Semi-Nonparametric Modeling and Estimation of First-Price Auction Models with Auction-Specific Heterogeneity*

Herman J. Bierens†
Pennsylvania State University
University Park, PA 16802, USA.

Hosin Song
Department of Economics
Ewha Womans University
Seoul, Korea.

January 19, 2018

Abstract

In this paper we extend and generalize the semi-nonparametric modeling and sieve estimation approach of Bierens and Song (2012) for independently and identically distributed first-price auctions to first-price auction models with observed auction-specific heterogeneity. The latter will be incorporated via a linear model in the auction-specific covariates for the log values with unknown error distribution. It will be shown that under some mild regularity conditions the parameters of

*We are grateful to Quang Vuong and Joris Pinkse for their constructive comments on an earlier version of this paper.

†Professor Emeritus of Economics, Pennsylvania State University, University Park, PA 16802. A previous version of this paper was presented at Syracuse University and the University of Southern California. Please address correspondence by email only to hbierens@psu.edu, as I no longer have an office at PennState.

the auction-specific covariates can be estimated consistently, unbiasedly, and asymptotic normally with standard convergence rate, without knowing the error distribution of the log-value model involved. Using these results, the conditional value distribution can now be estimated consistently via an integrated squared simulated Laplace moment sieve estimation approach, similar to Bierens and Song (2012).

Keywords: First price auctions, heterogeneity, log-linear values, Laplace transform, semiparametric estimation, semi-nonparametric modeling, simulated method of moments, sieve estimation, consistency, asymptotic normality.

JEL codes: C14, C21, C51, D44

1 Introduction

In many repeated first-price auctions the objects to be auctioned off are different across auctions. Consequently, the value distributions are then different across auctions. However, if we observe the auction-specific characteristics in the form of covariates, and the value distributions conditional on these covariates have the same functional form, the conditional bid distribution given the auction-specific covariates will be the same for all auctions. The question then arises how to incorporate the observable characteristics into the auction model. Laffont et al. (1995) incorporate covariates in the value distribution by specifying a linear regression model for the log of values with zero-mean normal errors. Donald and Paarsch (1996) parameterize the upper bound of the values as a function of covariates. Li (2005) specifies the value distribution as the exponential distribution with mean a linear function of covariates. Guerre et al. (2000) propose a two-stage nonparametric kernel density estimation approach, where in the first stage the bid distribution and density conditional on the covariates are estimated nonparametrically, which then is used in inverse form to generate values given the actual bids and the covariates. The generated values are then used to estimate the conditional value distribution nonparametrically.

In this paper, we propose an alternative semi-nonparametric (SNP hereafter) approach to estimate first-price auction models with observed auction-specific heterogeneity and private, symmetric and independent values conditional on a vector of auction specific covariates. This approach extends the

SNP integrated simulated moments estimation method of Bierens and Song (2012) to the heterogenous case with observable auction-specific covariates. We consider a first-price auction model where the log values takes the form of a linear regression-type model conditional on covariates, with unknown error distribution.

The setup of the paper is as follows. In section 2 we introduce the model and data-generating process. It will be assumed that the data consist of a random sample of first-price auctions, where in each auction the bids and the auction-specific covariates are observed. The model is a standard first-price auction model without binding reservation price, where the conditional value distribution is derived from a linear model in the covariates for the log values, with independent errors. The main reason for focusing on the nonbinding reservation price case is that in the case of a binding reservation price the value distribution for values below the reservation price may not identified, which is a problem for SNP modeling. Also, in section 2 we will show that under standard conditions the model is semi-nonparametrically identified. In section 3 we will set forth mild conditions such that the parameters of this log-linear value model can be estimated consistently, unbiasedly, and asymptotic normally, without knowing the exact form of the error distribution. In section 4 the conditional bid function will be modeled semi-nonparametrically. In section 5 we discuss the integrated simulated moments objective function to be used to estimate the value distribution semi-nonparametrically. The main difference with the setup in Bierens and Song (2012) is two-fold. First, we now take into account that different auctions have different numbers of bidders, assuming that the distribution of the latter has finite support. Secondly, the simulated moment condition involved is now based on empirical Laplace transforms rather than the empirical characteristic functions approach in Bierens and Song (2012). In section 6 we briefly discuss the infinite-dimensional parameter space involved, and in section 7 we show that the sieve estimator of the model is consistent, without requiring that the sieve estimator is confined to an infinite-dimensional compact set. The latter was assumed in Bierens and Song (2012), following the consistency proof in Bierens (2008).

2 Model and data-generating process

2.1 The model

Given a vector X of auction-specific characteristics, let $\Gamma_0(v|X)$ be the conditional distribution of the private value V that each potential bidder has for the object to be auctioned off in an auction characterized by a vector X of auction-specific covariates. As said before, we will only consider auctions without a reservation price. Therefore, the number of potential bidders, $I_X \geq 2$, is the same as the actual number of bidders in this auction, which is assumed to be ex-ante known to all potential bidders.

As is well-known, the equilibrium bid function of a first-price sealed bid auction without a reservation price, where values are independent and private, and bidders are symmetric and risk-neutral, takes the form

$$b_0\left(v|X, I_X\right) = v - \frac{1}{\Gamma_0(v|X)^{I_X-1}} \int_0^v \Gamma_0(z|X)^{I_X-1} \mathrm{d}z \tag{1}$$

See for example Riley and Samuelson (1981) or Krishna (2002). We will refer to this model as the Heterogeneous First Price Auction (HFPA) model.

We will assume that all the bids in the auctions involved are observed by the econometrician. Then conditional on X and I_X the conditional value distribution $\Gamma_0(v|X)$ is nonparametrically identified from the conditional distribution of the bids. This follows from Theorem 4 of Guerre et al. (2000). However, since the model in this paper is semi-nonparametric (SNP) rather than nonparametric, the identification issue will be addressed in section 2.4 below, because it appears than an additional condition is required for SNP identification.

2.2 Data-generating process

As to the data, it will be assumed that

Assumption 1. *We observe L auctions, labeled $k = 1, 2, ..., L$, each characterized by a vector $X_k \in \mathbb{R}^p$ of auction-specific covariates and a number $I_k = I_{X_k} \geq 2$ of bidders. The random vectors $X_k^+ = (X_k', I_k)'$, $k = 1, 2, ..., L$, are independently and identically distributed as $X^+ = (X', I_X)'$. The bids $B_{i,k}$, $i = 1, 2, ..., I_k$, in auction k are observed and are distributed as $b_0(V_{i,k}|X_k^+)$,* *where for each k the values $V_{i,k}$, $i = 1, 2, ..., I_k$, are independent random drawings from the conditional value distribution $\Gamma_0(v|X_k)$. Moreover, the sellers in these auctions do not ex-ante reveal their reservation prices, if any.*

The independence condition exclude the case where the same bidder bids in different auctions in this sample. Therefore, one should interpret these auctions as geographically so far apart that their sets of bidders don't overlap. Also, Assumption 1 excludes auctions that are repeated over time in the same locations. In the latter case the data set gets a panel data structure, which is beyond the scope of the current paper.

2.3 Log-linear values

The problem with the nonparametric two-step approach of Guerre et al. (2000) is that it is difficult to display the nonparametric estimation results for the density $\gamma_0(v|x)$ of $\Gamma_0(v|x)$ if the dimension of x is larger than 1. Therefore, we propose to specify the log values as a linear model in the covariates. In particular, we will assume that

Assumption 2. *Each value $V_{i,k}$ of bidder i in auction k is generated by*

$$\ln(V_{i.k}) = \beta_0' X_k + \ln(W_{i,k}), \tag{2}$$

where within and across auctions the $W_{i,k}$'s are drawn independently from an absolutely continuous distribution $F_0(w)$ with continuous density $f_0(w)$ and support $(0, \infty)$. Moreover,

$$\int_0^\infty w f_0(w) \mathrm{d}w < \infty. \tag{3}$$

This assumption is akin to the setup of Laffont et al. (1995), except that these authors assume that the $\ln(W_{i,k})$'s are i.i.d. zero mean normal, so that in their case $f_0(w)$ is the log-normal density, with only parameter the variance of the normal distribution involved. The moment condition (3) guarantees that conditional on $X^+ = (X', I_X)'$ the corresponding actual bids are bounded, with upper bound depending on X^+. See equation (10) below. Of course, the lower bound of the bids is zero. Moreover, since at this stage no further normalization conditions are imposed on $F_0(w)$, any constant term in X can be absorbed by $F_0(w)$, so that $\beta_0'X$ should not contain a constant term.

Remark 1. As to condition (3), consider the case $\beta_0 = 0$, $I_X = I_0$, so that $F_0(v)$ is the actual value distribution function. Riley and Samuelson (1981) assume that F_0 has bounded support. This condition is also adopted by Donald and Paarsch (1996) and Guerre et al. (2000), among others. However, Riley and Samuelson (1981) do not use this condition explicitly, but only its implication that then the expected revenue of the seller,

$$\overline{R} = I_0. \int_0^\infty \left(vf_0(v) + F_0(v) - 1\right) F_0(v)^{I_0-1}\mathrm{d}v,$$

is finite. Cf. Proposition 1 in Riley and Samuelson (1981). Since $\overline{R} < I_0. \int_0^\infty vf_0(v)\mathrm{d}v$, condition (3) guarantees that $\overline{R} < \infty$. On the other hand, condition (3) is rather restrictive in that certain frequently used distributions in econometric modeling are excluded. For example, suppose that the $\ln(W_{i,k})$'s have a standard logistic distribution, so that $F_0(w) = (1+w)^{-1}w$ with density $f_0(w) = (1+w)^{-2}$ and inverse $F_0^{-1}(u) = u/(1-u)$. Then

$$\int_0^\infty wf_0(w)\mathrm{d}w = \int_0^1 F_0^{-1}(u)\mathrm{d}u = \infty.$$

Finally, as will become clear, the approach in this paper does not hinge on condition (3), but only on the condition that $E[(\ln(W_{i,k}))^2] < \infty$. Cf. Assumption 4 below.

The conditional distribution of V given X now takes the form

$$\Gamma_0(v|X) = F_0(v.\exp(-\beta_0'X)), \tag{4}$$

with support $(0,\infty)$. Then the bid function can be written as

$$\begin{aligned} b_0\left(v|X^+\right) &= v - \frac{\exp(\beta_0'X)}{F_0(v.\exp(-\beta_0'X))^{I_X-1}} \int_0^{v\exp(-\beta_0'X)} F_0(y)^{I_X-1}\mathrm{d}y \quad (5) \\ &= \frac{(I_X-1)\exp(\beta_0'X)}{F_0(v.\exp(-\beta_0'X))^{I_X-1}} \int_0^{v\exp(-\beta_0'X)} yF_0(y)^{I_X-2}f_0(y)\mathrm{d}y \\ &= \frac{(I_X-1)\exp(\beta_0'X)}{F_0(v.\exp(-\beta_0'X))^{I_X-1}} \int_0^{F_0(v\exp(-\beta_0'X))} z^{I_X-2}F_0^{-1}(z)\mathrm{d}z \end{aligned}$$

for $v > 0$, where the second equality follows from integration by parts and the third one by the change of variable $z = F_0(y)$.

Given X, the values V_X of the bidders in this auction are independent random drawing from the conditional distribution function $F_0(v \exp(-\beta_0' X))$. As is well-known, such a value can be generated by solving $F_0(V_X . \exp(-\beta_0' X)) = U$, hence

$$V_X = \exp(\beta_0' X) F_0^{-1}(U), \tag{6}$$

where U is a random drawing from the uniform distribution on $[0,1]$, hereafter indicated by $\mathcal{U}(0,1)$. Thus, conditional on X^+, the bids in this auction are independently distributed as $\overline{b}_0\left(U|X^+\right)$, where for $u \in [0,1]$,

$$\overline{b}_0\left(u|X^+\right) = \exp(\beta_0' X)\left(F_0^{-1}(u) - \frac{\int_0^{F_0^{-1}(u)} F_0(y)^{I_X - 1}\mathrm{d}y}{u^{I_X - 1}}\right) \tag{7}$$

$$= \frac{\exp(\beta_0' X)}{u^{I_X - 1}} \int_0^u (I_X - 1) z^{I_X - 2} F_0^{-1}(z)\mathrm{d}z \tag{8}$$

$$= \exp(\beta_0' X) \int_0^1 (I_X - 1) z^{I_X - 2} F_0^{-1}(u.z)\mathrm{d}z. \tag{9}$$

Clearly, $\overline{b}_0\left(u|X^+\right)$ is strictly monotonic increasing in u, with upper bound

$$\overline{b}_0\left(1|X^+\right) = (I_X - 1)\exp(\beta_0' X) \int_0^1 z^{I_X - 2} F_0^{-1}(z)\mathrm{d}z < \infty \text{ a.s.} \tag{10}$$

where the inequality follows from (3), i.e.,

$$\infty > \int_0^\infty w f_0(w)\mathrm{d}w = \int_0^\infty F_0^{-1}(F_0(w))\mathrm{d}F_0(w) = \int_0^1 F_0^{-1}(z)\mathrm{d}z. \tag{11}$$

2.4 Semi-nonparametric identification

We will now show that F_0 and β_0 are identified, as follows. Suppose that there exists an alternative absolutely continuous distribution function F with support $(0, \infty)$, density f and inverse F^{-1} and an alternative parameter vector β, with corresponding bid function

$$\overline{b}\left(u|X^+\right) = \frac{(I_X - 1)\exp(\beta' X)}{u^{I_X - 1}} \int_0^u z^{I_X - 2} F^{-1}(z)\mathrm{d}z \tag{12}$$

such that for $U \sim \mathcal{U}(0,1)$, $\overline{b}_0\left(U|X^+\right)$ and $\overline{b}\left(U|X^+\right)$ have the same conditional distribution. Thus,

$$\Pr[\overline{b}_0\left(U|X^+\right) \leq b|X^+] = \Pr[\overline{b}\left(U|X^+\right) \leq b|X^+] \text{ a.s.}$$

for all $b \in (0, \overline{b}_0(1|X^+))$. Since $\overline{b}_0\left(u|X^+\right)$ is continuous and strictly monotonic increasing in u, it is invertible, with inverse $\overline{b}_0^{-1}\left(b|X^+\right)$ for $0 \leq b \leq \overline{b}_0\left(1|X^+\right)$. Then for all $u \in (0,1)$, and $b = \overline{b}_0\left(u|X^+\right)$,

$$u = \overline{b}_0^{-1}\left(b|X^+\right) = \Pr[U \leq u|X^+] = \Pr[\overline{b}_0^{-1}(\overline{b}\left(U|X^+\right)|X^+) \leq u|X^+] \text{ a.s.}$$

The latter implies that $\overline{b}_0^{-1}(\overline{b}\left(u|X^+\right)|X^+) = u$ for all $u \in (0,1)$, hence

$$\overline{b}_0\left(u|X^+\right) = \overline{b}\left(u|X^+\right) \text{ a.s. for all } u \in (0,1),$$

which implies that for all $u \in (0,1)$,

$$\exp(\beta_0'X)\int_0^u z^{I_X-2}F_0^{-1}(z)\mathrm{d}z = \exp(\beta'X)\int_0^u z^{I_X-2}F^{-1}(z)\mathrm{d}z \text{ a.s.}$$

Taking the derivative to u yields the equality

$$\exp(\beta_0'X)F_0^{-1}(u) = \exp(\beta'X)F^{-1}(u) \tag{13}$$

so that $(\beta-\beta_0)'X = \ln(F_0^{-1}(u)/F^{-1}(u))$, hence $(\beta-\beta_0)'(X-E[X]) = 0$ a.s. and thus

$$(\beta-\beta_0)'(X-E[X])(X-E[X])'(\beta-\beta_0) = 0 \text{ a.s.}$$

Taking the expectation the latter implies $(\beta-\beta_0)'\mathrm{Var}(X)(\beta-\beta_0) = 0$, so that if

Assumption 3. $\mathrm{Var}(X)$ *is finite and nonsingular,*

then $\beta = \beta_0$, and thus by (13), $F \equiv F_0$.

