)$$

Note that $\left\{\sum_{j=1}^{J_L} \left[\hat{\mu}_j(i,x) - \mu_j(i,x)\right]^2\right\}^{1/2} = \left[\sum_{j=1}^{J_L} \hat{m}_j(i,x)^2\right]^{1/2} / \hat{f}_{IX,h_\mu}(i,x)$. Pick an arbitrary large constant $\bar{C} > 0$ and consider the expression

$$P\left(\frac{1}{\hat{f}_{IX,h_\mu}(i,x)} \left[\sum_{j=1}^{J_L} \hat{m}_j(i,x)^2\right]^{1/2} \ge \bar{C} J_L^{1/2} L^{-\frac{S+1}{2S+2+D}}\right). \tag{B.23}$$

By Assumption 1.3 and since $\hat{f}_{IX,h_\mu}(i,x)$ is a consistent estimator of $f_{IX}(i,x)$, there is a constant $\underline{c}_{IX} > 0$ such that the event $\hat{f}_{IX,h_\mu}(i,x) > \underline{c}_{IX}$ occurs w.p.a.1. Hence, expression (B.23) can be bounded above by

$$P\left(\left[\sum_{j=1}^{J_L} \hat{m}_j(i,x)^2\right]^{1/2} > \underline{c}_{IX} \frac{\bar{C} J_L^{1/2}}{L^{\frac{S+1}{2S+2+D}}}\right) + P\left(\hat{f}_{IX,h_\mu}(i,x) \leq \underline{c}_{IX}\right).$$

As the second term converges to 0, the rest of the proof focuses only on the first one.

Observe that

$$\hat{\mu}_j(i,x) - \mu_j(i,x) = \sum_{l=1}^{L} \frac{1}{I_l} \sum_{p=1}^{I_l} \left[\pi_j\left(B_{lp}^*\right) - \mu_j(i,x)\right] \omega_l(i,x)$$

because of $\sum_{l=1}^{L}(1/I_l)\sum_{p=1}^{I_l}\omega_l(i,x) = 1$. Then, write $\hat{m}_j(i,x) = \hat{m}_{j,1}(i,x) + \hat{m}_{j,2}(i,x)$ with

$$\begin{aligned}
\hat{m}_{j,1}(i,x) &= \frac{1}{Lh_\mu^D} \sum_{l=1}^{L} \frac{1}{I_l} \sum_{p=1}^{I_l} \left[\mu_j(I_l,X_l) - \mu_j(i,x)\right] \mathbb{1}\{I_l = i\} K\left(\frac{x - X_l}{h_\mu}\right) \\
&= \frac{1}{Lh_\mu^D} \sum_{l=1}^{L} \left[\mu_j(I_l,X_l) - \mu_j(i,x)\right] \mathbb{1}\{I_l = i\} K\left(\frac{x - X_l}{h_\mu}\right), \\
\hat{m}_{j,2}(i,x) &= \frac{1}{Lh_\mu^D} \sum_{l=1}^{L} \frac{1}{I_l} \sum_{p=1}^{I_l} W_{j,lp} \mathbb{1}\{I_l = i\} K\left(\frac{x - X_l}{h_\mu}\right),
\end{aligned}$$

$W_{j,lp} = \pi_j\left(\tilde{B}_{lp}^*\right) - \mu_j(I_l,X_l)$, and $\mu_j(I_l,X_l) = E\left[\pi_j\left(\tilde{B}_{lp}^*\right) \middle| I_l, X_l\right]$. By a standard triangle inequality, we can bound

$$P\left(\left[\sum_{j=1}^{J_L} \hat{m}_j(i,x)^2\right]^{1/2} > \underline{c}_{IX} \frac{\bar{C} J_L^{1/2}}{L^{\frac{S+1}{2S+2+D}}}\right) \leq \sum_{\tau=1}^{2} P\left(\left[\sum_{j=1}^{J_L} \hat{m}_{j,\tau}(i,x)^2\right]^{1/2} \geq \frac{\underline{c}_{IX}}{2} \frac{\bar{C} J_L^{1/2}}{L^{\frac{S+1}{2S+2+D}}}\right).$$

For each $\tau = 1,2$, observe that

$$\begin{aligned}
P\left(\left[\sum_{j=1}^{J_L} \hat{m}_{j,\tau}(i,x)^2\right]^{1/2} \geq \frac{\underline{c}_{IX}}{2} \frac{\bar{C} J_L^{1/2}}{L^{\frac{S+1}{2S+2+D}}}\right) &\leq \frac{4L^{\frac{2(S+1)}{2S+2+D}}}{(\underline{c}_{IX}\bar{C})^2 J_L} \sum_{j=1}^{J_L} E\left[\hat{m}_{j,\tau}(i,x)^2\right] \\
&\leq \frac{4L^{\frac{2(S+1)}{2S+2+D}}}{(\underline{c}_{IX}\bar{C})^2} \sup_{j \in \mathbb{N}} E\left[\hat{m}_{j,\tau}(i,x)^2\right].
\end{aligned}$$

