Semiparametric Estimation of First-Price Auction Models

Aryal, Gaurab and Gabrielli, Maria F. and Vuong, Quang

University of Chicago, CONICET and Universidad Nacional de Cuyo, New York University

12 July 2014

Semiparametric Estimation of First-Price Auction Models *

Gaurab Aryal
University of Chicago
aryalg@uchicago.edu

Maria Florencia Gabrielli †
CONICET &
Universidad Nacional de Cuyo
florgabrielli@gmail.com

Quang Vuong
New York University
qvuong@nyu.edu

Abstract

We propose a semiparametric estimator within the class of indirect methods. Specifically, we model private valuations through a set of conditional moment restrictions. Our econometric model calls for a two step procedure. In the first step we recover a sample of pseudo private values while using a Local Polynomial Estimator. In the second step we use a GMM procedure to obtain an estimate for the parameter of interest. The proposed semiparametric estimator is shown to have desirable statistical properties namely, it is consistent and has an asymptotic normal distribution. Moreover, the estimator attains the parametric rate of convergence.

Keywords: Auctions, Structural Approach, Semiparametric Estimator, Local Polynomial, GMM.
JEL Codes: C14, C71, D44.

*We thank Isabelle Perrigne for her insightful discussions and comments. The usual caveats apply.
†Corresponding Author

1 Introduction

From a theoretical point of view, auctions are modeled as games of incomplete information in which asymmetric information among players (seller/buyer and bidders) is one of the key features. See for example Krishna (2002), McAfee and McMillan (1987) and Wilson (1992). From an applied perspective, since auctions are widely used mechanisms to allocate goods and services which are often public, many data sets are available for empirical research. By assuming that observed bids are the equilibrium outcomes of the underlying auction model under consideration, the structural approach to analyze auction data provides a framework in which the theoretical model and its empirical counterpart are closely related. The main objective of this approach is then to recover the structural elements of the auction model. This line of research has been considerably developed in the last fifteen years. The difficulties in estimating auction models are many. First, auction models lead to highly nonlinear econometric models through the equilibrium strategies. Second, auction models may not lead to tractable solutions rendering even more difficult the derivation of an econometric model. Third, the estimation of auction models often requires the numerical computation of the equilibrium strategy and its inverse.

As documented by Perrigne and Vuong (1999, 2008) several contributions to the structural estimation of first-price auction models are available in the literature.[1] We distinguish two kinds of methods for estimating structural auction models: Direct methods and indirect methods. Direct Methods were first developed in the literature relying on parametric econometric models. Starting from a specification of the underlying distribution of private values, the objective of direct methods is to estimate the parameter vector characterizing such a distribution. Within this class of methods, there are between two major estimation procedures. The first methodology introduced by Paarsch (1992) and Donald and Paarsch (1993) is a fully parametric setup that uses Maximum Likelihood (ML)-based estimation procedures requiring the computation of the equilibrium strategy. This in turn could be highly computationally demanding, as recognized by Donald and Paarsch (1993),

[1] See also Paarsch and Hong (2006) for an extensive survey on structural estimation of auction models within the Independent Private Value (IPV) paradigm.

and thus only very simple distributions are considered in practice. In particular, because the upper bound of the bid distribution depends on the parameter(s) of the underlying distribution, the ML estimator has a nonstandard limiting distribution. In view of this, Donald and Paarsch (1993) develop a so-called piecewise pseudo ML estimator requiring the computation of the equilibrium strategy that can be obtained using specific parametric distribution(s). Laffont, Ossard and Vuong (1995) introduced a second methodology, which is more computationally convenient. Relying on the revenue equivalence theorem, the authors propose a simulation-based method that avoids computation of the equilibrium strategy and therefore allows for more general parametric specifications for the private value distribution.

More recently Guerre, Perrigne and Vuong (2000) (GPV (2000) hereafter) have developed a fully nonparametric indirect procedure introducing the use of indirect methods for the structural estimation of auction models. This alternative methodology relies on a simple but crucial observation, namely each private value can be expressed as a function of the corresponding bid, the distribution of observed bids and its density using the first-order condition of the bidder's optimization problem. Based on this equation, which defines the inverse of the equilibrium strategy, the authors show that the model is nonparametrically identified. Other papers by the same authors and others follow entertaining other auction models in a similar fashion such as models with affiliation among private values, models with asymmetric bidders, dynamic auction models and models with risk averse bidders. Therefore, in contrast to direct methods, indirect methods start from the distribution of observed bids in order to estimate the distribution of unobserved private values without computing the Bayesian Nash equilibrium strategy or its inverse explicitly. This calls naturally for a two step procedure. In the first step, a sample of pseudo private valuations is obtained while using (say) kernel estimators for the distribution and density of observed bids. With this pseudo sample at hand, the second step consists in estimating nonparametrically the density of bidders' private values. GPV (2000) also establish some asymptotic properties of their estimator, namely its uniform consistency and the achievement of the optimal consistency rate by appropriate vanishing rates for bandwidths, while the optimal rate is derived using the minimax theory as developed by Ibragimov and Has'minskii (1981).