Note that this condition implies that X does not contain a constant 1. Moreover, the finiteness of $\mathrm{Var}(X)$ requires that $E[X'X] < \infty$, which is therefore implied by Assumption 3.

Summarizing, the following result holds.

Theorem 1. *Under Assumptions 1-3 the HFPA model is semi-nonparametrically identified.*

Remark 2. This identification result holds more generally, as follows. Specify the log-value model (2) nonparametrically as

$$\ln(V_{i.k}) = \ln(g_0(X_k)) + \ln(W_{i,k}), \tag{14}$$

where $g_0(x)$ is an unknown positive valued function. Then the identification result (13) becomes $g_0(X)F_0^{-1}(u) = g(X)F^{-1}(u)$ for all $u \in [0,1]$, where $g(X)$ and F are alternative versions of $g_0(X)$ and F_0, respectively. If we interpret (14) as a nonparametric median regression model, for example, then $F_0(1) = 0.5$, and imposing this restriction on F as well yields

$$g(X) = g(X)F^{-1}(0.5) = g_0(X)F_0^{-1}(0.5) = g_0(X) \text{ a.s.}$$

Thus, in principle we could model $g_0(X)$ (semi-)nonparametrically under the condition that $g_0(X)$ is the conditional median of V. However, this will make the estimation of the bid function unduly complicated.

3 Semi-parametric estimation of the log-linear value model

In this section we will set forth additional conditions for the construction of a consistent estimator $\widetilde{\beta}_L$ of the parameter vector β_0 in the log-linear value model (2) such that for $L \to \infty$,

$$\sqrt{L}\left(\widetilde{\beta}_L - \beta_0\right) \xrightarrow{\mathrm{d}} N_p(0, \Sigma), \tag{15}$$

without the need to know the distribution function F_0.

3.1 Consistency and asymptotic normality

Recall that the bids $B_{i,k}$, $i = 1, 2, ..., I_k$, in auction k are distributed as

$$\overline{b}_0\left(U_{i,k}|X_k^+\right) = \exp(\beta_0' X_k) \int_0^1 (I_k - 1) z^{I_k - 2} F_0^{-1}(U_{i,k}.z) \mathrm{d}z$$

where the $U_{i,k}$'s are random drawings from the $\mathcal{U}(0,1)$ distribution. Hence

$$b_{i,k} \stackrel{\text{def.}}{=} \ln(B_{i,k}) \sim \beta_0' X_k + Z_{i,k} \tag{16}$$

where

$$Z_{i,k} = \ln\left(\int_0^1 (I_k - 1) z^{I_k - 2} F_0^{-1}(U_{i,k}.z)\mathrm{d}z\right). \tag{17}$$

Now suppose that in addition to (or instead of) condition (3),

Assumption 4. $\int_0^\infty (\ln(w))^2 f_0(w)\mathrm{d}w < \infty$,

so that the errors $\ln(W_{i,k})$ in (2) have finite second moments. This is a standard condition for least squares estimation of log-linear regression models.

Remark 3. Note that condition (3) implies $\int_1^\infty (\ln(w))^2 f_0(w)\mathrm{d}w < \infty$, because $\sup_{w \geq 1} w^{-1}(\ln(w))^2 = 4e^{-2}$, hence

$$\int_1^\infty (\ln(w))^2 f_0(w)\mathrm{d}w \leq 4e^{-2} \int_1^\infty w f_0(w)\mathrm{d}w.$$

Moreover, a sufficient condition for $\int_0^1 (\ln(w))^2 f_0(w)\mathrm{d}w < \infty$ is that

$$\int_0^1 w^{-c} f_0(w)\mathrm{d}w < \infty \text{ for some } c \in (0,1), \tag{18}$$

because $\lim_{w \downarrow 0} w^c(\ln(w))^2 = 0$, hence $\sup_{0 \leq w < 1} w^c(\ln(w))^2 < \infty$, so that

$$\int_0^1 (\ln(w))^2 f_0(w)\mathrm{d}w \leq \sup_{0 \leq x < 1} x^c(\ln(x))^2 . \int_0^1 w^{-c} f_0(w)\mathrm{d}w < \infty.$$

Furthermore, (18) holds if $\lim_{w \downarrow 0} f_0(w) < \infty$, as is the case for the exponential density, because then $\int_0^1 w^{-c} f_0(w)\mathrm{d}w \leq \sup_{0 \leq x \leq 1} f_0(x).(1-c)^{-1} < \infty$. Of course, the condition $\lim_{w \downarrow 0} f_0(w) < \infty$ is not a necessary condition for (18), because the latter also holds if for some $\delta \in (0,1)$, $f_0(w) \propto w^{-\delta}$ as $w \downarrow 0$, with $c \in (0, 1-\delta)$.

The reason for Assumption 4 is the following result.

Lemma 1. *Under Assumption 4, the errors* $Z_{i,k}$ *in* (16) *satisfy* $E[Z_{i,k}^2] < \infty$.

Proof. Appendix

Next, suppose that

Assumption 5. *The marginal distribution of the number of bidders, I_X, has finite support, i.e., or some fixed $M \in \mathbb{N}\backslash\{1\}$, $\Pr[I_X = K_m] = p_m > 0$, $\sum_{m=1}^{M} p_m = 1$, where $2 \leq K_1 < K_2 < ... < K_M$.*

The trick is split up the data in sub-samples according to the values of I_k, as follows. Denote for $K \in \mathbb{N}\backslash\{1\}$,

$$\Omega(u|K) = \ln\left((K-1)\int_0^1 z^{K-2}F_0^{-1}(u.z)\mathrm{d}z\right), \tag{19}$$

and note that

$$Z_{i,k} = \Omega(U_{i,k}|I_k).$$

Then it follows from (16) that, with $\mathbf{1}(.)$ the indicator function,

$$\begin{aligned}
&(b_{i,k} - \beta' X_k)\,\mathbf{1}\left(I_k = K_m\right) \\
&= (\beta_0 - \beta)' X_k \mathbf{1}\left(I_k = K_m\right) + \Omega(U_{i,k}|K_m)\mathbf{1}\left(I_k = K_m\right), \\
&i = 1, 2, ..., K_m.
\end{aligned} \tag{20}$$

Next, denote for $m = 1, 2, 3, ..., M$,

$$\begin{aligned}
\widehat{p}_{m,L} &= \frac{1}{L}\sum_{j=1}^{L}\mathbf{1}\left(I_j = K_m\right), \\
\overline{X}_{m,L} &= \frac{\frac{1}{L}\sum_{j=1}^{L} X_j.\mathbf{1}\left(I_j = K_m\right)}{\widehat{p}_{m,L}}, \\
\overline{b}_{i,m,L} &= \frac{\frac{1}{L}\sum_{j=1}^{L} b_{i,j}.\mathbf{1}\left(I_j = K_m\right)}{\widehat{p}_{m,L}}.
\end{aligned} \tag{21}$$

Note that by (20),

$$\begin{aligned}
\overline{b}_{i,m,L} &= \beta_0' \frac{\frac{1}{L}\sum_{j=1}^{L} X_j.\mathbf{1}\left(I_j = K_m\right)}{\widehat{p}_{m,L}} + \frac{\frac{1}{L}\sum_{j=1}^{L}\Omega(U_{i,j}|K_m).\mathbf{1}\left(I_j = K_m\right)}{\widehat{p}_{m,L}} \\
&= \beta_0'\overline{X}_{m,L} + \frac{\frac{1}{L}\sum_{j=1}^{L}\Omega(U_{i,j}|K_m).\mathbf{1}\left(I_j = K_m\right)}{\widehat{p}_{m,L}},
\end{aligned}$$

hence,

$$\left(\frac{1}{K_m}\sum_{i=1}^{K_m}(b_{i,k}-\overline{b}_{i,m,L})-(X_k-\overline{X}_{m,L})'\beta\right)\mathbf{1}\left(I_k=K_m\right)$$
$$=(X_k-\overline{X}_{m,L})'(\beta_0-\beta).\mathbf{1}\left(I_k=K_m\right)$$
$$+\frac{1}{K_m}\sum_{i=1}^{K_m}\Omega(U_{i,k}|K_m)\mathbf{1}\left(I_k=K_m\right)$$
$$-\frac{1}{L}\sum_{j=1}^{L}\left(\frac{1}{K_m}\sum_{i=1}^{K_m}\Omega(U_{i,j}|K_m).\right)\mathbf{1}\left(I_j=K_m\right)\times\frac{\mathbf{1}\left(I_k=K_m\right)}{\widehat{p}_{m,L}}$$

and thus

$$\sum_{m=1}^{M}\frac{1}{L}\sum_{k=1}^{L}(X_k-\overline{X}_{m,L})\frac{1}{K_m}\sum_{i=1}^{K_m}(b_{i,k}-\overline{b}_{i,m,L})\mathbf{1}\left(I_k=K_m\right)$$
$$-\sum_{m=1}^{M}\left(\frac{1}{L}\sum_{k=1}^{L}(X_k-\overline{X}_{m,L})(X_k-\overline{X}_{m,L})'\mathbf{1}\left(I_k=K_m\right)\right)\beta$$
$$=\sum_{m=1}^{M}\left(\frac{1}{L}\sum_{k=1}^{L}(X_k-\overline{X}_{m,L})(X_k-\overline{X}_{m,L})'\mathbf{1}\left(I_k=K_m\right)\right)(\beta_0-\beta)$$
$$+\sum_{m=1}^{M}\frac{1}{L}\sum_{k=1}^{L}\frac{1}{K_m}\sum_{i=1}^{K_m}\Omega(U_{i,k}|K_m)(X_k-\overline{X}_{m,L})\mathbf{1}\left(I_k=K_m\right)$$
$$-\sum_{m=1}^{M}\left(\frac{1}{L}\sum_{j=1}^{L}\left(\frac{1}{K_m}\sum_{i=1}^{K_m}\Omega(U_{i,j}|K_m)\right)\mathbf{1}\left(I_j=K_m\right)\right)$$
$$\times\frac{1}{\widehat{p}_{m,L}}\frac{1}{L}\sum_{k=1}^{L}(X_k-\overline{X}_{m,L})\mathbf{1}\left(I_k=K_m\right).$$

The last sum is zero, because it follows trivially from the definition of $\overline{X}_{m,L}$ that

$$\frac{1}{L}\sum_{k=1}^{L}(X_k-\overline{X}_{m,L})\mathbf{1}\left(I_k=K_m\right)\equiv 0. \tag{22}$$

Thus, denoting

$$\widetilde{\Sigma}_{1,L} \quad = \quad \sum_{m=1}^{M}\frac{1}{L}\sum_{k=1}^{L}(X_k-\overline{X}_{m,L})(X_k-\overline{X}_{m,L})'\mathbf{1}\left(I_k=K_m\right), \tag{23}$$

$$\widetilde{\eta}_L = \sum_{m=1}^{M} \frac{1}{L} \sum_{k=1}^{L} (X_k - \overline{X}_{m,L}) \frac{1}{K_m} \sum_{i=1}^{K_m} \left(b_{i,k} - \overline{b}_{i,m,L} \right) \mathbf{1}\left(I_k = K_m \right) \quad (24)$$

and

$$\begin{aligned} \widetilde{\rho}_L &= \sum_{m=1}^{M} \frac{1}{K_m} \sum_{i=1}^{K_m} \Omega(U_{i,k}|K_m) \frac{1}{L} \sum_{k=1}^{L} (X_k - \overline{X}_{m,L}) \mathbf{1}\left(I_k = K_m \right) \\ &= \frac{1}{L} \sum_{k=1}^{L} \sum_{m=2}^{M} \frac{1}{K_m} \sum_{i=1}^{K_m} \left(\Omega(U_{i,k}|K_m) - \int_0^1 \Omega(u|K_m) \mathrm{d}u \right) \\ &\quad \times (X_k - \overline{X}_{m,L}) \mathbf{1}\left(I_k = K_m \right) \end{aligned} \quad (25)$$

where the second equality follows from (22), we have

$$\widetilde{\eta}_L - \widetilde{\Sigma}_{1,L}\beta = \widetilde{\Sigma}_{1,L}(\beta_0 - \beta) + \widetilde{\rho}_L$$

If it can be shown that $\widetilde{\rho}_L = o_p(1)$ and $\widetilde{\Sigma}_{1,L} \xrightarrow{\mathrm{p}} \Sigma_1$, where $\det(\Sigma_1) > 0$, then the least squares estimator

$$\widetilde{\beta}_L = \widetilde{\Sigma}_{1,L}^{-1} \widetilde{\eta}_L = \beta_0 + \widetilde{\Sigma}_{1,L}^{-1} \widetilde{\rho}_L \quad (26)$$

is consistent, and if in addition $\sqrt{L}\widetilde{\rho}_L \xrightarrow{\mathrm{d}} N_p\left(0, \Sigma_2\right)$, where also Σ_2 is nonsingular, then

$$\sqrt{L}\left(\widetilde{\beta}_L - \beta_0 \right) \xrightarrow{\mathrm{d}} N_p\left(0, \Sigma_1^{-1} \Sigma_2 \Sigma_1^{-1} \right) \quad (27)$$

Moreover, since the $U_{i,k}$'s are independent of the X_k's and the I_k's, we have by the law of iterated expectations that $E\left[\widetilde{\Sigma}_{1,L}^{-1} \widetilde{\rho}_L \right] = 0$, hence then $\widetilde{\beta}_L$ is *unbiased*!