In what follows, for each $\tau = 1, 2$, we show that there exists a constant $\bar{C}_{\mu,\tau} > 0$ such that $\sup\{E[\hat{m}_{j,\tau}(i,x)^2] : j \in \mathbb{N}\} \leq \bar{C}_{\mu,\tau} L^{\frac{-2(S+1)}{2S+2+D}}$. We employ Li and Racine (2007)'s arguments. For simplicity and without loss of generality, the rest of the proof assumes $\gamma_X = 1$ and that L is sufficiently large so that $x + h_\mu t \in \text{interior}(\mathcal{X})$ for every $t \in [-1,1]^D$.

Consider first $\hat{m}_{j,1}(i,x)$ and write $E[\hat{m}_{j,1}(i,x)^2] = E[\hat{m}_{j,1}(i,x)]^2 + \text{var}[\hat{m}_{j,1}(i,x)]$. Let $[-\underline{c}_\wp, \underline{c}_\wp]^D$ be a rectangle such that $\underline{c}_\wp > 0$ and $x + y \in \text{interior}(\mathcal{X})$ for any $y \in [-\underline{c}_\wp, \underline{c}_\wp]^D$. Define the function $\wp_j(y) = [\mu_j(i, x+y) - \mu_j(i,x)] f_{IX}(i, x+y)$ for $y \in [-\underline{c}_\wp, \underline{c}_\wp]^D$, so we can write

$$\begin{aligned} E[\hat{m}_{j,1}(i,x)] &= \frac{1}{h_\mu^D} E\left\{[\mu_j(I_l, X_l) - \mu_j(i,x)] \mathbb{1}\{I_l = i\} K\left(\frac{x - X_l}{h_\mu}\right)\right\} \\ &= \frac{P(I_l = i)}{h_\mu^D} E\left\{[\mu_j(i, X_l) - \mu_j(i,x)] K\left(\frac{x - X_l}{h_\mu}\right) \middle| I_l = i\right\} \\ &= \frac{1}{h_\mu^D} \int_{\mathbb{R}^{\mathbb{D}}} [\mu_j(i,y) - \mu_j(i,x)] K\left(\frac{x-y}{h_\mu}\right) f_{IX}(i,y) dy \\ &= \int_{[-1,1]^D} \wp_j(h_\mu t) K(t) dt. \end{aligned}$$

By Assumption 1.3, $\wp_j(\cdot)$ admits $S+1$ continuously bounded partial derivatives on $[-\underline{c}_\wp, \underline{c}_\wp]^D$ for every $j \in \mathbb{N}$. Moreover,

$$\begin{aligned} |\wp_j(y)| &= \left|\int_0^1 \pi_j(t)[g^*(t|i, x+y) - g^*(t|i,x)]dt\right| f_{IX}(i, x+y) \\ &\leq \left\{\int_0^1 \pi_j(t)^2 dt\right\}^{1/2} \left\{\int_0^1 [g^*(t|i,x+y) - g^*(t|i,x)]^2 dt\right\}^{1/2} f_{IX}(i, x+y). \end{aligned}$$

Since $\int_0^1 \pi_j(t)^2 dt = 1$ for all j, $|\wp_j(\cdot)|$ is uniformly bounded across $j \in \mathbb{N}$. Proceeding in a similar manner, we can show that the partial derivatives of $\wp_j(\cdot)$ are also uniformly bounded across $j \in \mathbb{N}$. As the kernel $K(\cdot)$ is of order $S+1$, using a Taylor expansion for $\wp_j(h_\mu t)$, it can be shown that $|E[\hat{m}_{j,1}(i,x)]| \leq \bar{C}_\wp h_\mu^{S+1}$ for any $j \in \mathbb{N}$ and some constant $\bar{C}_\wp > 0$. This implies $E[\hat{m}_{j,1}(i,x)]^2 \leq \bar{C}_\wp^2 h_\mu^{2(S+1)}$ for every $j \in \mathbb{N}$; see Li and Racine (2007), eq. (2.8).