Though a fully nonparametric estimator has some desirable properties such as flexibility and robustness to misspecification, it has a number of drawbacks such as a slow consistency rate and the difficulty to consider a large number of covariates by the so-called curse of dimensionality. In this paper, we propose a semiparametric estimator within the class of indirect methods. In the same spirit as GPV (2000) our procedure is in two steps. Unlike GPV (2000), our second step is fully parametric and therefore our resulting model falls within the category of semiparametric models. Specifically, we propose to model private valuations through a set of conditional moment restrictions.[2]

For sake of simplicity, we consider a symmetric first-price sealed-bid auction model within the independent private value (IPV) paradigm with a nonbinding reserve price. As shown below, our estimation procedure applies to a more general class of auction models, namely symmetric and asymmetric affiliated private value models. More generally, our method extends to models which have been estimated using a nonparametric indirect procedure. Let $V_{p\ell}$, $p = 1, \ldots, I_\ell$, $\ell = 1, \ldots, L$ denote the private value of the pth bidder for the ℓth auctioned object. Let $Z_\ell \equiv (X_\ell, I_\ell) \in I\!R^{d+1}$ denote the vector of exogenous variables, it includes the number of bidders I_ℓ and variables X_ℓ characterizing object heterogeneity across auctions.[3] We model private values by the following set of conditional moment restrictions

$$\mathrm{E}[M(V, Z; \theta_0)|Z] = 0, \tag{1}$$

for some known function $M(\cdot, \cdot; \theta) : I\!R^{d+2} \to I\!R^q$ and $\theta \in I\!R^p$. However, such moment restrictions are infeasible in practice since private values are unobserved.

[2]A noticeable exception is found in Jofre-Bonet and Pesendorfer (2003) which use a fully parametric indirect procedure to entertain a large number of covariates. The authors do not provide, however, any asymptotic properties for their estimator.

[3]The dependence of the private value distribution on I_ℓ captures the idea that private values and the number of bidders can be dependent in general. For instance, objects of higher value may attract more bidders. The number of bidders may capture some unobserved heterogeneity. It may also result from endogenous participation. Note also that the reserve price is nonbinding so that the number of potential bidders I_ℓ in the ℓth auction is known. This assumption can be relaxed and our results can be straightforwardly extended in a similar way as in GPV (2000, Section 4).

From the theoretical model we know that the equilibrium bid, B, can be expressed as $B = s(V, Z; \theta_0, \gamma_0)$, where γ_0 could be an infinite dimensional parameter. Therefore the equilibrium strategy depends on the parameter vector θ_0 both directly since $B \sim G(\cdot|Z; \theta_0, \gamma_0)$ (say) and indirectly through V since $V \sim F(\cdot|Z; \theta_0, \gamma_0)$. A natural way of expressing the above conditions would be

$$\mathrm{E}\{M[s^{-1}(B, Z; \theta_0, \gamma_0), Z; \theta_0]|Z\} = 0.$$

The above condition requires the computation of the equilibrium strategy as well as of its inverse. This could be highly computationally demanding for two different reasons. First, such computation has to be carried out for any trial value of the parameters (θ, γ). When γ is an infinite dimensional parameter the model falls within the class of models analyzed by Ai and Chen (2003). Second, in a more general class of models, the computation of the equilibrium strategy $s(\cdot, \cdot; \theta_0, \gamma_0)$ and of its inverse is much more involved such as for affiliated private value models. Furthermore, in asymmetric models this computation becomes intractable.

We therefore propose to replace V in (1) by its nonparametric estimator $\hat{V} = \hat{\xi}(B, Z)$ to make the moment condition operational. Thus the "conditional moment restriction" that we use in our second step is as follows

$$\mathrm{E}\{M[\hat{\xi}(B, Z), Z; \theta_0]|Z\} \approx 0.$$

It is worth noting that $\hat{V} = \hat{\xi}(B, Z)$ is obtained nonparametrically by means of local polynomial fitting. Unlike other nonparametric estimators, such us kernels, local polynomial fitting is an attractive method from a theoretical and practical point of view. As pointed out by Fan and Gijbels (1995), local polynomial estimators (LPE) have some advantages over other commonly used nonparametric estimators including the absence of boundary effects. As it is well known, many nonparametric estimators are ill-behaved close to the boundaries of the support. A common way to deal with this problem is to trim out observations in these problematic regions. In this paper, we use a LPE which is not subject to boundary effects and therefore we do not have to trim out observations. This is a remarkable advantage of our procedure since otherwise the trimming would have to be performed on the endogenous variable of the model introducing an additional technical difficulty because this trimming on the bids, which implies an automatic trimming on private values, will affect the moments of the latter. In a standard econometric framework the trimming is usually applied on the exogenous variables. See for instance, Lavergne and Vuong (1996) and Robinson (1988).

In line with GPV (2000) our econometric model calls for a two step procedure. The first step is similar to the first step in GPV (2000) in which we recover a sample of pseudo private values while using a LPE. The second step departs from the fully nonparametric second step in GPV (2000) since we use instead a GMM procedure to obtain an estimate for θ_0. Thus our procedure is semiparametric.