To prove that the asymptotic normality result (27) holds, we need the following auxiliary results. Recall that

$$\begin{aligned} \widetilde{\rho}_L &= \frac{1}{L} \sum_{k=1}^{L} \sum_{m=1}^{M} \frac{1}{K_m} \sum_{i=1}^{K_m} \left(\Omega(U_{i,k}|K_m) - \int_0^1 \Omega(u|K_m) \mathrm{d}u \right) \\ &\quad \times (X_k - \overline{X}_{m,L}) \mathbf{1}\left(I_k = K_m \right) \end{aligned} \quad (28)$$

and

$$\overline{X}_{m,L} = \frac{\frac{1}{L} \sum_{j=1}^{L} X_j . \mathbf{1}\left(I_j = K_m \right)}{\widehat{p}_{m,L}}$$

$$
\begin{aligned}
&= \frac{\frac{1}{L}\sum_{j=1}^{L}\left(X_j.\mathbf{1}\left(I_j = K_m\right) - E\left[X.\mathbf{1}(I_X = K_m)\right]\right)}{\widehat{p}_{m,L}} \\
&\quad + E\left[X|I_X = K_m\right].p_m/\widehat{p}_{m,L} \\
&= \frac{\frac{1}{L}\sum_{j=1}^{L}\left(X_j.\mathbf{1}\left(I_j = K_m\right) - E\left[X.\mathbf{1}(I_X = K_m)\right]\right)}{\widehat{p}_{m,L}} \\
&\quad + \frac{E\left[X|I_X = K_m\right]\left(p_m - \widehat{p}_{m,L}\right)}{\widehat{p}_{m,L}} + E\left[X|I_X = K_m\right]
\end{aligned}
$$

Since $p_m > 0$ for each $m \in \{1, 2, 3, ..., M\}$, and

$$
\begin{aligned}
\frac{1}{\sqrt{L}}\sum_{j=1}^{L}\left(X_j.\mathbf{1}\left(I_j = K_m\right) - E\left[X.\mathbf{1}(I_X = K_m)\right]\right) &= O_p(1). \\
\sqrt{L}\left(\widehat{p}_{m,L} - p_m\right) &= O_p(1),
\end{aligned}
$$

it follows easily that

$$
\begin{aligned}
\widehat{p}_{m,L} &= p_m + O_p(1/\sqrt{L}), && (29) \\
\overline{X}_{m,L} &= E\left[X|I_X = K_m\right] + O_p(1/\sqrt{L}) && (30)
\end{aligned}
$$

where the O_p terms are uniform in m. Consequently,

$$
\begin{aligned}
\sqrt{L}\widetilde{\rho}_L &= \frac{1}{\sqrt{L}}\sum_{k=1}^{L}\sum_{m=1}^{M}\frac{1}{K_m}\sum_{i=1}^{K_m}\left(\Omega(U_{i,k}|K_m) - \int_0^1 \Omega(u|K_m)\mathrm{d}u\right) \\
&\quad \times(X_k - \overline{X}_{m,L})\mathbf{1}\left(I_k = K_m\right) \\
&= \frac{1}{\sqrt{L}}\sum_{k=1}^{L}\sum_{m=1}^{M}\frac{1}{K_m}\sum_{i=1}^{K_m}\left(\Omega(U_{i,k}|K_m) - \int_0^1 \Omega(u|K_m)\mathrm{d}u\right) \\
&\quad \times\left(X_k - E\left[X|I_X = K_m\right]\right)\mathbf{1}\left(I_k = K_m\right) \\
&\quad - \frac{1}{\sqrt{L}}\sum_{k=1}^{L}\sum_{m=1}^{M}\frac{1}{K_m}\sum_{i=1}^{K_m}\left(\Omega(U_{i,k}|K_m) - \int_0^1 \Omega(u|K_m)\mathrm{d}u\right) \\
&\quad \times\left(\overline{X}_{m,L} - E\left[X|I_X = K_m\right]\right)\mathbf{1}\left(I_k = K_m\right) \\
&= \frac{1}{\sqrt{L}}\sum_{k=1}^{L}\sum_{m=1}^{M}\frac{1}{K_m}\sum_{i=1}^{K_m}\left(\Omega(U_{i,k}|K_m) - \int_0^1 \Omega(u|K_m)\mathrm{d}u\right) \\
&\quad \times\left(X_k - E\left[X|I_X = K_m\right]\right)\mathbf{1}\left(I_k = K_m\right)
\end{aligned}
$$

$$
\begin{aligned}
& -\frac{1}{\sqrt{L}} \sum_{k=1}^{L} \sum_{m=1}^{M} \frac{1}{K_m} \sum_{i=1}^{K_m} \left(\Omega(U_{i,k}|K_m) - \int_0^1 \Omega(u|K_m) \mathrm{d}u \right) \\
& \times \mathbf{1}\left(I_k = K_m\right) \times O_p(1/\sqrt{L}) \\
= \; & \frac{1}{\sqrt{L}} \sum_{k=1}^{L} \sum_{m=1}^{M} \frac{1}{K_m} \sum_{i=1}^{K_m} \left(\Omega(U_{i,k}|K_m) - \int_0^1 \Omega(u|K_m) \mathrm{d}u \right) \\
& \times \left(X_k - E\left[X|I_X = K_m\right]\right) \mathbf{1}\left(I_k = K_m\right) + O_p(1/\sqrt{L}),
\end{aligned}
$$

where the third equality follows from (30) and the last equality follows from the easy fact that under Assumptions 1 and 4,

$$
\begin{aligned}
\frac{1}{\sqrt{L}} \sum_{k=1}^{L} \sum_{m=1}^{M} \frac{1}{K_m} \sum_{i=1}^{K_m} \left(\Omega(U_{i,k}|K_m) - \int_0^1 \Omega(u|K_m) \mathrm{d}u \right) \mathbf{1}\left(I_k = K_m\right) \\
= O_p(1).
\end{aligned}
$$

Finally, denote

$$
\begin{aligned}
Z_k \;=\; & \sum_{m=1}^{M} \frac{1}{K_m} \sum_{i=1}^{K_m} \left(\Omega(U_{i,k}|K_m) - \int_0^1 \Omega(u|K_m) \mathrm{d}u \right) \\
& \times \left(X_k - E\left[X|I_X = K_m\right]\right) \mathbf{1}\left(I_k = K_m\right),
\end{aligned}
$$

so that

$$
\sqrt{L}\widetilde{\rho}_L = \frac{1}{\sqrt{L}} \sum_{k=1}^{L} Z_k + O_p(1/\sqrt{L})
$$

Note that under Assumption 1 the Z_k's are i.i.d., with $E[Z_k] = 0$ and variance matrix

$$
\begin{aligned}
\Sigma_2 \;=\; & E[Z_k Z_k'] \\
=\; & \sum_{m=1}^{M} E\left[\left(\frac{1}{K_m} \sum_{i=1}^{K_m} \left(\Omega(U_{i,k}|K_m) - \int_0^1 \Omega(u|K_m) \mathrm{d}u \right) \right)^2 \right] \\
& \times E\left[\left(X_k - E\left[X|I_X = K_m\right]\right) \left(X_k - E\left[X|I_X = K_m\right]\right)' \mathbf{1}\left(I_k = K_m\right) \right] \\
=\; & \sum_{m=1}^{M} \frac{1}{K_m} \left(\int_0^1 \left(\Omega(u|K_m)\right)^2 \mathrm{d}u - \left(\int_0^1 \Omega(u|K_m) \mathrm{d}u \right)^2 \right) \\
& \times E\left[\left(X - E\left[X|I_X = K_m\right]\right) \left(X - E\left[X|I_X = K_m\right]\right)' |I_X = K_m \right]
\end{aligned}
$$

$$
\begin{aligned}
&\times p_m \\
=\ & \sum_{m=1}^{M} \frac{p_m}{K_m} \left(\int_0^1 (\Omega(u|K_m))^2 \, \mathrm{d}u - \left(\int_0^1 \Omega(u|K_m) \mathrm{d}u \right)^2 \right) \\
& \times \mathrm{Var}\,(X|I_X = K_m)\,. \qquad (31)
\end{aligned}
$$

Moreover, note that by Assumption 4, $\int_0^1 (\Omega(u|K_m))^2 \mathrm{d}u < \infty$, and since $\Omega(u|K_m)$ cannot be constant we have

$$
\int_0^1 (\Omega(u|K_m))^2 \, \mathrm{d}u > \left(\int_0^1 \Omega(u|K_m) \mathrm{d}u \right)^2 .
$$

Therefore, Σ_2 is finite and nonsingular if in addition,

Assumption 6. *For at least one* m, $\mathrm{Var}(X|I_X = K_m)$ *is nonsingular.*

It follows now from the standard central limit theorem that under Assumptions 1-6,

$$
\sqrt{L}\widetilde{\rho}_L \xrightarrow{\mathrm{d}} N_p[0, \Sigma_2], \qquad (32)
$$

where

$$
\begin{aligned}
\Sigma_2 \ =\ & \sum_{m=1}^{M} \frac{1}{K_m} \left(\int_0^1 (\Omega(u|K_m))^2 \, \mathrm{d}u - \left(\int_0^1 \Omega(u|K_m) \mathrm{d}u \right)^2 \right) \\
& \times \mathrm{Var}\,(X|I_X = K_m) \times p_m
\end{aligned}
$$

To derive a consistent estimator of Σ_2, note first that by (29) and (30),

$$
\begin{aligned}
& \widetilde{\Gamma}_{L,m} \overset{\text{def.}}{=} \frac{1}{L} \sum_{k=1}^{L} (X_k - \overline{X}_{m,L})(X_k - \overline{X}_{m,L})' \mathbf{1}\,(I_k = K_m) \qquad (33) \\
& = \frac{1}{L} \sum_{k=1}^{L} X_k X_k' . \mathbf{1}\,(I_k = K_m) - \left(\frac{1}{L} \sum_{k=1}^{L} X_k \mathbf{1}\,(I_k = K_m) \right) \overline{X}'_{m,L} \\
& - \overline{X}_{m,L} \frac{1}{L} \sum_{k=1}^{L} X_k' \mathbf{1}\,(I_k = K_m) + \overline{X}_{m,L} \overline{X}'_{m,L} \left(\frac{1}{L} \sum_{k=1}^{L} \mathbf{1}\,(I_k = K_m) \right) \\
& = \frac{1}{L} \sum_{k=1}^{L} X_k X_k' . \mathbf{1}\,(I_k = K_m) - \widehat{p}_{m,L} \overline{X}_{m,L} \overline{X}'_{m,L}
\end{aligned}
$$

$$= \frac{1}{L}\sum_{k=1}^{L} X_k X_k'.\mathbf{1}\left(I_k = K_m\right) - p_m.\left(E\left[X|I_X = K_m\right]\right)\left(E\left[X|I_X = K_m\right]\right)' + o_p(1)$$

$$= \left(E\left[XX'|I_X = K_m\right] - \left(E\left[X|I_X = K_m\right]\right)\left(E\left[X|I_X = K_m\right]\right)'\right) p_m + o_p(1)$$

$$= \mathrm{Var}\left(X|I_X = K_m\right) \times p_m + o_p(1)$$

So it remains to derive a consistent estimator $\widetilde{\sigma}^2_{L,m}$ of

$$\sigma_m^2 = \int_0^1 \left(\Omega(u|K_m)\right)^2 \mathrm{d}u - \left(\int_0^1 \Omega(u|K_m)\mathrm{d}u\right)^2$$

as then

$$\widetilde{\Sigma}_{2,L} = \sum_{m=1}^{M} \frac{1}{K_m}\widetilde{\sigma}^2_{L,m}\widetilde{\Gamma}_{L,m} \xrightarrow{\mathrm{p}} \Sigma_2. \tag{34}$$

Without loss of generality we may assume that in each auction the bids are sorted in random order. Then

$$\left(b_{1,k} - b_{2,k}\right)\mathbf{1}\left(I_k = K_m\right) = \left(\Omega(U_{1,k}|K_m) - \Omega(U_{2,k}|K_m)\right)\mathbf{1}\left(I_k = K_m\right)$$

where the $U_{1,k}$'s and $U_{2,k}$'s are i.i.d. $\mathcal{U}(0,1)$. Hence by the law of large numbers,

$$\frac{1}{L}\sum_{k=1}^{L}\left(b_{1,k} - b_{2,k}\right)^2\mathbf{1}\left(I_k = K_m\right)$$

$$= \frac{1}{L}\sum_{k=1}^{L}\left(\Omega(U_{1,k}|K_m) - \Omega(U_{2,k}|K_m)\right)^2\mathbf{1}\left(I_k = K_m\right)$$

$$= E\left[\left(\Omega(U_{1,k}|K_m) - \Omega(U_{2,k}|K_m)\right)^2\right] p_m + o_p(1)$$

$$= 2\left(\int_0^1 \left(\Omega(u|K_m)\right)^2 \mathrm{d}u - \left(\int_0^1 \Omega(u|K_m)\mathrm{d}u\right)^2\right) p_m + o_p(1)$$

$$= 2\sigma_m^2 p_m + o_p(1).$$

Therefore,

$$\widetilde{\sigma}^2_{L,m} = \frac{\frac{1}{L}\sum_{k=1}^{L}\left(b_{1,k} - b_{2,k}\right)^2\mathbf{1}\left(I_k = K_m\right)}{2.\widehat{p}_{m,L}} \xrightarrow{\mathrm{p}} \sigma_m^2.$$

Finally, observe from (23) and (33) that

$$\widetilde{\Sigma}_{1,L} = \sum_{m=1}^{M} \widetilde{\Gamma}_{L,m} \xrightarrow{\text{p}} \sum_{m=2}^{M} \text{Var}\left(X|I_X = K_m\right) \times p_m = \Sigma_1,$$

say, where by Assumption 6, Σ_1 is positive definite.