Regarding the variance of $\hat{m}_{j,1}(i,x)$, we can write

$$\begin{aligned}
\operatorname{var}\left[\hat{m}_{j,1}(i,x)\right] &= \frac{1}{Lh_\mu^{2D}}\operatorname{var}\left\{\left[\mu_j(I_l,X_l)-\mu_j(i,x)\right]\mathbb{1}\{I_l=i\}K\left(\frac{x-X_l}{h_\mu}\right)\right\} \\
&= \frac{1}{Lh_\mu^{2D}}E\left[\left\{\left[\mu_j(I_l,X_l)-\mu_j(i,x)\right]\mathbb{1}\{I_l=i\}K\left(\frac{x-X_l}{h_\mu}\right)\right\}^2\right] \\
&\quad -\frac{1}{Lh_\mu^{2D}}\left[E\left\{\left[\mu_j(I_l,X_l)-\mu_j(i,x)\right]\mathbb{1}\{I_l=i\}K\left(\frac{x-X_l}{h_\mu}\right)\right\}\right]^2.
\end{aligned}$$

The first term in the second equality can be written as

$$\frac{1}{Lh_\mu^D}\int_{[-1,1]^D}\left[\frac{\wp_j(h_\mu t)}{f_{IX}(i,x+h_\mu t)}K(t)\right]^2 dt = h_\mu^{2(S+1)}\int_{[-1,1]^D}\left[\frac{\wp_j(h_\mu t)}{f_{IX}(i,x+h_\mu t)}K(t)\right]^2 dt$$

by a change of variables and by the form of the bandwidth h_μ (Assumption 4); recall that here we are assuming $\gamma_X = 1$. The second term equals $\{E[\hat{m}_{j,1}(i,x)]\}^2/L$, so it can be bounded above by $\bar{C}_\wp^2 h_\mu^{2(S+1)}$. Combining together all previous results and by the form of the bandwidth, we obtain $\sup\{E\left[\hat{m}_{j,1}(i,x)^2\right] : j\in\mathbb{N}\} \le \bar{C}_{\mu,1}L^{\frac{-2(S+1)}{2S+2+D}}$ with

$$\bar{C}_{\mu,1} = 2\bar{C}_\wp^2 + \int_{[-1,1]^D}\left[\frac{\wp_j(h_\mu t)}{f_{IX}(i,x+h_\mu t)}K(t)\right]^2 dt < +\infty,$$

which is clearly independent of j.

Now consider $\hat{m}_{j,2}(i,x)$ and note $E\left[\hat{m}_{j,2}(i,x)\right]=0$ by the law of iterated expectations. Using standard arguments, we write

$$\begin{aligned}
E\left[\hat{m}_{j,2}(i,x)^2\right] &= \frac{1}{iLh_\mu^{2D}}E\left[W_{j,lp}^2\mathbb{1}\{I_l=i\}K\left(\frac{x-X_l}{h_\mu}\right)^2\right] \\
&= \frac{1}{iLh_\mu^{2D}}\int_{\mathbb{R}^D}\Sigma_j(i,y)K\left(\frac{x-y}{h_\mu}\right)^2 f_{IX}(i,y)dy \\
&= \frac{1}{iLh_\mu^{D}}\int_{[0,1]^D}\Sigma_j(i,x+h_\mu t)K\left(t\right)^2 f_{IX}(i,x+h_\mu t)dt
\end{aligned}$$

where $\Sigma_j(i,x) \equiv E\left(W_{j,lp}^2 | I_l = i, X_l = x\right) = E\left[\pi_j\left(\tilde{B}_{lp}^*\right)^2 | I_l = i, X_l = x\right] - \mu_j(i,x)^2$. Observe that

$$E\left[\pi_j\left(\tilde{B}_{lp}^*\right)^2 | I_l = i, X_l = x\right] = \int_0^1 \pi_j(y)^2 g^*(y|i,x)dy \leq \bar{C}_{B^*} \int_0^1 \pi_j(y)^2 dy = \bar{C}_{B^*},$$

for every (j,i,x), where $\bar{C}_{B^*} = \max\{g^*(y|i,x) : (y,i,x) \in [0,1]\times\mathcal{I}\times\mathcal{X}\} < +\infty$. Moreover,

$$|\mu_j(i,x)| = \left|\int_0^1 \pi_j(y) g^*(y|i,x)dy\right| \leq \left[\int_0^1 \pi_j(y)^2 dy\right]^{1/2} \left[\int_0^1 g^*(y|i,x)^2 dy\right]^{1/2} \leq \bar{C}_{B^*}.$$

by Cauchy-Schwarz inequality. By the form of the h_μ and as $\bar{C}_{B^*}$ is independent of (j,x), we obtain $\sup\{E\left[\hat{m}_{j,2}(i,x)^2\right] : j \in \mathbb{N}\} \leq \bar{C}_{\mu,2} L^{\frac{-2(S+1)}{2S+2+D}}$ with

$$\bar{C}_{\mu,2} = \left(\bar{C}_{B^*} + \bar{C}_{B^*}^2\right) \max_{(i,x)\in\mathcal{I}\times\mathcal{X}} f_{IX}(i,x) \int_{[0,1]^D} K(t)^2 dt < +\infty. \tag{B.24}$$