We establish the consistency and asymptotic normality of our semiparametric estimator. Specifically, we show that our estimator converges uniformly at the parametric $\sqrt{L}$ rate. As it is well known, nonparametric estimators converge at a slower rate than $\sqrt{L}$ and their rates are usually negatively related to the dimension of the vector of exogenous variables. This makes these estimators less desirable in applications, especially when a limited number of observations is available and/or when the number of exogenous variables is relatively large.[4] The estimator we propose in this paper does not share this drawback since it is not subject to the so-called curse of dimensionality. In other words, the convergence rate of our estimator is independent of the dimension of the exogenous variables. A second major advantage of our estimation procedure is that it can be used to estimate more general auction models. Indirect methods in general do not require neither the computation of the equilibrium strategy nor of its inverse. Therefore these methods are specially convenient when there is no closed form solution to the differential equation(s) characterizing the equilibrium

[4]Examples of semiparametric estimators attaining $\sqrt{L}$ rate can be found in Newey and McFadden (1994) and Powell (1994). Some notable exceptions are the estimators proposed by Manski (1985), Horowitz (1992), Kryriazidou (1997) and Honoré and Kryriazidou (2000). An example of a semiparametric estimator converging at a slower than the parametric rate but not subject to the curse of dimensionality, i.e. its rate is independent of d, is given by Campo, Guerre, Perrigne and Vuong (2006).

strategy as is the case in asymmetric auction models which lead to intractable expressions for the first-order conditions.

The rest of the paper is organized as follows. In Section 2, we introduce the theoretical model from which the structural econometric model is derived as well as our semiparametric two-step estimator. Section 3 establishes the asymptotic properties of our estimator and presents some Monte Carlo experiments to illustrate the properties of our procedure in small samples and to assess its advantages relative to a nonparametric procedure. Section 4 discusses how to extend our procedure in a more general class of auction models. Section 5 concludes and indicates some future lines of research. An appendix collects the proofs of our results.

2 The Model

2.1 The Symmetric IPV Model

We present the benchmark theoretical model underlying our structural econometric model, namely the symmetric IPV model with a nonbinding reserve price. Although this model is somehow restrictive for applications, it allows us to develop in a more transparent way our econometric procedure. We postpone until Section 4 possible extensions. A single and indivisible object is auctioned to I_ℓ bidders who are assumed to be ex ante identical (i.e. the game is symmetric) and risk neutral. We allow for the possibility that the total number of bidders varies across auctions as well. Unobserved private values are denoted by V and more precisely we assume that each valuation $V_{p\ell}$, $\ell = 1, \ldots, L$, $p = 1, \ldots, I_\ell$, is distributed according to $F(\cdot|Z_\ell; \theta_0, \gamma_0)$, where $\theta_0 \in I\!R^p$ is the parameter of interest and γ_0 could be in principle infinite or finite dimensional or even an empty set. The support of $F(\cdot|\cdot)$ is $[\underline{V}_\ell, \overline{V}_\ell]$, with $0 \leq \underline{V}_\ell = \underline{V}(Z_\ell) < \overline{V}_\ell = \overline{V}(Z_\ell) < \infty$. Among others, Riley and Samuelson (1981) have characterized the unique symmetric differentiable Bayesian Nash equilibrium. In particular, for every ℓ, $I_\ell \geq 2$ the equilibrium bid $B_{p\ell}$ in the ℓth auction is

$$B_{p\ell} = s_0(V_{p\ell}, Z_\ell) = V_{p\ell} - \frac{1}{F(V_{p\ell}|Z_\ell; \theta_0, \gamma_0)^{I_\ell - 1}} \int_{\underline{v}_\ell}^{V_{p\ell}} F(v|Z_\ell; \theta_0, \gamma_0)^{I_\ell - 1} dv, \tag{2}$$

for any $V_{p\ell}$ since $p_0 < \underline{v}(Z_\ell)$ subject to the boundary condition $s(\underline{v}_\ell) = \underline{v}_\ell$.

The distribution and density of observed bids in the ℓth auction is given by $G(\cdot|Z_\ell;\theta_0,\gamma_0) \equiv G_0(\cdot|Z_\ell)$ and $g(\cdot|Z_\ell;\theta_0,\gamma_0) \equiv g_0(\cdot|Z_\ell)$, respectively. Observed bids are assumed to be the equilibrium outcome of the game. From GPV (2000), the observed bids and private values are related by the following equilibrium expression

$$V_{p\ell} = \xi_0(B_{p\ell}, Z_\ell) = B_{p\ell} + \frac{1}{I_\ell - 1}\frac{G_0(B_{p\ell}|Z_\ell)}{g_0(B_{p\ell}|Z_\ell)}, \tag{3}$$

for $\ell = 1,\dots,L$, $p = 1,\dots,I_\ell$. Equation (3) constitutes the basis for the identification result in GPV (2000), i.e. the authors show that the model is nonparametically identified.