Summarizing, the following result holds.

Theorem 2. *Under Assumptions 1-6 the proposed estimator* $\widetilde{\beta}_L$ *in* (26) *is unbiased, i.e.,* $E[\widetilde{\beta}_L] = \beta_0$, *consistent, and asymptotically normal*:

$$\sqrt{L}\left(\widetilde{\beta}_L - \beta_0\right) \xrightarrow{\text{d}} N_p\left(0, \Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}\right),$$

where

$$\begin{aligned}
\Sigma_1 &= \sum_{m=1}^{M} \text{Var}\left(X|I_X = K_m\right) \times p_m, \ \det(\Sigma_1) > 0, \\
\Sigma_2 &= \sum_{m=1}^{M} \frac{1}{K_m}\left(\int_0^1 \left(\Omega(u|K_m)\right)^2 \text{d}u - \left(\int_0^1 \Omega(u|K_m)\text{d}u\right)^2\right) \\
&\quad \times \text{Var}\left(X|I_X = K_m\right) \times p_m, \ \det(\Sigma_2) > 0.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\widetilde{\Sigma}_{1,L} &= \sum_{m=1}^{M} \frac{1}{L}\sum_{k=1}^{L}(X_k - \overline{X}_{m,L})(X_k - \overline{X}_{m,L})'\mathbf{1}\left(I_k = K_m\right) \xrightarrow{\text{p}} \Sigma_1, \\
\widetilde{\Sigma}_{2,L} &= \sum_{m=1}^{M} \frac{1}{K_m} \frac{\frac{1}{L}\sum_{k=1}^{L}\left(b_{1,k} - b_{2,k}\right)^2 \mathbf{1}\left(I_k = K_m\right)}{2.\widehat{p}_{m,L}} \\
&\quad \times \frac{1}{L}\sum_{k=1}^{L}(X_k - \overline{X}_{m,L})(X_k - \overline{X}_{m,L})'\mathbf{1}\left(I_k = K_m\right) \xrightarrow{\text{p}} \Sigma_2,
\end{aligned}$$

hence $\widetilde{\Sigma}_{1,L}^{-1}\widetilde{\Sigma}_{2,L}\widetilde{\Sigma}_{1,L}^{-1} \xrightarrow{\text{p}} \Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}$ *as* $L \to \infty$.

3.2 A numerical example

In order to see how well the estimation procedure in Theorem 2 works, we have generated two sets of auctions, for $L = 100$ and $L = 500$, as follows.

- In each auction k the vector $X_k = (X_{1,k}, X_{2,k}, X_{3,k}, X_{4,k}, X_{5,k})'$ of auction-specific covariates have been generated by drawing each $X_{i,k}$ from the standard normal distribution, conditional on $|X_{i,k}| < 1$.

- The corresponding number of bidders I_k has been generated as $I_k = 2 + R(X_{1,k})$, where $R(X_{1,k})$ is a random drawing from the Binomial-Logit$(3, p(X_{1,k}))$ distribution, with $p(x) = 1/(1 + \exp(-x))$. Thus, M and K_m in Assumption 5 are

$$M = 4,\ K_m = 1 + m,\quad m = 1, 2, 3, 4.$$

- The parameter vector involved has been chosen as $\beta_0 = (1, 0, 0, 0, 0)'$.

- The distribution function F_0 in Assumption 2 has been chosen as $F_0(w) = 1 - \exp(-w^2)$, with density $f_0(w) = 2w\exp(-w^2)$ and inverse

$$F_0^{-1}(u) = \sqrt{\ln\left(1/(1-u)\right)}.$$

In this case $\int_0^\infty w f_0(w)\mathrm{d}w = \sqrt{\pi}/2$,[1] and it is not hard to verify that $\int_0^\infty (\ln(w))^2 f_0(w)\mathrm{d}w < \infty$.

- The bids $B_{i,k}$, $i = 1, 2, ..., I_k$, in auction k are generated according to bid function (9) as

$$B_{i,k} = \exp(X_{1,k}) \int_0^1 (I_k - 1) z^{I_k - 2} F_0^{-1}(U_{i,k}.z)\mathrm{d}z,$$

where the $U_{i,k}$'s are independent random drawings from the $\mathcal{U}(0,1)$ distribution.

[1] Because

$$\begin{aligned}\int_0^\infty w f_0(w)\mathrm{d}w &= 2\int_0^\infty w^2\exp(-w^2)\mathrm{d}w \\ &= \frac{1}{\sqrt{2}}\int_0^\infty x^2\exp(-x^2/2)\mathrm{d}x \\ &= \frac{\sqrt{2\pi}}{2\sqrt{2}}\int_{-\infty}^\infty x^2\frac{\exp(-x^2/2)}{\sqrt{2\pi}}\mathrm{d}x = \sqrt{\pi}/2.\end{aligned}$$

The estimation results are presented in Table 1.

Table 1. Estimation results for β_0

	$L = 100$		$L = 500$	
β_0	$\widetilde{\beta}_L$	*t-value*	$\widetilde{\beta}_L$	*t-value*
1	1.054020941	15.925	1.004394570	33.158
0	0.006968898	0.103	−0.000557115	−0.020
0	0.113922071	1.831	0.012743293	0.465
0	−0.004020778	−0.058	0.020242694	0.746
0	0.039059912	0.602	−0.027952404	−1.039

Clearly, the estimator $\widetilde{\beta}_L$ does a good job in this particular case.

4 SNP modeling of the bid function

The question we will address now is: How can we model the distribution function F_0 and its inverse F_0^{-1} in the bid function (7), i.e.,

$$\overline{b}_0\left(u|X^+\right) = \exp(\beta_0' X)\left(F_0^{-1}(u) - \frac{\int_0^{F_0^{-1}(u)} F_0(y)^{I_X - 1}\mathrm{d}y}{u^{I_X - 1}}\right),$$

semi-nonparametrically?

However, recall from (8) and (9) that $\overline{b}_0\left(u|X^+\right)$ can also be written as

$$\begin{aligned}\overline{b}_0\left(u|X^+\right) &= \exp(\beta_0' X)\frac{\int_0^u (I_X - 1)z^{I_X - 2}F_0^{-1}(z)\mathrm{d}z}{u^{I_X - 1}} \\ &= \exp(\beta_0' X)\int_0^1 (I_X - 1)z^{I_X - 2}F_0^{-1}(u.z)\mathrm{d}z,\end{aligned}$$

so that in modeling $\overline{b}_0\left(u|X^+\right)$ semi-nonparametrically it suffices to only model $F_0^{-1}(z)$ semi-nonparametrically.

4.1 Transforming $F_0(w)$ and/or $F_0^{-1}(u)$ to a distribution function on the unit interval

In Bierens and Song (2012) we proposed to write

$$F_0(w) = H_0(G(w)), \tag{35}$$

where $G(w)$ is an a priori chosen absolutely continuous distribution function on $[0, \infty)$ with density $g(w)$ satisfying $g(w) > 0$ for $w > 0$, so that $G(w)$ is invertible, with inverse $G^{-1}(u)$. Then

$$H_0(u) = H_0(G(G^{-1}(u))) = F_0(G^{-1}(u))$$

is an absolutely continuous distribution function on $[0, 1]$ with density

$$h_0(u) = \frac{f_0(G^{-1}(u))}{g(G^{-1}(u))}, \tag{36}$$

hence,

$$f_0(w) = h_0(G(w))g(w).$$

Since $h_0(u) > 0$ on $(0, 1)$ because $f_0(w) > 0$ and $g(w) > 0$ on $(0, \infty)$, H_0 is invertible with inverse H_0^{-1}, which itself is an absolutely continuous distribution function on $[0, 1]$, with density

$$\mathrm{d}H_0^{-1}(u)/\mathrm{d}u = \frac{1}{h_0(H_0^{-1}(u))}.$$

Moreover, the inverse $F_0^{-1}(u)$ can be computed by solving $u = H_0(G(w))$, hence

$$F_0^{-1}(u) = G^{-1}(H_0^{-1}(u)). \tag{37}$$

Clearly, it is advisable to choose G such that its inverse G^{-1} has a closed form expression. For example, let $G(w)$ be the standard exponential distribution function, i.e., $G(w) = 1 - \exp(-w)$, with density $g(w) = \exp(-w)$ and inverse $G^{-1}(u) = \ln(1/(1-u))$. In this case $g(G^{-1}(u)) = 1 - u$, so that

$$h_0(u) = \frac{f_0(\ln(1/(1-u)))}{1-u}, \quad H_0(u) = F_0(\ln(1/(1-u))), \tag{38}$$

and by (37),

$$F_0^{-1}(u) = \ln\left(\frac{1}{1 - H_0^{-1}(u)}\right).$$

Consequently, given the choice of G, the actual bid function $\overline{b}_0\left(u|X^+\right)$ can now be written as

$$\overline{b}_0\left(u|X^+\right) = \exp(\beta_0' X) \int_0^1 (I_X - 1) z^{I_X - 2} G^{-1}(H_0^{-1}(u.z))\mathrm{d}z. \tag{39}$$

which in the case of the standard exponential distribution function G takes the form

$$\overline{b}_0\left(u|X^+\right) = -\exp(\beta_0'X)\int_0^1 (I_X - 1)z^{I_X-2}\ln(1 - H_0^{-1}(u.z))\mathrm{d}z.$$

4.2 SNP modeling of density and distribution functions on the unit interval

The reason for the transformation (35) is that there are various ways to model $H_0(u)$ and $h_0(u)$ semi-nonparametrically. In particular, $\sqrt{h_0(u)} \in L^2(0,1)$, where the latter is the Hilbert space of square-integrable real functions on $(0,1)$ endowed with innerproduct $\langle f,g\rangle = \int_0^1 f(u)g(u)\mathrm{d}u$, norm $||f|| = \sqrt{\langle f,f\rangle}$ and associated metric $||f-g||$. As is well-known, there exist many complete orthonormal sequences $\{\rho_m(u)\}_{m=0}^\infty$ in $L^2(0,1)$ with $\rho_0(u) \equiv 1$. For such a sequence, denote

$$\delta_{0,m} = \frac{\int_0^1 \rho_m(u)\sqrt{h_0(u)}\mathrm{d}u}{\int_0^1 \sqrt{h_0(z)}\mathrm{d}z},\ m \in \mathbb{N}, \tag{40}$$

which satisfies

$$\sum_{m=1}^\infty \delta_{0,m}^2 = \left(\int_0^1 \sqrt{h_0(z)}\mathrm{d}z\right)^{-1} - 1 < \infty, \tag{41}$$

and let for $n \in \mathbb{N}$,

$$h(u|\pi_n\boldsymbol{\delta}^0) = \frac{\left(1 + \sum_{m=1}^n \delta_{0,m}\rho_m(u)\right)^2}{1 + \sum_{m=1}^n \delta_{0,m}^2},$$

where $\boldsymbol{\delta}^0 = \{\delta_{0,m}\}_{m=1}^\infty$ is the infinite-dimensional true parameter and π_n is the truncation operator, i.e., π_n applied to $\boldsymbol{\delta} = \{\delta_m\}_{m=1}^\infty$ as $\pi_n\boldsymbol{\delta}$ sets all the δ_m's for $m > n$ to zeros. Then it can be shown[2] that

$$\lim_{n\to\infty}\int_0^1 \left|h(u|\pi_n\boldsymbol{\delta}^0) - h_0(u)\right| \mathrm{d}u = 0, \tag{42}$$

[2] See, for example, Bierens (2014a).

hence, denoting $H(u|\pi_n\boldsymbol{\delta}^0) = \int_0^u h(z|\pi_n\boldsymbol{\delta}^0)\mathrm{d}z$, we have

$$\lim_{n\to\infty} \sup_{0\le u\le 1} \left|H(u|\pi_n\boldsymbol{\delta}^0) - H_0(u)\right| = 0$$

and thus

$$\lim_{n\to\infty} \sup_{w>0} \left|F(w|\pi_n\boldsymbol{\delta}^0) - F_0(w)\right| = 0, \text{ where } F(w|\pi_n\boldsymbol{\delta}^0) = H(G(w)|\pi_n\boldsymbol{\delta}^0).$$

Consequently,

$$H_0(u) \equiv H(u|\boldsymbol{\delta}^0) \text{ and } F_0(w) \equiv F(w|\boldsymbol{\delta}^0).$$

Bierens (2008), Bierens and Carvalho (2007) and Bierens and Song (2012) employ the Legendre polynomials for $\rho_m(u)$, with $\rho_0(u) \equiv 1$, and Bierens (2014a,b) advocate the use of the cosine series $\rho_m(u) = \sqrt{2}\cos(m\pi u)$, $m \in \mathbb{N}$, with $\rho_0(u) \equiv 1$. In the case of the Legendre polynomials, $H(u|\pi_n\boldsymbol{\delta}^0)$ has to be computed by numerical integration, whereas in the case of the cosine sequence, $H(u|\pi_n\boldsymbol{\delta}^0)$ has the following closed form:

$$\begin{aligned} H(u|\pi_n\boldsymbol{\delta}^0) &= u \\ &+ \frac{1}{1+\sum_{m=1}^n \delta_{0,m}^2}\left[2\sqrt{2}\sum_{k=1}^n \delta_{0,k}\frac{\sin(k\pi u)}{k\pi} + \sum_{m=1}^n \delta_{0,m}^2\frac{\sin(2m\pi u)}{2m\pi}\right. \\ &\qquad + 2\sum_{k=2}^n\sum_{m=1}^{k-1}\delta_{0,k}\delta_{0,m}\frac{\sin((k+m)\pi u)}{(k+m)\pi} \\ &\qquad \left. + 2\sum_{k=2}^n\sum_{m=1}^{k-1}\delta_{0,k}\delta_{0,m}\frac{\sin((k-m)\pi u)}{(k-m)\pi}\right]. \end{aligned} \tag{43}$$

See Bierens (2014a,b). Therefore, we will base our present approach on the cosine series $\rho_m(u) = \sqrt{2}\cos(m\pi u)$, $m \in \mathbb{N}$, with $\delta_{0,m}$'s in (43) defined by (40).