B.5 Proof of Lemma A.5

This proof follows closely the arguments of Matzkin (2003)'s Theorem 1. Pick any $y \in [0,\bar{y}]$. After applying the mean value theorem to $\lambda^{-1}(\cdot)$, we obtain

$$\left[\hat{\lambda}(y) - \lambda(y)\right] \times \left(\lambda^{-1}\right)'(\lambda^*) = \lambda^{-1}[\hat{\lambda}(y)] - \lambda^{-1}[\lambda(y)]$$

for some $\lambda^* \geq 0$ that lies between $\hat{\lambda}(y)$ and $\lambda(y)$. Since $\lambda(y) \leq \lambda(\bar{y}) < \bar{u}$, $\hat{\lambda}(y) \leq \hat{\lambda}(\bar{y}) \leq \bar{u}$ w.p.a.1, and $(\lambda^{-1})'(\lambda^*)$ is bounded away from zero on $[0,\bar{u}]$, there is a finite constant $\bar{C} > 0$ such that

$$|\hat{\lambda}(y) - \lambda(y)| \leq \bar{C}|\lambda^{-1}[\hat{\lambda}(y)] - \lambda^{-1}[\lambda(y)]| = |\lambda^{-1}[\hat{\lambda}(y)] - \hat{\lambda}^{-1}[\hat{\lambda}(y)]| \leq \|\hat{\lambda}^{-1} - \lambda^{-1}\|_{[0,\bar{u}],\infty}.$$

The equality follows by $\lambda^{-1}[\lambda(y)] = y = \hat{\lambda}^{-1}[\hat{\lambda}(y)]$, while the second inequality holds w.p.a.1 because $\hat{\lambda}(y) \leq \hat{\lambda}(\bar{y}) \leq \bar{u}$ w.p.a.1.

Acknowledgements

This article is developed from the first chapter of my Ph.D. dissertation at UCLA. I am grateful to Rosa L. Matzkin, Jinyong Hahn, and Ichiro Obara for their guidance and encouragement. For useful comments and detailed discussions, I thank Arie Beresteanu, Stephen Cosslett, Tong Li, Zhipeng Liao, JF Richards, Enrique Seira, Artyom Shneyerov, and Walter Sosa-Escudero. The editor, the associate editor, and three anonymous referees provided helpful comments that improve the quality of the paper. I am also grateful to the seminar participants at the institutions where this manuscript was presented. Financial support form the UCLA Dissertation Year Fellowship and the Welton Graduate Prize is gratefully acknowledge. I thank the CCPR at UCLA for its hospitality during the completion of this paper.

References

D. Ackerberg, K. Hirano, and Q. Shahriar. Identification of time and risk preferences in buy price auctions. Working Paper, University of Arizona, 2011.

A. Aradillas-Lopéz, A. Gandhi, and D. Quint. Identification and inference in ascending auctions with correlated private values. *Econometrica*, 81(2):489–534, 2013.

S. Athey and P. A. Haile. Nonparametric approaches to auctions. In J. J. Heckman and E. E. Leamer, editors, *Handbook of Econometrics*, volume 6A, chapter 60, pages 3847–3965. Elsevier, 2007.

A. R. Barron and C. Sheu. Approximation of density functions by sequences of exponential families. *Annals of Statistics*, 19(3):1347–1369, 1991.

H. J. Bierens and H. Song. Semi-nonparametric estimation of independently and identically repeated first-price auctions via an integrated simulated moments method. *Journal of Econometrics*, 168:108–119, 2012.

S. Campo, E. Guerre, I. Perrigne, and Q. Vuong. Semiparametric estimation of first-price auctions with risk-averse bidders. *Review of Economic Studies*, 78:112–147, 2011.

X. Chen. Large sample sieve estimation of semi-nonparametric models. In J. J. Heckman and E. E. Leamer, editors, *Handbook of Econometrics*, volume 6B, chapter 76, pages 5549–5632. Elsevier, 2007.

V. Chernozhukov, H. Hong, and E. Tammer. Estimation and confidence regions for parameter sets in econometric models. *Econometrica*, 75(5):1243–1284, 2007.