2.2 The Two Step Estimator

In line with GPV (2000), equation (3) is the basis for our econometric model. The difference with GPV (2000) is to model private values as a set of moment conditions. Therefore knowledge of $G_0(\cdot|\cdot)$ and $g_0(\cdot|\cdot)$ would lead us to a GMM framework. However, these functions are unknown in practice but can be easily estimated from observed bids. This suggests the following two-step procedure.

The first step is similar to the first step in GPV (2000) in which we recover a sample of pseudo private values by using nonparametric local polynomial estimators. The second step departs from the nonparametric second step of GPV (2000) since we use instead a GMM procedure to obtain an estimator for θ_0. Before presenting our two-step estimator, it is worth mentioning that some of our assumptions are similar or even identical to those in GPV (2000). This is not surprising since our methodology follows closely their methodology. In particular we follow GPV (2000) and indicate when some modifications are necessary. Our first two assumptions deal with the underlying data generating process and the smoothness of the latent joint distribution of $(V_{p\ell}, Z_\ell)$ for any $p = 1,\dots,I_\ell$.

Assumption A1:

(i) $Z_\ell = (X_\ell, I_\ell) \in I\!R^{d+1}$, $\ell = 1, 2, \dots$ *are independently and identically distributed as* $F_m(\cdot,\cdot)$

with density $f_m(\cdot, \cdot)$.

(ii) For each ℓ, $V_{p\ell}$, $p = 1, \ldots, I_\ell$ are independently and identically distributed conditionally upon Z_ℓ as $F(\cdot|\cdot; \theta_0, \gamma_0)$ with density $f(\cdot|\cdot; \theta_0, \gamma_0)$ where $\theta_0 \in I\!R^p$ and γ_0 is potentially infinite, finite or empty.

Let $\mathcal{I}$ be the set of possible values for I_ℓ. We denote by $S(*)$ the support of $*$ and by $S_i(*)$ the support when the number of bidders is equal to i.

Assumption A2: *$\mathcal{I}$ is a bounded subset of $\{2, 3, \ldots\}$, and:*

(i) for each $i \in \mathcal{I}$, $S_i(F) = \{(v, x) : x \in [\underline{x}, \overline{x}], v \in [\underline{v}(x), \overline{v}(x)]\}$, with $\underline{x} < \overline{x}$,

(ii) for $(v, x, i) \in S(F)$, $f(v|x, i; \theta_0, \gamma_0) \geq c_f > 0$, and for $(x, i) \in S(F_m)$, $f_m(x, i) \geq c_f > 0$,

(iii) for each $i \in \mathcal{I}$, $F(\cdot|\cdot, i; \theta_0, \gamma_0)$ and $f_m(\cdot, i)$ admit up to $R + 1$ continuous bounded partial derivatives on $S_i(F)$ and $S_i(F_m)$, with $R > d + 1$.

These assumptions can be found in GPV (2000) as well, though A2-(iii) is stronger in our case. That is, we need to require R to be sufficiently large with respect to the dimension of X, i.e. $R > d + 1$. This kind of condition is typically encountered in the semiparametric literature. Note that Proposition 1 in GPV (2000) still holds under our assumption A2.

The next two assumptions are on kernels and bandwidths that we use in the first stage.

Assumption A3:

(i) The kernels $K_G(\cdot)$, $K_{1g}(\cdot)$ and $K_{2g}(\cdot)$ are symmetric with bounded hypercube supports and twice continuous bounded derivatives with respect to their arguments,

(ii) $\int K_G(x)dx = 1$, $\int K_{1g}(x)dx = 1$, $\int K_{2g}(b)db = 1$

(iii) $K_G(\cdot)$, $K_{1g}(\cdot)$ and $K_{2g}(\cdot)$ are of order $R - 1$. Thus moments of order strictly smaller than $R - 1$ vanish.

This is similar to assumption A3 in GPV (2000). It is a standard assumption in the nonparametric literature.

Assumption A4: *The bandwidths h_G, h_{1g} and h_{2g} satisfy*

(i) $h_G \to 0$ *and* $\dfrac{Lh_G^d}{\log L} \to \infty$, *as* $L \to \infty$,

(ii) $h_{1g} \to 0$, $h_{2g} \to 0$ *and* $\dfrac{Lh_{1g}^d h_{2g}}{\log L} \to \infty$, *as* $L \to \infty$.

As shown below, for consistency of our estimator it is possible to choose the optimal bandwidths in the first step, i.e. the bandwidths proposed in Stone (1982). Unlike GPV (2000) we do not need to specify a "boundary bandwidth" since the local polynomial method does not require knowledge of the location of the endpoints of the support. Therefore, it is not necessary to estimate the boundary of the support of the bid distribution. This is necessary when one needs to trim out observations, which we do not given that our first step estimator is not subject to the so-called boundary effect.

For simplicity of presentation, in the rest of the paper we treat only univariate case for X, that is $d = 1$.