4.3 SNP bid function

The true bid function (7) can now be written as

$$\overline{b}_0\left(u|X^+, \beta_0, \boldsymbol{\delta}^0\right) = \exp(\beta_0' X).\Lambda(u|I_X, \boldsymbol{\delta}^0), \tag{44}$$

with SNP version

$$\overline{b}_0\left(u|X^+,\beta,\pi_n\boldsymbol{\delta}\right)=\exp(\beta'X)\Lambda(u|I_X,\pi_n\boldsymbol{\delta}), \tag{45}$$

where for $\boldsymbol{\delta}=\{\delta_m\}_{m=1}^{\infty}$ satisfying $\sum_{m=1}^{\infty}\delta_m^2<\infty$ and $K\in\mathbb{N}\backslash\{1\}$,

$$\Lambda(u|K,\boldsymbol{\delta})=\int_0^1(K-1)z^{K-2}G^{-1}(H^{-1}(u.z|\boldsymbol{\delta}))\mathrm{d}z \tag{46}$$

with $H^{-1}(u|\boldsymbol{\delta})$ the inverse of $H(u|\boldsymbol{\delta})$.

In particular, if we choose for G the standard exponential distribution function then (46) reads

$$\Lambda(u|K,\boldsymbol{\delta})=-\int_0^1(K-1)z^{K-2}\ln(1-H^{-1}(u.z|\boldsymbol{\delta}))\mathrm{d}z. \tag{47}$$

Remark 4. Given $H(z|\boldsymbol{\delta})$, the inverse $H^{-1}(u|\boldsymbol{\delta})$ can easily be approximated via the following simple algorithm. Starting from $a_1=0$, $b_1=1$, $c_1=H(0.5|\boldsymbol{\delta})$, suppose that after k steps we have narrowed down $H^{-1}(u|\boldsymbol{\delta}))$ to

$$H^{-1}(u|\boldsymbol{\delta}))\in(a_k,b_k)\subset(0,1).$$

Let $c_k=H(a_k+(b_k-a_k)/2|\boldsymbol{\delta})$. If $c_k>u$ then $a_k+(b_k-a_k)/2>H^{-1}(u|\boldsymbol{\delta})>a_k$, hence set

$$a_{k+1}=a_k,\ b_{k+1}=a_k+(b_k-a_k)/2.$$

If $c_k<u$ then $a_k+(b_k-a_k)/2<H^{-1}(u|\boldsymbol{\delta})<b_k$, hence set

$$a_{k+1}=a_k+(b_k-a_k)/2,\ b_{k+1}=b_k.$$

If $c_k=u$ then we are done, of course, but otherwise repeat this iteration until b_k-a_k is smaller than a preset small threshold, and choose $a_k+(b_k-a_k)/2$ as the approximation of $H^{-1}(u|\boldsymbol{\delta})$.

5 Integrated simulated Laplace moments

For the case $\beta_0=\beta=0$ and I_X is a.s. constant, $\boldsymbol{\delta}^0$ can be consistently estimated by the integrated squared simulated moments sieve estimation approach in Bierens and Song (2012).

Under the conditions of Theorem 2, this approach can be adapted as follows. For each bid $B_{i,k}$, $i = 1, 2, ..., I_k$, denote

$$Y_{i,k} = \exp(-\widetilde{\beta}'_L X_k) B_{i,k} = \exp\left(-\left(\widetilde{\beta}_L - \beta_0\right)' X_k\right) \exp(-\beta'_0 X_k) B_{i,k} \quad (48)$$

and recall that

$$\exp(-\beta'_0 X_k) B_{i,k} \sim \Lambda(U^0_{i,k} | I_k, \boldsymbol{\delta}^0). \quad (49)$$

where the $U^0_{i,k}$'s are i.i.d. $\mathcal{U}(0,1)$ distributed.

Next, denote

$$\widetilde{Y}_{i,k}(\boldsymbol{\delta}) = \Lambda(\widetilde{U}_{i,k} | I_k, \boldsymbol{\delta}), \quad (50)$$

where the $\widetilde{U}_{i,k}$'s are drawn independently from the $\mathcal{U}(0,1)$ distribution. Note that the same $\widetilde{U}_{i,k}$'s should be used for different $\boldsymbol{\delta}$'s. Moreover, denote for $m = 1, 2, 3, ..., M$ and $t \in (0, \infty)$,

$$\varphi_{m,L}(t) = \widehat{p}^{-1}_{m,L} \frac{1}{L} \sum_{k=1}^{L} \frac{1}{K_m} \sum_{i=1}^{K_m} \exp\left(-t.Y_{i,k}\right) \mathbf{1}\left(I_k = K_m\right), \quad (51)$$

$$\widehat{\varphi}_{m,L}(t|\boldsymbol{\delta}) = \widehat{p}^{-1}_{m,L} \frac{1}{L} \sum_{k=1}^{L} \frac{1}{K_m} \sum_{i=1}^{K_m} \exp\left(-t.\widetilde{Y}_{i,k}(\boldsymbol{\delta})\right) \mathbf{1}\left(I_k = K_m\right), \quad (52)$$

where $\widehat{p}_{m,L}$ is defined by (21).

Assuming that

Assumption 7. *The auction-specific covariates X_k have bounded support,*

it is a standard exercise to shown that

Lemma 2. *Under Assumptions 1-7,*

$$p \lim_{L\to\infty} \sup_{0\le t\le\tau} \left|\varphi_{m,L}(t) - \varphi_m(t|\boldsymbol{\delta}^0)\right| = 0, \text{ where} \quad (53)$$

$$\varphi_m(t|\boldsymbol{\delta}^0) = \int_0^1 \exp\left(-t.\Lambda(u|K_m, \boldsymbol{\delta}^0)\right) \mathrm{d}u,$$

and pointwise in $\boldsymbol{\delta}$,

$$p \lim_{L\to\infty} \sup_{0\le t\le\tau} \left|\widehat{\varphi}_{m,L}(t|\boldsymbol{\delta}) - \varphi_m(t|\boldsymbol{\delta})\right| = 0, \text{ where} \quad (54)$$

$$\varphi_m(t|\boldsymbol{\delta}) = \int_0^1 \exp\left(-t.\Lambda(u|K_m, \boldsymbol{\delta})\right) \mathrm{d}u,$$

for any $t \in (0, \infty)$.

Note that both $\varphi_m(t|\boldsymbol{\delta}^0)$ and $\varphi_m(t|\boldsymbol{\delta})$ are Laplace transforms of nonnegative random variables. As is well-known, two Laplace transforms are equal on $(0, \infty)$ if and only if their distributions are equal. More generally, the following result holds.

Lemma 3. *Let* V_1 *and* V_2 *be nonnegative random variables with Laplace transforms* $\mathcal{L}_1(t) = E[\exp(-t.V_1)]$ *and* $\mathcal{L}_2(t) = E[\exp(-t.V_2)]$, *respectively, where* $t \geq 0$. *Suppose that on an open interval* $T \subset (0, \infty)$, $\mathcal{L}_1(t) = \mathcal{L}_2(t)$ *for all* $t \in T$. *Then* V_1 *and* V_2 *have the same distribution. Consequently,* $\mathcal{L}_1(t) = \mathcal{L}_2(t)$ *for all* t *in an open subset of* $(0, \infty)$ *implies that* $\mathcal{L}_1(t) = \mathcal{L}_2(t)$ *for all* $t \geq 0$.

Proof. The proof of this result can likely be found somewhere in the literature, so we do not claim originality. However, a Google Scholar search of the "uniqueness of the Laplace transform of nonnegative random variables" did not yield a relevant reference to this result. Therefore, the proof is given in the Appendix.

Choosing $T = (0, 1)$, Lemma 3 implies that if $\varphi_m(t|\boldsymbol{\delta}^0) = \varphi_m(t|\boldsymbol{\delta})$ for all $t \in (0, 1)$, then for two mutually independent $\mathcal{U}(0, 1)$ distributed random variables U_1 and U_2, $\Lambda(U_1|K_m, \boldsymbol{\delta}^0) \sim \Lambda(U_2|K_m, \boldsymbol{\delta})$, which implies that

$$\Lambda(u|K_m, \boldsymbol{\delta}^0) \equiv \Lambda(u|K_m, \boldsymbol{\delta}) \text{ on } [0, 1],$$

hence $H(u|\boldsymbol{\delta}) \equiv H(u|\boldsymbol{\delta}^0)$ on $[0, 1]$ and therefore $\boldsymbol{\delta} = \boldsymbol{\delta}^0$.

These results suggests to estimate $\boldsymbol{\delta}^0$ by minimizing the integrated squared simulated Laplace moments (ISSLM) objective function

$$\begin{aligned} \widehat{Q}_L(\boldsymbol{\delta}) &= \frac{1}{M}\sum_{m=1}^{M}\int_0^1 \left(\widehat{\varphi}_{m,L}(t|\boldsymbol{\delta}) - \varphi_{m,L}(t)\right)^2 \mathrm{d}t \qquad (55) \\ &= \frac{1}{M}\sum_{m=1}^{M}\widehat{p}_{m,L}^{-2}\frac{1}{L^2}\sum_{k_1=1}^{L}\sum_{k_2=1}^{L}\mathbf{1}\left(I_{k_1} = K_m\right)\mathbf{1}\left(I_{k_2} = K_m\right) \\ &\quad \times\frac{1}{K_m^2}\sum_{i_1=1}^{K_m}\sum_{i_2=1}^{K_m}\left\{\frac{1 - \exp\left(-2\left(Y_{i_1,k_1} + Y_{i_2,k_2}\right)\right)}{2\left(Y_{i_1,k_1} + Y_{i_2,k_2}\right)}\right. \end{aligned}$$

$$+\frac{1-\exp\left(-2\left(\widetilde{Y}_{i_1,k_1}(\boldsymbol{\delta})+\widetilde{Y}_{i_2,k_2}(\boldsymbol{\delta})\right)\right)}{\widetilde{Y}_{i_1,k_1}(\boldsymbol{\delta})+\widetilde{Y}_{i_2,k_2}(\boldsymbol{\delta})}$$
$$\left.-2\frac{1-\exp\left(-\left(Y_{i_1,k_1}+Y_{i_2,k_2}+\widetilde{Y}_{i_1,k_1}(\boldsymbol{\delta})+\widetilde{Y}_{i_2,k_2}(\boldsymbol{\delta})\right)\right)}{Y_{i_1,k_1}+Y_{i_2,k_2}+\widetilde{Y}_{i_1,k_1}(\boldsymbol{\delta})+\widetilde{Y}_{i_2,k_2}(\boldsymbol{\delta})}\right\} \quad (56)$$

via a sieve estimation approach, similar to Bierens and Song (2012).

Note that computing $\widehat{Q}_L(\boldsymbol{\delta})$ exactly by (56) is very computational intensive, in particular for large L. On the other hand, the integral in (55) can be computed quite accurately via a Gaussian quadrature approach.

Observe from (51) and (52) that $\widehat{\varphi}_{m,L}(t|\boldsymbol{\delta}) \in [0,1]$ and $\varphi_{m,L}(t) \in [0,1]$, hence $(\widehat{\varphi}_{m,L}(t|\boldsymbol{\delta}) - \varphi_{m,L}(t))^2 \leq 1$, and thus for all $\boldsymbol{\delta}$,

$$0 \leq \widehat{Q}_L(\boldsymbol{\delta}) \leq 1 \quad (57)$$

Moreover, note that by Lemma 2,

$$p\lim_{L\to\infty} \widehat{Q}_L(\boldsymbol{\delta}) = Q(\boldsymbol{\delta}), \quad (58)$$

pointwise in $\boldsymbol{\delta}$, where

$$Q(\boldsymbol{\delta}) = \frac{1}{M}\sum_{m=1}^{M}\int_0^1 \left(\varphi_m(t|\boldsymbol{\delta}) - \varphi_m(t|\boldsymbol{\delta}^0)\right)^2 \mathrm{d}t, \quad (59)$$

and that by (53) and (54),

$$0 \leq Q(\boldsymbol{\delta}) \leq 1 \quad (60)$$

for all $\boldsymbol{\delta}$.

Remark 5. The approach in Bierens and Song (2012) was based on matching the empirical characteristic functions of the actual bids and the simulated bids, but in hindsight we could have used empirical Laplace transforms as well. However, at that time we were unaware of the result in Lemma 3.

6 The infinite-dimensional parameter space

It has been shown in Bierens (2014a, Theorem 1.19) that there exist possibly uncountable many sequences $\boldsymbol{\delta} = \{\delta_m\}_{m=1}^{\infty}$ such that for a given density $h(u)$

on $[0,1]$,

$$h(u) = h(u|\boldsymbol{\delta}) = \frac{\left(1+\sum_{m=1}^{\infty}\delta_m\sqrt{2}\cos(m\pi u)\right)^2}{1+\sum_{m=1}^{\infty}\delta_m^2} \text{ a.e. on } [0,1]. \quad (61)$$

On the other hand, if $h(u)$ is continuous and positive on $(0,1)$ then $\boldsymbol{\delta} = \{\delta_m\}_{m=1}^{\infty}$ is unique, and defined by

$$\delta_m = \sqrt{2}\frac{\int_0^1 \cos(m\pi u)\sqrt{h(u)}\mathrm{d}u}{\int_0^1 \sqrt{h(u)}\mathrm{d}u}, \ m \in \mathbb{N}. \quad (62)$$

See Bierens (2014a, Theorem 1.20). As follows from the proof of the latter theorem, of all the possible sequences $\boldsymbol{\delta} = \{\delta_m\}_{m=1}^{\infty}$ for which (61) hold, the one generated by (62) has the smallest value of $\sum_{m=1}^{\infty}\delta_m^2$, namely

$$\sum_{m=1}^{\infty}\delta_m^2 = \left(\int_0^1 \sqrt{h(z)}\mathrm{d}z\right)^{-1} - 1 \quad (63)$$

Note that by Assumption 2 the density $h_0(u)$ in (36) is continuous and positive on $(0,1)$, so that the corresponding sequence $\boldsymbol{\delta}^0 = \{\delta_{0,m}\}_{m=1}^{\infty}$ is unique. Consequently, (**??**) implies that $\boldsymbol{\delta} = \boldsymbol{\delta}^0$.