M. R. Delgado. Understanding overbidding: using the neural circuitry of reward to design economic auctions. *Science*, 321:1849–1852, 2008.

R. A. DeVore and G. G. Lorentz. *Constructive Approximation*. Springer-Verlag, 1993.

V. N. Gabushin. Inequalities for the norms of a function and its derivatives in metric L_p. *Matematicheskie Zametki*, 1(3):291–298, 1967.

E. Guerre, I. Perrigne, and Q. Vuong. Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574, 2000.

E. Guerre, I. Perrigne, and Q. Vuong. Nonparametric identification of risk aversion in first-price auctions under exclusion restrictions. *Econometrica*, 77(4):1193–1227, 2009.

D. J. Henderson, J. A. List, D. L. Millimet, C. F. Parmeter, and M. K. Price. Empirical implementation of nonparametric first-price auction models. *Journal of Econometrics*, 168:17–28, 2012.

K. Hendricks and H. J. Paarsch. A survey of recent empirical work concerning auctions. *Canadian Journal of Economics*, 28(2):403–426, 1995.

K. Hendricks and R. Porter. An empirical perspective on auctions. In M Armstrong and R. Porter, editors, *Handbook of Industrial Organization*, volume 3, chapter 32, pages 2073–2143. Elsevier, 2007.

B. R. Hickman and T. P. Hubbard. Replacing sample trimming with boundary correction in nonparametric estimation of first-price auctions. *Journal of Applied Econometrics*, forthcoming, 2014.

A. Hu, S. A. Matthews, and L. Zou. Risk aversion and optimal reserve prices in first- and second-price auctions. *Journal of Economic Theory*, 145:1188–1202, 2010.

D. Kim. Nonparametric estimation of utility function in first-price sealed-bid auctions. *Economics Letters*, 126:101–106, 2015.

A. P. Korostelev and A. B. Tsybakov. *Minimax Theory of Image Reconstruction*. Springer-Verlag, 1993.

J. Laffont. Game theory and empirical economics: the case of auction data. *European Economic Review*, 41(1):1–35, 1997.

Q. Li and J. S. Racine. *Nonparametric Econometrics*. Princeton University Press, 2007.

T. Li, I. Perrigne, and Q. Vuong. Semiparametric estimation of the optimal reserve price in first-price auctions. *Journal of Business and Economic Statistics*, 21(1): 53–64, 2003.

N. Liu and Y. Luo. A nonparametric test of exogenous participation in first-price auctions. Working Paper, University of Toronto, 2014.

J. Lu. Stochastic private values in auctions: identification and estimation. Working Paper, University of Southern California, 2004.

J. Lu and I. Perrigne. Estimating risk aversion from ascending and sealed-bid auctions: the case of timber. *Journal of Applied Econometrics*, 23:871–896, 2008.

V. Marmer and A. Shneyerov. Quantile-based nonparametric inference for first-price auctions. *Journal of Econometrics*, 167:345–357, 2012.

R. L. Matzkin. Restrictions of economic theory in nonparametric methods. In R. F. Engle and D.L. McFadden, editors, *Handbook of Econometrics*, volume 4, chapter 42, pages 2523–2558. Elsevier, 1994.

R. L. Matzkin. Nonparametric estimation of nonadditive random functions. *Econometrica*, 71(5):1339–1375, 2003.

H. J. Paarsch, H. Hong, and M. R. Haley. *An Introduction to the Structural Econometrics of Auction Data*. MIT Press, 2006.

I. Perrigne and Q. Vuong. Structural econometrics of first-price auctions: a survey of methods. *Canadian Journal of Agricultural Economics*, 47(3):203–223, 1999.

L. Stocco, E. Salcudean, and F. Sassani. Fast constrained global minimax optimization of robot parameters. *Robotica*, 16:595–605, 1998.

N. L. Stokey, R. E. Lucas, and E. C. Prescott. *Recursive Methods in Economic Dynamics*. Harvard University Press, 1989.

C. J. Stone. Optimal global rates of convergence for nonparametric regression. *Annals of Statistics*, 10(2):1040–1053, 1982.

C. J. Stone. Large-sample inference for log-spline models. *Annals of Statistics*, 18(2): 717–741, 1990.

C. J. Stone. Asymptotics for doubly flexible logspline response models. *Annals of Statistics*, 19(4):1832–1854, 1991.

X. Tang. Bounds on revenue distributions in counterfactual auctions with reserves prices. *RAND Journal of Economics*, 42(1):175–203, 2011.

X. Wu. Exponential series estimator of multivariate densities. *Journal of Econometrics*, 156:354–366, 2010.