In order to describe our two–step estimator, we observe first that, analogously to the first step in GPV (2000), our objective is to estimate the ratio $\psi(\cdot|\cdot) = G_0(\cdot|\cdot)/g_0(\cdot|\cdot)$ by $\hat{\psi} = \hat{G}(\cdot|\cdot)/\hat{g}(\cdot|\cdot)$ (see (3) above). We use a local polynomial estimator (LPE) for each function. From Proposition 1 in GPV (2000) we know that $G_0(\cdot|\cdot)$ is $R+1$ times continuously differentiable on its entire support and therefore $g_0(\cdot|\cdot)$ is R times continuously differentiable on its entire support as well.[5] Given the smoothness of each function we propose to use a LPE(R), i.e. a LPE of degree R, for $G_0(\cdot|\cdot)$ and a LPE($R-1$) for $g_0(\cdot|\cdot)$. We introduce first some notation.

Let $P_\rho(X;\beta)$ denote a polynomial of degree ρ in X with parameter β. Then for each each i we

[5] Observe that by Proposition 1 in GPV (2000) we also know that the conditional density $g_0(\cdot|\cdot)$ is $R+1$ times continuously differentiable on a closed subset of the interior of the support and thus the degree of smoothness closed to the boundaries and at the boundaries of the support is not $R+1$.

have

$$\hat{G}(b|x) = \arg\min_{\beta_G} \sum_{\{\ell: I_\ell = i\}}^{L} \sum_{p=1}^{i} \left\{ Y_{p\ell}^{G} - P_R(X_\ell - x; \beta_G) \right\}^2 \frac{1}{h_G} K_G \left(\frac{X_\ell - x}{h_G} \right)$$

where $Y_{p\ell}^{G} = \mathbb{1}(B_{pl} \leq b)$, and

$$\hat{g}(b|x) = \arg\min_{\beta_g} \sum_{\{\ell: I_\ell = i\}}^{L} \sum_{p=1}^{i} \left\{ Y_{p\ell}^{g} - P_{R-1}(X_\ell - x; \beta_g) \right\}^2 \frac{1}{h_{1g}} K_{1g} \left(\frac{X_\ell - x}{h_{1g}} \right)$$

where $Y_{p\ell}^{g} = \frac{1}{h_{2g}} K_{2g} \left(\frac{B_{p\ell} - b}{h_{2g}} \right)$.

More precisely we have,

$$\begin{aligned}
\hat{G}(b|x,i) &= \frac{1}{h_G} \sum_{\{\ell: I_\ell = i\}}^{L} \sum_{p=1}^{i} e_1^T (X_{i,R+1}^T W_x^G X_{i,R+1})^{-1} \\
&\quad X_{R+1,\ell} K_G \left(\frac{X_\ell - x}{h_G} \right) \mathbb{1}(B_{p\ell} \leq b) \\
&= \frac{1}{L h_G} \frac{L}{n_i} \sum_{\{\ell: I_\ell = i\}}^{L} \sum_{p=1}^{i} e_1^T \left(\frac{X_{i,R+1}^T W_x^G X_{i,R+1}}{n_i} \right)^{-1} \\
&\quad X_{R+1,\ell} K_G \left(\frac{X_\ell - x}{h_G} \right) \mathbb{1}(B_{p\ell} \leq b)
\end{aligned} \tag{4}$$

and

$$\begin{aligned}
\hat{g}(b|x,i) &= \frac{1}{h_{1g} h_{2g}} \sum_{\{\ell: I_\ell = i\}}^{L} \sum_{p=1}^{i} e_1^T (X_{i,R}^T W_x^g X_{i,R})^{-1} X_{R,\ell} K_{1g} \left(\frac{X_\ell - x}{h_{1g}} \right) \\
&\quad K_{2g} \left(\frac{B_{p\ell} - b}{h_{2g}} \right) \\
&= \frac{1}{L h_{1g} h_{2g}} \frac{L}{n_i} \sum_{\{\ell: I_\ell = i\}}^{L} \sum_{p=1}^{i} e_1^T \left(\frac{X_{i,R}^T W_x^g X_{i,R}}{n_i} \right)^{-1} X_{R,\ell} K_{1g} \left(\frac{X_\ell - x}{h_{1g}} \right) \\
&\quad K_{2g} \left(\frac{B_{p\ell} - b}{h_{2g}} \right)
\end{aligned} \tag{5}$$

where for $s = \{R, R+1\}$,

e_1 is the unit vector in $I\!R^s$ containing a 1 in its first entry,

$n_i = iL_i$,

$L_i = \#\{\ell : I_\ell = i\}$,

$X_{s,\ell} = [1 \quad (X_\ell - x) \ldots (X_\ell - x)^{s-1}]^T$, is a $s \times 1$ vector and,

$$X_{i,s} = \begin{pmatrix} 1 & (X_1 - x) & \ldots & (X_1 - x)^{s-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (X_{n_i} - x) & \ldots & (X_{n_i} - x)^{s-1} \end{pmatrix}$$

is the matrix of regressors of dimension $n_i \times s$ with the first i rows identical and similarly for the other rows,

$W_x^G = \text{diag}\left\{\frac{1}{h_G} K_G\left(\frac{X_\ell - x}{h_G}\right)\right\}$,

$W_x^g = \text{diag}\left\{\frac{1}{h_{1g}} K_{1g}\left(\frac{X_\ell - x}{h_{1g}}\right)\right\}$,

where $K_G(\cdot)$, $K_{1g}(\cdot)$ and $K_{2g}(\cdot)$ are some kernels with bounded support and h_G, h_{1g},h_{2g} are some bandwidths (see A3 and A4). Given (4) and (5), a natural way of recovering pseudo private values is

$$\hat{V}_{p\ell} = B_{p\ell} + \frac{1}{I_\ell - 1}\hat{\psi}(B_{p\ell}|Z_\ell).$$

Unlike in GPV (2000), $\hat{\psi}$ is not subject to the so-called boundary effect, a typical problem encountered in kernel estimation. Thus, as mentioned above, we do not need to trim out observations that are "too close" to the boundary of the support of the joint distribution of $(B_{p\ell}, Z_\ell)$.