As to the space of suitable $\boldsymbol{\delta}$'s, the minimum requirement is that $\boldsymbol{\delta}$ corresponds to a density $h(u)$ satisfying (61), and the minimum requirement for this is that $\boldsymbol{\delta} = \{\delta_m\}_{m=1}^{\infty}$ satisfies $\sum_{m=1}^{\infty}\delta_m^2 < \infty$, i.e., $\boldsymbol{\delta} \in \Delta$, where

$$\Delta = \left\{\boldsymbol{\delta} = \{\delta_m\}_{m=1}^{\infty} : \sum_{m=1}^{\infty}\delta_m^2 < \infty\right\}.$$

A natural norm on Δ is the pseudo-Euclidean norm

$$||\boldsymbol{\delta}|| = ||\{\delta_m\}_{m=1}^{\infty}|| = \sqrt{\sum_{m=1}^{\infty}\delta_m^2}$$

with associated metric $||\boldsymbol{\delta}_1 - \boldsymbol{\delta}_2||$, so that Δ becomes a metric space.

We can also define the associated innerproduct

$$\langle\boldsymbol{\delta}_1, \boldsymbol{\delta}_2\rangle = \sum_{m=1}^{\infty}\delta_{1,m}\delta_{2,m} \text{ for } \boldsymbol{\delta}_i = \{\delta_{i,m}\}_{m=1}^{\infty}, \ i = 1, 2,$$

on Δ, so that $||\boldsymbol{\delta}|| = \sqrt{\langle\boldsymbol{\delta}, \boldsymbol{\delta}\rangle}$, and then it can be shown that Δ becomes a Hilbert space.[3]

[3]By showing that every Cauchy sequence in Δ takes a limit in Δ,

7 Sieve estimation and consistency

For $n \in \mathbb{N}$, let

$$\Delta_n = \left\{ \boldsymbol{\delta} = \{\delta_m\}_{m=1}^{\infty} : \sum_{m=1}^{n} \delta_m^2 \leq B_n,\ \delta_m = 0 \text{ for } m > n \right\},$$

where B_n is an a priori chosen monotonic increasing sequence of positive constants converging (slowly) to ∞ for $n \to \infty$. Note that Δ_n is compact because it is isomorph to the hyperball $||x||^2 \leq B_n$ in $\mathbb{R}^n$ which is closed and bounded and therefore compact. Moreover, $\Delta_n \subset \Delta_{n+1}$, and $\Delta = \overline{\cup_{n=1}^{\infty} \Delta_n}$, where the bar denotes the closure. The sequence Δ_n is called the "sieve", and each Δ_n is called a sieve space.

Now the sieve estimator of $\boldsymbol{\delta}^0$ is

$$\widehat{\boldsymbol{\delta}}_{n_L} = \arg\min_{\boldsymbol{\delta} \in \Delta_{n_L}} \widehat{Q}_L(\boldsymbol{\delta}), \tag{64}$$

where n_L is a subsequence of the sample size L satisfying $\lim_{L\to\infty} n_L = \infty$, $\lim_{L\to\infty} n_L/L = 0$.

Our goal is to prove $p\lim_{L\to\infty} ||\widehat{\boldsymbol{\delta}}_{n_L} - \boldsymbol{\delta}^0|| = 0$. In Bierens and Song (2012) we have assumed that the infinite-dimensional parameter space involved is compact, and then the consistency of $\widehat{\boldsymbol{\delta}}_{n_L}$ follows straightforwardly from Theorem 4.1 in Bierens (2014b). However, in the present case the consistency proof is much simpler, does not require additional conditions beyond Assumption 7, and does not require to confine $\widehat{\boldsymbol{\delta}}_{n_L}$ to a compact infinite-dimensional subset of Δ, as in Bierens and Song (2012).

The sieve consistency proof employs the following result.

Lemma 4. *For an arbitrary $\varepsilon > 0$, let $\Delta^+(\varepsilon) = \{\boldsymbol{\delta} \in \Delta : ||\boldsymbol{\delta} - \boldsymbol{\delta}^0|| \geq \varepsilon\}$. Then*

$$E\left[\inf_{\boldsymbol{\delta} \in \Delta_{n_L} \cap \Delta^+(\varepsilon)} \widehat{Q}_L(\boldsymbol{\delta})\right] = \inf_{\boldsymbol{\delta} \in \Delta_{n_L} \cap \Delta^+(\varepsilon)} Q(\boldsymbol{\delta}) + o(1).$$

Proof. Appendix

On the basis is this result, it will be shown that $\widehat{\boldsymbol{\delta}}_{n_L}$ is consistent:

Theorem 3. *Let n_L be a subsequence of the number L of auctions such that $\lim_{L\to\infty} n_L = \infty$ and $\lim_{L\to\infty} n_L/L = 0$. Then under Assumptions 1-7 the sieve estimator $\widehat{\boldsymbol{\delta}}_{n_L}$ defined by* (64) *is consistent*: $p\lim_{L\to\infty} ||\widehat{\boldsymbol{\delta}}_{n_L} - \boldsymbol{\delta}^0|| = 0$.

Proof. Appendix.

The crux of the proof of Theorem 3 is three-fold. First the boundedness of $\widehat{Q}_L(\boldsymbol{\delta})$ and $Q(\boldsymbol{\delta})$ is essential. Cf. (57) and (60). Second, the facts that $Q(\boldsymbol{\delta}^0) = 0$ and $\widehat{Q}_L(\widehat{\boldsymbol{\delta}}_{n_L}) \geq 0$ play a crucial role. Third, the a.s. continuity of $\widehat{Q}_L(\boldsymbol{\delta})$ and the continuity of $Q(\boldsymbol{\delta})$ in each $\boldsymbol{\delta} \in \Delta$ is essential. The last condition is also necessary for consistency of sieve estimators in other situations, like the sieve ML approach in Bierens (2014b). Therefore, the result of Theorem 3 is typical for our SNP auction model, and may not be applicable to other SNP models. As to the latter, see Chen (2007) for a review of various SNP models.

Finally, recall that

$$f_0(w) = h_0(G(w))g(w)$$

and

$$\lim_{n\to\infty} \int_0^1 \left| h(u|\pi_n \boldsymbol{\delta}^0) - h_0(u) \right| \mathrm{d}u = 0.$$

Moreover, it is not hard to verify that $p\lim_{L\to\infty} ||\widehat{\boldsymbol{\delta}}_{n_L} - \boldsymbol{\delta}^0|| = 0$ implies

$$p\lim_{L\to\infty} \int_0^1 \left| h(u|\widehat{\boldsymbol{\delta}}_{n_L}) - h(u|\pi_{n_L} \boldsymbol{\delta}^0) \right| \mathrm{d}u = 0,$$

hence

$$p\lim_{L\to\infty} \int_0^1 \left| h(u|\widehat{\boldsymbol{\delta}}_{n_L}) - h_0(u) \right| \mathrm{d}u = 0.$$

Thus, denoting

$$\widehat{f}_L(w) = h(G(w)|\widehat{\boldsymbol{\delta}}_{n_L})g(w),\ \widehat{F}_L(w) = H(G(w)|\widehat{\boldsymbol{\delta}}_{n_L}),$$

we have

$$\int_0^\infty \left| \widehat{f}_L(w) - f_0(w) \right| \mathrm{d}w = \int_0^1 \left| h(u|\widehat{\boldsymbol{\delta}}_{n_L}) - h_0(u) \right| \mathrm{d}u \xrightarrow{\mathrm{p}} 0,$$

$$\sup_{w>0} \left| \widehat{F}_L(w) - F_0(w) \right| \xrightarrow{\mathrm{p}} 0.$$

8 To be continued

9 Appendix: Proofs

9.1 Proof of Lemma 1

Note that for $\varepsilon \in (0,1)$,

$$\begin{aligned}
\int_0^1 (I_k-1)z^{I_k-2}F_0^{-1}\left(U_{i,k}.z\right)\mathrm{d}z &\geq \int_\varepsilon^1 (I_k-1)z^{I_k-2}F_0^{-1}\left(U_{i,k}.z\right)\mathrm{d}z \\
&\geq F_0^{-1}\left(U_{i,k}.\varepsilon\right)\int_\varepsilon^1 (I_k-1)z^{I_k-2}\mathrm{d}z \\
&= F_0^{-1}\left(U_{i,k}.\varepsilon\right)\left(1-\varepsilon^{I_k-1}\right) \\
&\geq F_0^{-1}\left(U_{i,k}.\varepsilon\right)(1-\varepsilon)
\end{aligned}$$

hence

$$\begin{aligned}
Z_{i,k} &\geq \ln\left(F_0^{-1}\left(U_{i,k}.\varepsilon\right)\right)+\ln(1-\varepsilon) \\
&\geq -\left|\ln\left(F_0^{-1}\left(U_{i,k}.\varepsilon\right)\right)\right|+\ln(1-\varepsilon)
\end{aligned}$$

Moreover,

$$\begin{aligned}
\int_0^1 (I_k-1)z^{I_k-2}F_0^{-1}\left(U_{i,k}.z\right)\mathrm{d}z &\leq F_0^{-1}\left(U_{i,k}\right)\int_0^1 (I_k-1)z^{I_k-2}\mathrm{d}z \\
&= F_0^{-1}\left(U_{i,k}\right),
\end{aligned}$$

hence

$$Z_{i,k} \leq \ln\left(F_0^{-1}\left(U_{i,k}\right)\right) \leq \left|\ln\left(F_0^{-1}\left(U_{i,k}\right)\right)\right|.$$

Thus,

$$|Z_{i,k}| \leq \max\left\{\left|\ln\left(F_0^{-1}\left(U_{i,k}\right)\right)\right|, \left|\ln\left(F_0^{-1}\left(U_{i,k}.\varepsilon\right)\right)\right| + \ln\left(1/\left(1-\varepsilon\right)\right)\right\}$$

hence

$$\begin{aligned}
Z_{i,k}^2 &\leq \max\left\{\left(\ln\left(F_0^{-1}\left(U_{i,k}\right)\right)\right)^2, \left(\left|\ln\left(F_0^{-1}\left(U_{i,k}.\varepsilon\right)\right)\right| + \ln\left(1/\left(1-\varepsilon\right)\right)\right)^2\right\} \\
&\leq \left(\ln\left(F_0^{-1}\left(U_{i,k}\right)\right)\right)^2 + 2\left(\ln\left(F_0^{-1}\left(U_{i,k}.\varepsilon\right)\right)\right)^2 + \left(\ln\left(1/\left(1-\varepsilon\right)\right)\right)^2.
\end{aligned}$$

Assumption 4 implies that

$$\begin{aligned}
\infty &> \int_0^{\infty} (\ln(w))^2 f_0(w)\mathrm{d}w \\
&\geq \int_0^{F_0^{-1}(1/\varepsilon)} (\ln(w))^2 f_0(w)\mathrm{d}w \\
&= \int_0^{F_0^{-1}(1/\varepsilon)} \left(\ln\left(F_0^{-1}\left(F_0(w)\right)\right)\right)^2 \mathrm{d}F_0(w) \\
&= \int_0^{1/\varepsilon} \left(\ln\left(F_0^{-1}\left(u\right)\right)\right)^2 \mathrm{d}u \\
&= \varepsilon \int_0^{1} \left(\ln\left(F_0^{-1}\left(\varepsilon.u\right)\right)\right)^2 \mathrm{d}u,
\end{aligned}$$

hence

$$E\left[\left(\ln\left(F_0^{-1}\left(U_{i,k}.\varepsilon\right)\right)\right)^2\right] < \infty.$$

Similarly,

$$\begin{aligned}
\infty &> \int_0^{\infty} (\ln(w))^2 f_0(w)\mathrm{d}w \\
&= \int_0^{\infty} \left(\ln\left(F_0^{-1}\left(F_0(w)\right)\right)\right)^2 \mathrm{d}F_0(w) \\
&= \int_0^{1} \left(\ln\left(F_0^{-1}\left(u\right)\right)\right)^2 \mathrm{d}u.
\end{aligned}$$

hence

$$E\left[\left(\ln\left(F_0^{-1}\left(U_{i,k}\right)\right)\right)^2\right] < \infty.$$

Consequently, Assumption 4 implies that $E[Z_{i,k}^2] < \infty$.