The second step of our estimation procedure is as follows. We propose to use the sample of pseudo private values in the following conditional moment restrictions, namely

$$\mathrm{E}[M(\hat{V}, Z; \theta_0)|Z] \approx 0,$$

for some known function $M(\cdot, \cdot; \theta) : I\!R^3 \to I\!R^q$ and $\theta \in I\!R^p$.

This set of conditional moment restrictions translates into the following set of unconditional moment restrictions,

$$\mathrm{E}[m(\hat{V}, Z; \theta_0)] \approx 0, \tag{6}$$

where $m(\cdot, \cdot; \theta) : I\!R^3 \to I\!R^q, \theta \in \Theta \subset I\!R^p$, with $q \geq p$.

In view of (6), we propose to estimate θ_0 by $\hat{\theta}$ as follows

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \hat{S}_L^T(\theta) \Omega \hat{S}_L(\theta), \tag{7}$$

where $\hat{S}_L(\theta) = 1/L \sum_{\ell=1}^{L} 1/I_\ell \sum_{p=1}^{I_\ell} m(\hat{V}_{p\ell}, Z_\ell; \theta)$ and Ω is a positive definite matrix of order q.

Ideally, one would like to specify the following set of conditional moment restrictions

$$\mathrm{E}[M(V, Z; \theta_0)|Z] = 0,$$

which would lead to the following unconditional moment restrictions

$$\mathrm{E}[m(V, Z; \theta_0)] = 0.$$

Therefore, the infeasible estimator $\tilde{\theta}$, (say), is the solution to

$$\tilde{\theta} = \arg\min_{\theta \in \Theta} S_L^T(\theta) \Omega S_L(\theta),$$

where $S_L(\theta) = 1/L \sum_{\ell=1}^{L} 1/I_\ell \sum_{p=1}^{I_\ell} m(V_{p\ell}, Z_\ell; \theta)$.

As we show in the Appendix, the asymptotic distributions of $\hat{\theta}$ and $\tilde{\theta}$ are closely related though not the same.

3 Asymptotic Properties

In this section we show that our two-step semiparametric estimator $\hat{\theta}$ of θ_0 is consistent and asymptotically normal distributed. Moreover, we establish that our estimator attains the parametric uniform rate of convergence given an appropriate choice of the bandwidths used in the first step to estimate $G_0(\cdot|\cdot)$ and $g_0(\cdot|\cdot)$. As we will discuss below the optimal bandwidths, given by Stone (1982), i.e. the one-step bandwidths, cannot be chosen, instead our choice implies that in practice one needs to undersmooth.

We also discuss the assumptions under which our results hold.

3.1 Consistency

Our first result establishes that $\hat{\theta}$ is a (strongly) consistent estimator for θ_0. Moreover this is the case even if one uses the optimal bandwidths for estimating $G_0(\cdot|\cdot)$ and $g_0(\cdot|\cdot)$ in the first step, i.e. the bandwidths proposed by Stone (1982). To see this, we notice that the "optimal one-step" bandwidths satisfy our assumption A4 above (with $d = 1$) since they are of the form,

(i) $h_G = \lambda_G \left(\frac{\log L}{L}\right)^{1/(2R+3)}$

(ii) $h_{1g} = \lambda_{1g} \left(\frac{\log L}{L}\right)^{1/(2R+1)}$ *and* $h_{2g} = \lambda_{2g} \left(\frac{\log L}{L}\right)^{1/(2R+1)}$,

where λ_G, λ_{1g} *and* λ_{2g} *are strictly positive constants.*

As observed by GPV (2000), h_G, h_{1g} and h_{2g}, as given above are optimal bandwidth choices to estimate $G_0(\cdot|\cdot)$ and $g_0(\cdot|\cdot)$ given Proposition 1 and A2-(iii) in that paper.[6] Thus, A4 implies that our consistency result can be established when using LPE in the first stage that converge at the best possible rate. The remaining assumptions are standard in the literature.

Assumption A5:

(i) The parameter space $\Theta \subset I\!R^p$ *is compact and* θ_0 *is in the interior of* Θ*,*

[6] As pointed out before, A2-(iii) in our case is stronger than A2-(iii) in GPV (2000). Thus their Proposition 1 also holds in our framework.

(ii) Identifying assumption: $\mathrm{E}[m(V,Z;\theta)] = 0$ *if and only if* $\theta = \theta_0$,

(iii) $\sup_{\theta\in\Theta} \left| \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}\|m(V_{p\ell}, Z_\ell;\theta)\| - \mathrm{E}\,\|m(V,Z;\theta)\| \right| = o_{as}(1)$,

(iv) $m(V,Z;\theta)$ *is Lipschitz in* V*: there exists a measurable function* $K_1(Z)$ *such that*

$$\|m(V,Z;\theta) - m(V',Z;\theta)\| \leq K_1(Z)\,|V - V'|\,,$$

for every V, $V' \in [\underline{V}, \overline{V}]$ *and every* $\theta \in \Theta$*. Moreover* $\mathrm{E}[K_1(Z)] < \infty$.

We now state our consistency result.

Proposition 1: *Let* $\hat{\theta}$ *be defined as in (7). Then, under A1-A5, we have*

$$\hat{\theta} \xrightarrow{a.s} \theta_0.$$

Proposition 1 is important since it establishes that our estimator possess one of the desirable asymptotic properties. This is the first step in order to be able to establish the asymptotic distribution of the estimator. Moreover, there is no need to undersmooth the distribution and density functions in the first step in order for $\hat{\theta}$ to be consistent.

3.2 Asymptotic Normality

Given that $\hat{\theta}$ is a (strongly) consistent estimator for θ_0, we can now establish its asymptotic distribution along with its uniform convergence rate. This is the purpose of Proposition 2 below. Before, we introduce some additional assumptions.

We need to modify our choice of bandwidths as mentioned earlier. As it is well known, among the typical properties encountered in semiparametric estimation, one usually finds that the optimal bandwidth choice is excluded and in particular one needs to undersmooth.[7] Our semiparametric

[7] Another typical property usually encountered has to do with a sufficiently large degree of smoothness relative to the dimension of the exogenous variables, as reflected by A2-(iii).

estimator is not an exception in this respect. Thus, as it is made clear by A4.AN below, in order for $\hat{\theta}$ to achieve the parametric uniform rate of convergence we need to specify bandwidths for our first step that rule out the optimal choice and moreover that imply undersmoothed estimates for $\hat{G}(\cdot|\cdot)$ and $\hat{g}(\cdot|\cdot)$.

Assumption A4.AN: *The bandwidths h_G, h_{1g} and h_{2g} satisfy*

(i) $\sqrt{L}h_G^{R+1} \to 0$ *and* $\dfrac{\log L}{\sqrt{L}h_G} \to 0,$ *as* $L \to \infty$

(ii) $\sqrt{L}h_{1g}^R \to 0,$ $\sqrt{L}h_{2g}^R \to 0$ *and* $\dfrac{\log L}{\sqrt{L}h_{1g}h_{2g}} \to 0,$ *as* $L \to \infty.$

(iii) $h_{1g} = h_{2g}$

The assumption that h_{1g} and h_{2g} vanish at the same rate, is to simplify the notation in the proof. In fact it is enough to choose any pair of bandwidths strictly smaller than their optimal counterparts.

Let $m_k(\cdot)$ be the partial derivative of $m(\cdot)$ with respect to the k^{th} argument. The next set of assumptions is standard.

Assumption A6:

(i) $m_3(V, Z; \theta)$ *is Lipschitz in* V*: there exists a measurable function* $K_3(Z)$ *such that*

$$\|m_3(V, Z; \theta) - m_3(V', Z; \theta)\| \leq K_3(Z)|V - V'|,$$

for every V*,* $V' \in [\underline{V}, \overline{V}]$ *and* $\theta \in \Theta$*. Moreover* $\mathrm{E}[K_3(Z)] < \infty$*,*

(ii) $m_3(V, Z; \theta)$ *is Lipschitz in* θ*: there exists a measurable function* $K_4(Z)$ *such that*

$$\|m_3(V, Z; \theta) - m_3(V, Z; \theta')\| \leq K_4(Z)\left\|\theta - \theta'\right\|,$$

for every θ*,* $\theta' \in \Theta$ *and* $V \in [\underline{V}, \overline{V}]$*. Moreover* $\mathrm{E}[K_4(Z)] < \infty$*,* *(iii)* $\sup_{\theta\in\Theta}\left\|\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}m_3(V_{p\ell},Z_\ell;\theta)-\mathrm{E}[m_3(V,Z;\theta)]\right\| = o_{as}(1)$ *and*

$\mathrm{E}[m_3'(V,Z;\theta)]\Omega\mathrm{E}[m_3(V,Z;\theta)]$ *is non singular,*

(iv) $\sup_{\theta\in\Theta}\|m_3(V,Z;\theta)\| \leq K_5(V,Z)$ *with* $\mathrm{E}[K_5(V,Z)] < \infty$,

(v) $m_1(V,Z;\theta)$ *is Lipschitz in* V*: there exists a measurable function* $K_6(Z)$ *such that*

$$\|m_1(V,Z;\theta) - m_1(V',Z;\theta)\| \leq K_6(Z)|V - V'|,$$

for every V, $V' \in [\underline{V},\overline{V}]$ *and* $\theta\in\Theta$. *Moreover* $\mathrm{E}[K_6(Z)] < \infty$,

(vi) $\sup_{\theta\in\Theta}\|m_1(V,Z;\theta)\| \leq K_7(V,Z)$ *with* $\mathrm{E}[K_7(V,Z)^2] < \infty$.