9.2 Proof of Lemma 2

By Assumption 7, $\Pr[||X_k|| \leq C] = 1$ for some constant C, and by Theorem 1, $\lim_{L\to\infty} \Pr[||\widetilde{\beta}_L - \beta_0|| < \varepsilon] = 1$ for every $\varepsilon > 0$. Then by (48) and (51),

$$\varphi_{m,L}(t) = \widehat{p}_{m,L}^{-1} \frac{1}{L} \sum_{k=1}^{L} \mathbf{1}\left(I_k = K_m\right)$$

$$\times \sum_{i=1}^{K_m} \exp\left(-t.\exp\left(-\left(\widetilde{\beta}_L - \beta_0\right)' X_k\right)\exp(-\beta_0' X_k)B_{i,k}\right)$$

$$\leq \widehat{p}_{m,L}^{-1} \frac{1}{L}\sum_{k=1}^{L}\sum_{i=1}^{K_m} \mathbf{1}\left(I_k = K_m\right)\exp\left(-t.\exp\left(-\varepsilon C\right)\exp(-\beta_0' X_k)B_{i,k}\right)$$

with probability converging to 1. By (49) we have,

$$\begin{aligned} &E\left[\mathbf{1}\left(I_k = K_m\right)\exp\left(-t.\exp\left(-\varepsilon C\right)\exp(-\beta_0' X_k)B_{i,k}\right)\right] \\ &\quad = E\left[\exp\left(-t.\exp\left(-\varepsilon C\right)\Lambda(U_{i,k}|K_m, \boldsymbol{\delta}^0)\right)\mathbf{1}\left(I_k = K_m\right)\right] \\ &\quad = \int_0^1 \exp\left(-t.\exp\left(-\varepsilon C\right)\Lambda(u|K_m, \boldsymbol{\delta}^0)\right)\mathrm{d}u \times p_m \\ &\quad = \varphi_m(t\exp\left(-\varepsilon C\right)|\boldsymbol{\delta}^0) \times p_m \end{aligned}$$

so that by the uniform strong law of numbers,

$$\begin{aligned} \sup_{0\leq t\leq 1}\Bigg|\frac{1}{L}\sum_{k=1}^{L}\frac{1}{K_m}\sum_{i=1}^{K_m}\mathbf{1}\left(I_k = K_m\right)\exp\left(-t.\exp\left(-\varepsilon C\right)\exp(-\beta_0' X_k)B_{i,k}\right) \\ - \varphi_m(t\exp\left(-\varepsilon C\right)|\boldsymbol{\delta}^0) \times p_m\Bigg| \overset{\text{a.s.}}{\to} 0 \text{ as } L\to\infty \end{aligned}$$

Since $\operatorname{plim}_{L\to\infty}\widehat{p}_{m,L}^{-1} = p_m^{-1}$ it follows now that with probability converging to 1,

$$\begin{aligned} &\sup_{0\leq t\leq 1}\left|\varphi_{m,L}(t) - \varphi_m(t|\boldsymbol{\delta}^0)\right| \\ &\qquad \leq \sup_{0\leq t\leq 1}\left|\varphi_m(t\exp\left(-\varepsilon C\right)|\boldsymbol{\delta}^0) - \varphi_m(t|\boldsymbol{\delta}^0)\right| \end{aligned}$$

Now for an arbitrary small $\kappa > 0$, choose ε so small that the right-hand side is less than κ, so that

$$\lim_{L\to\infty}\Pr\left[\sup_{0\leq t\leq 1}\left|\varphi_{m,L}(t) - \varphi_m(t|\boldsymbol{\delta}^0)\right| < \kappa\right] = 1.$$

This proves (53). The proof of (54) is similar, by replacing $\boldsymbol{\delta}^0$ by $\boldsymbol{\delta}$ and $\exp(-\beta_0' X_k)B_{i,k}$ by $\Lambda(U_{i,k}|I_k, \boldsymbol{\delta})$.

9.3 Proof of Lemma 3

For $t > 0$ it follows by the dominated convergence theorem that for all nonnegative integers k,

$$\begin{aligned}
\frac{\mathrm{d}^k \mathcal{L}_1(t)}{(\mathrm{d}t)^k} &= (-1)^k E[V_1^k \exp(-tV_1)], \\
\frac{\mathrm{d}^k \mathcal{L}_2(t)}{(\mathrm{d}t)^k} &= (-1)^k E[V_2^k \exp(-tV_2)],
\end{aligned}$$

so that for $t \in T$ and $k = 0, 1, 2, ...,$

$$E[V_1^k \exp(-tV_1)] = E[V_2^k \exp(-tV_2)]. \tag{65}$$

Hence, for any $\delta \in (0, t)$,

$$\begin{aligned}
E\left[\exp(-(t-\delta).V_1\right] &= \sum_{k=0}^{\infty} \frac{\delta^k}{k!} E\left[V_1^k \exp(-t.V_1)\right] \\
&= \sum_{k=0}^{\infty} \frac{\delta^k}{k!} E\left[V_2^k \exp(-t._2)\right] \\
&= E\left[\exp(-(t-\delta).V_2\right]. \qquad (66)
\end{aligned}$$

Therefore, with $t_0 \in T$, (65) holds for all $t \in (0, t_0)$.

Now for all $\xi \in \mathbb{R}$, all $t \in (0, t_0)$, and with $\mathbf{i} = \sqrt{-1}$, it follows similar to (66) that

$$\begin{aligned}
E[\exp(\mathbf{i}\xi V_1)\exp(-t.V_1)] &= \sum_{k=0}^{\infty} \frac{(\mathbf{i}.\xi)^k}{k!} E\left[V_1^k \exp(-t.V_1)\right] \\
&= \sum_{k=0}^{\infty} \frac{(\mathbf{i}.\xi)^k}{k!} E\left[V_2^k \exp(-t.V_2)\right] \\
&= E[\exp(\mathbf{i}\xi V_2)\exp(-t.V_2)],
\end{aligned}$$

hence, letting $t \downarrow 0$, it follows by bounded convergence that

$$E[\exp(\mathbf{i}\xi V_1)] = E[\exp(\mathbf{i}\xi V_2)].$$

Thus, the characteristic functions of V_1 and V_2 are equal, which implies that their distributions are equal.

9.4 Proof of Lemma 4

Denote

$$\widehat{\varphi}^{(1)}_{m,L}(t|\boldsymbol{\delta}) = p_m^{-1}\frac{1}{L}\sum_{k=1}^{L}\frac{1}{K_m}\sum_{i=1}^{K_m}\exp\left(-t.\Lambda(U_{i,k}|K_m,\boldsymbol{\delta})\right)\mathbf{1}\left(I_k = K_m\right)$$

Then

$$\sup_{t\geq 0,\ \boldsymbol{\delta}\in\Delta}\left|\widehat{\varphi}_{m,L}(t|\boldsymbol{\delta}) - \widehat{\varphi}^{(1)}_{m,L}(t|\boldsymbol{\delta})\right| \leq \left|\widehat{p}_{m,L}^{-1} - p_m^{-1}\right| = \frac{|\widehat{p}_{m,L} - p_m|}{p_m.\widehat{p}_{m,L}} = O_p(L^{-1/2}),$$

because $\sqrt{L}(\widehat{p}_{m,L} - p_m) \xrightarrow{\text{d}} N(0, p_m - p_m^2)$, hence

$$\sup_{\boldsymbol{\delta}\in\Delta}\int_0^1\left(\widehat{\varphi}_{m,L}(t|\boldsymbol{\delta}) - \widehat{\varphi}^{(1)}_{m,L}(t|\boldsymbol{\delta})\right)^2\mathrm{d}t = O_p(L^{-1}). \tag{67}$$

Next, denote

$$\widehat{\varphi}^{(2)}_{m,L}(t|\boldsymbol{\delta}) = \frac{1}{L}\sum_{k=1}^{L}\frac{1}{K_m}\sum_{i=1}^{K_m}\exp\left(-t.\Lambda(U^*_{i,k}|K_m,\boldsymbol{\delta})\right).$$

Then

$$\begin{aligned}
&E\left[\int_0^1\left(\widehat{\varphi}^{(1)}_{m,L}(t|\boldsymbol{\delta}) - \widehat{\varphi}^{(2)}_{m,L}(t|\boldsymbol{\delta})\right)^2\mathrm{d}t\right]\\
&= p_m^{-2}E\left[\left(\frac{1}{L}\sum_{k=1}^{L}\frac{1}{K_m}\sum_{i=1}^{K_m}\exp\left(-t.\Lambda(U^*_{i,k}|K_m,\boldsymbol{\delta})\right)\left(\mathbf{1}\left(I_k = K_m\right) - p_m\right)\right)^2\right]\\
&\leq p_m^{-2}\frac{1}{L^2}\sum_{k=1}^{L}\int_0^1 E\left[\frac{1}{K_m}\sum_{i=1}^{K_m}\exp\left(-t.\Lambda(U^*_{i,k}|K_m,\boldsymbol{\delta})\right)^2\mathrm{d}t\right.\\
&\qquad\qquad\left.\times\left(\mathbf{1}\left(I_k = K_m\right) - p_m\right)^2\right]\\
&\leq p_m^{-2}\frac{1}{L^2}\sum_{k=1}^{L}E\left[\left(\mathbf{1}\left(I_k = K_m\right) - p_m\right)^2\right] = (p_m^{-1} - 1)/L,
\end{aligned}$$

so that

$$E\left[\int_0^1\left(\widehat{\varphi}^{(1)}_{m,L}(t|\boldsymbol{\delta}) - \widehat{\varphi}^{(2)}_{m,L}(t|\boldsymbol{\delta})\right)^2\mathrm{d}t\right] = O(L^{-1}). \tag{68}$$

Moreover, it follows trivially that

$$E\left[\int_0^1 \left(\widehat{\varphi}^{(2)}_{m,L}(t|\boldsymbol{\delta}) - \varphi(t|\boldsymbol{\delta})\right)^2 \mathrm{d}t\right] = O(L^{-1}). \tag{69}$$

and it follows from Lemma 2 that $p\lim_{L\to\infty}\int_0^1(\varphi_{m,L}(t) - \varphi_m(t))^2\mathrm{d}t = 0$, which by bounded convergence implies

$$\lim_{L\to\infty} E\left[\int_0^1 (\varphi_{m,L}(t) - \varphi_m(t))^2\mathrm{d}t\right] = 0. \tag{70}$$

It follows now straightforwardly from (67) through (70) that

$$\begin{aligned} E\left[\widetilde{Q}_L(\boldsymbol{\delta})\right] &= \frac{1}{M-1}\sum_{m=2}^{M}\int_0^1 \left(\varphi_m(t|\boldsymbol{\delta}) - \varphi_m(t)\right)^2 \mathrm{d}t + o(1) \\ &= Q(\boldsymbol{\delta}) + o(1). \end{aligned} \tag{71}$$

However, the $o(1)$ term depends on $\boldsymbol{\delta}$ because the $O(L^{-1})$ terms in (68) and (69) depend on $\boldsymbol{\delta}$. On the other hand, (68) and (69) imply

$$\begin{aligned} E\left[\inf_{\boldsymbol{\delta}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)}\int_0^1 \left(\widehat{\varphi}^{(1)}_{m,L}(t|\boldsymbol{\delta}) - \widehat{\varphi}^{(2)}_{m,L}(t|\boldsymbol{\delta})\right)^2 \mathrm{d}t\right] &= O(L^{-1}), \\ E\left[\inf_{\boldsymbol{\delta}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)}\int_0^1 \left(\widehat{\varphi}^{(2)}_{m,L}(t|\boldsymbol{\delta}) - \varphi(t|\boldsymbol{\delta})\right)^2 \mathrm{d}t\right] &= O(L^{-1}), \end{aligned}$$

and now these $O(L^{-1})$ terms no longer depend on $\boldsymbol{\delta}$. Consequently,

$$E\left[\inf_{\boldsymbol{\delta}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)}\widehat{Q}_L(\boldsymbol{\delta})\right] = \inf_{\boldsymbol{\delta}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)} Q(\boldsymbol{\delta}) + o(1).$$

9.5 Proof of Theorem 3

Note first that

$$0 \le \widehat{Q}_L(\widehat{\boldsymbol{\delta}}_{n_L}) \le \widehat{Q}_L(\pi_{n_L}\boldsymbol{\delta}^0), \tag{72}$$

provided that L is so large that $\pi_{n_L}\boldsymbol{\delta}^0 \in \Delta_{n_L}$. This the case for L so large that $||\boldsymbol{\delta}^0||^2 \le B_{n_L}$ so that $||\pi_{n_L}\boldsymbol{\delta}^0||^2 \le B_{n_L}$ as well.

Next, it follows from Lemma 3 in Bierens and Song (2012) that

$$p\lim_{L\to\infty}\left|\widehat{Q}_L(\pi_{n_L}\boldsymbol{\delta}^0) - \widehat{Q}_L(\boldsymbol{\delta}^0)\right| = 0, \tag{73}$$

so that by (58),

$$p\lim_{L\to\infty}\widehat{Q}_L(\boldsymbol{\delta}^0)=Q(\boldsymbol{\delta}^0)=0, \tag{74}$$

hence by (72), $p\lim_{L\to\infty}\widehat{Q}_L(\widehat{\boldsymbol{\delta}}_{n_L})=0$ and therefore also

$$p\lim_{L\to\infty}\widehat{Q}_L(\widehat{\boldsymbol{\delta}}_{n_L})^2=0. \tag{75}$$

Since convergence in probability of bounded random variables implies convergence in expectation,[4] it follows now from (57) and (75) that

$$\lim_{L\to\infty}E[\widehat{Q}_L(\widehat{\boldsymbol{\delta}}_{n_L})^2]=0. \tag{76}$$

Thus, with $\Delta^+(\varepsilon)$ defined in Lemma 3, we have

$$\begin{aligned}
E\left[\widehat{Q}_L(\widehat{\boldsymbol{\delta}}_{n_L})^2\right] &= E\left[\widehat{Q}_L(\widehat{\boldsymbol{\delta}}_{n_L})^2\mathbf{1}\left(\widehat{\boldsymbol{\delta}}_{n_L}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)\right)\right] \\
&\quad +E\left[\widehat{Q}_L(\widehat{\boldsymbol{\delta}}_{n_L})^2\mathbf{1}\left(\widehat{\boldsymbol{\delta}}_{n_L}\in\Delta_{n_L}\backslash\Delta^+(\varepsilon)\right)\right] \\
&\geq E\left[\left(\inf_{\boldsymbol{\delta}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)}\widehat{Q}_L(\boldsymbol{\delta})\right)^2.\mathbf{1}\left(\widehat{\boldsymbol{\delta}}_{n_L}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)\right)\right] \\
&\geq \sqrt{E\left[\inf_{\boldsymbol{\delta}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)}\widehat{Q}_L(\boldsymbol{\delta})\right]}\times\Pr\left[\widehat{\boldsymbol{\delta}}_{n_L}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)\right] \\
&= \sqrt{\inf_{\boldsymbol{\delta}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)}Q(\boldsymbol{\delta})}\times\Pr\left[\widehat{\boldsymbol{\delta}}_{n_L}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)\right]+o(1) \\
&\geq \sqrt{\inf_{\boldsymbol{\delta}\in\Delta^+(\varepsilon)}Q(\boldsymbol{\delta})}\times\Pr\left[\widehat{\boldsymbol{\delta}}_{n_L}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)\right]+o(1),
\end{aligned}$$

where the second inequality follows from Schwarz inequality and the last equality follows from Lemma 4. Since $\inf_{\boldsymbol{\delta}\in\Delta^+(\varepsilon)}Q(\boldsymbol{\delta})>0$, it follows from (76) that $\lim_{L\to\infty}\Pr[\widehat{\boldsymbol{\delta}}_{n_L}\in\Delta_{n_L}\cap\Delta^+(\varepsilon)]=0$, which by $\Pr[\widehat{\boldsymbol{\delta}}_{n_L}\in\Delta_{n_L}]=1$ implies that

$$\lim_{L\to\infty}\Pr\left[\widehat{\boldsymbol{\delta}}_{n_L}\in\Delta^+(\varepsilon)\right]=0.$$

Since $\varepsilon>0$ was arbitrary, the theorem under review follows.

[4] See, for example, Theorem 6.4 in Bierens (2004).

References

Bierens, H. J.. (2004), *Introduction to the Mathematical and Statistical Foundations of Econometrics.* Cambridge University Press.

Bierens, H. J. (2008), "Semi-Nonparametric Interval Censored Mixed Proportional Hazard Models: Identification and Consistency Results", *Econometric Theory* 24, 749-794.

Bierens, H. J. (2014a), "The Hilbert Space Theoretical Foundation of Semi-Nonparametric Modeling", in J. Racine, L. Su and A. Ullah (eds.), *The Oxford Handbook of Applied Nonparametric and Semiparametric Econometrics and Statistics*, Chapter 1, Oxford University Press.

Bierens, H. J. (2014b), "Consistency and Asymptotic Normality of Sieve ML Estimators Under Low-Level Conditions", *Econometric Theory* 30, 1021-1076.

Bierens, H. J. and J. Carvalho (2007), "Semi-Nonparametric Competing Risks Analysis of Recidivism", *Journal of Applied Econometrics* 22, 971-993.

Bierens, H .J. and H. Song (2012), "Semi-Nonparametric Estimation of Independently and Identically Repeated First-Price Auctions via an Integrated Simulated Moments Method", *Journal of Econometrics,* 168, 108-119.

Chen, X. (2007), "Large sample sieve estimation of semi-nonparametric models, in J. Heckman and E. Leamer (eds.), *Handbook of Econometrics, Vol. 6,* Chapter 76, Elsevier.

Donald, G. S. S. and J. H. Paarsch (1996), "Identification, Estimation, and Testing in Parametric Empirical Models of Auctions within the Independent Private Values Paradigm", *Econometric Theory*, 12, 517-567.

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

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

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

Li, T. (2005), "Econometrics of First-Price Auctions with Entry and Binding Reservation Prices", *Journal of Econometrics*, 126, 173-200.

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

10 Scratch 1

$$\begin{aligned}
\varphi_{m,L}(t) &= \widehat{p}_{m,L}^{-1} \frac{1}{L} \sum_{k=1}^{L} \frac{1}{K_m} \sum_{i=1}^{K_m} \exp\left(-t.Y_{i,k}\right) \mathbf{1}\left(I_k = K_m\right), \\
\widehat{\varphi}_{m,L}(t|\boldsymbol{\delta}) &= \widehat{p}_{m,L}^{-1} \frac{1}{L} \sum_{k=1}^{L} \frac{1}{K_m} \sum_{i=1}^{K_m} \exp\left(-t.\widetilde{Y}_{i,k}(\boldsymbol{\delta})\right) \mathbf{1}\left(I_k = K_m\right),
\end{aligned}$$

$$\begin{aligned}
& \varphi_{m,L}(t) - \widehat{\varphi}_{m,L}(t|\boldsymbol{\delta}) \\
= \; & \widehat{p}_{m,L}^{-1} \frac{1}{L} \sum_{k=1}^{L} \frac{1}{K_m} \sum_{i=1}^{K_m} \left(\exp\left(-t.Y_{i,k}\right) - \exp\left(-t.\widetilde{Y}_{i,k}(\boldsymbol{\delta})\right)\right) \mathbf{1}\left(I_k = K_m\right)
\end{aligned}$$

$$\begin{aligned}
& \left(\varphi_{m,L}(t) - \widehat{\varphi}_{m,L}(t|\boldsymbol{\delta})\right)^2 \\
= \; & \widehat{p}_{m,L}^{-2} \frac{1}{L^2} \sum_{k_1=1}^{L} \sum_{k_2=1}^{L} \frac{1}{K_m^2} \\
& \times \sum_{i_1=1}^{K_m} \left(\exp\left(-t.Y_{i_1,k_1}\right) - \exp\left(-t.\widetilde{Y}_{i_1,k_1}(\boldsymbol{\delta})\right)\right) \mathbf{1}\left(I_{k_1} = K_m\right) \\
& \times \sum_{i_2=1}^{K_m} \left(\exp\left(-t.Y_{i_2,k_{21}}\right) - \exp\left(-t.\widetilde{Y}_{i_2,k_2}(\boldsymbol{\delta})\right)\right) \mathbf{1}\left(I_{k_2} = K_m\right) \\
= \; & \widehat{p}_{m,L}^{-2} \frac{1}{L^2} \sum_{k_1=1}^{L} \sum_{k_2=1}^{L} \frac{1}{K_m^2}
\end{aligned}$$

11 Scratch 2

Let

$$\begin{aligned}
Y &= \exp(-\beta_0' X) B \sim \frac{1}{U^{I_X - 1}} \int_0^U (I_X - 1) z^{I_X - 2} F_0^{-1}(z) \mathrm{d}z \\
&= \int_0^1 (I_X - 1) z^{I_X - 2} F_0^{-1}(U.z) \mathrm{d}z \sim \widetilde{Y}(F_0, I_X)
\end{aligned}$$

$$\begin{aligned}\widetilde{Y}(F, I_X) &= \frac{1}{\widetilde{U}^{I_X-1}} \int_0^{\widetilde{U}} (I_X - 1) z^{I_X-2} F^{-1}(z) \mathrm{d}z \\ &= \int_0^1 (I_X - 1) z^{I_X-2} F^{-1}(\widetilde{U}.z) \mathrm{d}z\end{aligned}$$

where U and $\widetilde{U}$ are independently $\mathcal{U}(0,1)$ distributed.

The conditional characteristic function of $\widetilde{Y}(F_0, I_X)$ given I_X is

$$E\left[\exp(\mathbf{i}.(t.\widetilde{Y}(F_0, I_X))|I_X\right] = \int_0^1 \exp\left(\mathbf{i}.t. \int_0^1 (I_X - 1) z^{I_X-2} F_0^{-1}(u.z) \mathrm{d}z\right) \mathrm{d}u$$

whereas the conditional characteristic function of $\widetilde{Y}(F, I_X)$ given I_X is

$$E\left[\exp(\mathbf{i}.(t.\widetilde{Y}(F, I_X))|I_X\right] = \int_0^1 \exp\left(\mathbf{i}.t. \int_0^1 (I_X - 1) z^{I_X-2} F^{-1}(u.z) \mathrm{d}z\right) \mathrm{d}u$$

If these conditional characteristic functions are equal then we must have that conditional on I_X,

$$\widetilde{Y}(F_0, I_X) \sim \widetilde{Y}(F, I_X)$$

which in its turn implies that for all $u \in [0,1]$, and given I_X,

$$\int_0^1 (I_X - 1) z^{I_X-2} F_0^{-1}(u.z) \mathrm{d}z = \int_0^1 (I_X - 1) z^{I_X-2} F^{-1}(u.z) \mathrm{d}z$$

See Remark 1 below. Recall that this equality can also be written as

$$u^{1-I_X} \int_0^u (I_X - 1) z^{I_X-2} F_0^{-1}(z) \mathrm{d}z = u^{1-I_X} \int_0^u (I_X - 1) z^{I_X-2} F^{-1}(z) \mathrm{d}z$$

hence

$$\int_0^u (I_X - 1) z^{I_X-2} F_0^{-1}(z) \mathrm{d}z = \int_0^u (I_X - 1) z^{I_X-2} F^{-1}(z) \mathrm{d}z$$

Taking the derivative to u yields

$$(I_X - 1) u^{I_X-2} F_0^{-1}(u) = (I_X - 1) u^{I_X-2} F^{-1}(u)$$

hence $F_0^{-1}(u) = F^{-1}(u)$ for all $u \in [0,1]$, which implies that $F \equiv F_0$.

Remark 1. Let $\rho_1(u)$ and $\rho_2(u)$ be two real functions on $(0,1)$ satisfying

$$\int_0^1 |\rho_1(u)|\mathrm{d}u < \infty, \ \int_0^1 |\rho_2(u)|\mathrm{d}u < \infty,$$

such that for two random drawings U_1 and U_2 from $\mathcal{U}(0,1)$,

$$\rho_1(U_1) \sim \rho_2(U_2)$$

Then for all $t \in \mathbb{R}$,

$$E\left[\rho_1(U_1)\exp(\mathbf{i}.t.U_1)\right] = E\left[\rho_2(U_2)\exp(\mathbf{i}.t.U_2)\right]$$

hence

$$\int_0^1 (\rho_1(u) - \rho_2(u))\exp(\mathbf{i}.t.u)\mathrm{d}u = 0,$$

which by Theorem 1 in Bierens (1982) implies that $\rho_1(u) = \rho_2(u)$ a.e. on $[0,1]$.

$$\begin{aligned}
& \left(\widetilde{Y}(F_0, I_X) - \widetilde{Y}(F, I_X)\right)^2 \\
& \left(\int_0^1 (I_X - 1)z^{I_X-2}F^{-1}(\widetilde{U}.z)\mathrm{d}z - \int_0^1 (I_X - 1)z^{I_X-2}F_0^{-1}(U.z)\mathrm{d}z\right)^2 \\
= & \left(\int_0^1 (I_X - 1)z^{I_X-2}F^{-1}(\widetilde{U}.z)\mathrm{d}z\right)^2 + \left(\int_0^1 (I_X - 1)z^{I_X-2}F_0^{-1}(U.z)\mathrm{d}z\right)^2 \\
& -2\int_0^1 (I_X - 1)z^{I_X-2}F^{-1}(\widetilde{U}.z)\mathrm{d}z \int_0^1 (I_X - 1)z^{I_X-2}F_0^{-1}(U.z)\mathrm{d}z
\end{aligned}$$

$$\begin{aligned}
& E\left[\left(\widetilde{Y}(F_0, I_X) - \widetilde{Y}(F, I_X)\right)^2 |I_X\right] \\
= & \int_0^1 \left(\int_0^1 (I_X - 1)z^{I_X-2}F^{-1}(uz)\mathrm{d}z\right)^2 \mathrm{d}u + \int_0^1 \left(\int_0^1 (I_X - 1)z^{I_X-2}F_0^{-1}(u.z)\mathrm{d}z\right) \mathrm{d}u \\
& -2\int_0^1 \int_0^1 (I_X - 1)z^{I_X-2} \int_0^1 F^{-1}(uz)\mathrm{d}z\mathrm{d}u \times \int_0^1 \int_0^1 (I_X - 1)z^{I_X-2} \int_0^1 F_0^{-1}(uz)\mathrm{d}z\mathrm{d}u
\end{aligned}$$

12 Scratch 3

$f_0(w) = 2w\exp(-w^2)$. In this case $\int_0^\infty wf_0(w)\mathrm{d}w = \sqrt{2\pi}$?

$$\begin{aligned}
\int_0^\infty wf_0(w)\mathrm{d}w &= 2\int_0^\infty w^2\exp(-w^2)\mathrm{d}w \\
&= \sqrt{2}\int_0^\infty w^2\exp(-w^2)\mathrm{d}\left(w\sqrt{2}\right) \\
x &= w\sqrt{2}, w = x/\sqrt{2} \\
&= \sqrt{2}\int_0^\infty (x^2/2)\exp(-x^2/2)\mathrm{d}x \\
&= \frac{1}{\sqrt{2}}\int_0^\infty x^2\exp(-x^2/2)\mathrm{d}x \\
&= \frac{1}{2\sqrt{2}}\int_{-\infty}^\infty x^2\exp(-x^2/2)\mathrm{d}x \\
&= \frac{\sqrt{2\pi}}{2\sqrt{2}}\int_{-\infty}^\infty x^2\frac{\exp(-x^2/2)}{\sqrt{2\pi}}\mathrm{d}x \\
&= \sqrt{\pi}/2
\end{aligned}$$

13 Scratch 4

For example, suppose that the $\ln(W_{i,k})$'s have a standard logistic distribution, so that $F_0(w) = (1+w)^{-1}w$ with density $f_0(w) = (1+w)^{-2}$ and inverse $F_0^{-1}(u) = u/(1-u)$. Then

$$\int_0^\infty wf_0(w)\mathrm{d}w = \int_0^1 F_0^{-1}(u)\mathrm{d}u = \infty.$$

Check

$$\Pr\left[\ln(W_{i,k}) \leq y\right] = \frac{1}{1+\exp(-y)}$$

$$\begin{aligned}
F_0(w) &= \Pr\left[W_{i,k} \leq w\right] = \Pr\left[\ln(W_{i,k}) \leq \ln(w)\right] \\
&= \frac{1}{1+\exp(-\ln(w))} = \frac{1}{1+\exp(\ln(1/w))} \\
&= \frac{1}{1+1/w} = \frac{w}{1+w}
\end{aligned}$$

$$f_0(w) = \frac{1}{1+w} - \frac{w}{(1+w)^2} = \frac{1}{(1+w)^2}$$

$$\begin{aligned}
\frac{w}{1+w} &= u \\
w &= (1+w)u = u + w.u \\
(1-u)w &= u \\
F_0^{-1}(u) &= \frac{u}{1-u}
\end{aligned}$$

For $M \in (0,1)$

$$\begin{aligned}
\int_0^M F_0^{-1}(u)\mathrm{d}u &= \int_0^M \frac{u}{1-u}\mathrm{d}u \\
&\geq \frac{1}{1-M}\int_0^M u\mathrm{d}u = \frac{1}{2}\frac{M^2}{1-M} \\
&\to \infty \text{ as } M \uparrow 1
\end{aligned}$$