Our last assumption concerns a technical condition we need for the proof of Proposition 2. Our problem brings some technical difficulties that are not usually found in the semiparametric literature. As a consequence, we need technical devices to get around some of these problems. This is the purpose of Assumption A7.[8]

Assumption A7:

(i) Technical condition: $\mathrm{E}[m_1(V,Z;\theta_0)] = 0$.

To see what it entails, consider the following simple example. Let $Z = X \in I\!R$ and $\mathrm{E}(V|Z) = Z\theta_0$. Then $\mathrm{E}(V - Z\theta_0|Z) = 0$, and the unconditional moment restriction becomes $\mathrm{E}(Z(V - Z\theta_0)) = 0$. Then the assumption satisfied whenever $\mathrm{E}(Z) = 0$. Next, we establish our main result.

Proposition 2: *Let* $\hat\theta$ *be defined as in (7). Then, under A1-A3, A4.AN and A5-A7, we have*

$$\sqrt{L}(\hat\theta - \theta_0) \xrightarrow{d} N(0,\Sigma).$$

[8]Assumptions A7 basically implies that we can apply a Taylor expansion around h at $h = 0$ to a bias term we need to deal with.

where, for each $i \in \mathcal{I}$

$$\Sigma = \mathrm{Var}(\psi_1),$$

$$\begin{aligned}\psi_1 &= -1/i\sum_{p=1}^{i}\Bigg\{(C^T\Omega C)^{-1}C\Omega m(V_{p1},X_1,i;\theta_0)+2\Bigg[\sum_i \frac{1}{i(i-1)}N(Y_{p1},i)\\ &\quad f_m^{-1}(X_1,i)g_0(Y_{p1},i)-\mathrm{E}\Big[\sum_i \frac{1}{i(i-1)}N(Y_{p1},i)f_m^{-1}(X_1,i)g_0(Y_{p1},i)\Big]\Bigg]\Bigg\},\end{aligned}$$

$$C = E[\partial m(V,X,i;\theta_0)/\partial\theta],$$

and

$$N(Y_{p1},i) = [m_1(V_{p1},X_1,i;\theta_0)/g_0(B_{p1}|X_1,i)^2]G_0(B_{p1}|X_1,i).$$

Proposition 2 is important for several reasons. First it establishes that our semiparametric estimator has a standard limiting distribution. Asymptotic Normality is fundamental since most of the econometric tests rely on it. Second, although slow estimators are used in the first step of our estimation procedure to recover pseudo private values, the estimator of the parameter of interest converges at the best possible rate. Third, our semiparametric estimator is not subject to the curse of dimensionality. Finally, Proposition 2 can be used to conduct inference on θ_0.

3.3 Monte Carlo Experiments

In this section we present the results of a set of Monte Carlo simulations in order to asses the performance of our semiparametric estimator relative to the nonparametric estimator proposed in GPV (2000). In line with the theoretical framework analyzed in the previous section we consider a setup with observed object heterogeneity ($d = 1$). We use $L = 200$ with $I = 5$ bidders, which gives a total of 1000 observed bids. The choice $L = 200$ corresponds to realistic size of auction data and makes the comparison with GPV (2000) easier since these authors also use this sample size although they consider homogeneous auctions. In order to account for object heterogeneity, we generate X from a log-normal distribution with mean 0 and variance 1 truncated at 0.055 and 30 to satisfy A2-(i). The true distribution $F(\cdot|Z;\theta_0,\gamma_0)$ of private values is also lognormal. Moreover we have that conditional on X, private values are log-normally distributed with mean $1+X$ and variance 1. We consider $\theta_0=(1,1)^T$ and $\gamma_0=\emptyset$. To satisfy assumption A2-(ii) we truncate this distribution at 0.055 and 30. We consider that this function has 4 continuous bounded partial derivatives, so that $R=3$. In line with assumption A3, we choose the triweight kernel $(35/32)(1-u^2)^3 1\!\!1(|u|\leq 1)$ for the three kernels involved in our first step estimators.

We choose the bandwidths according to A4.AN. In particular we use $h_G=1.06\hat{\sigma}_x(IL)^{-1/6.5}$, $h_{1g}=1.06\hat{\sigma}_x(IL)^{-1/4.5}$, $h_{2g}=1.06\hat{\sigma}_b(IL)^{-1/4.5}$,where $\hat{\sigma}_b$ and $\hat{\sigma}_x$ are the estimated standard deviations of observed bids and object heterogeneity, respectively. The factor 1.06 follows from the so-called rule of thumb (see Hardle (1991)). The use of I arises because we have I bidders per auction.[9]

We use 1000 replications. More precisely, for each replication we first generate randomly IL private values using the truncated log–normal distribution. Next we compute the corresponding bids $B_{p\ell}$ using (2). With these observed bids we can now apply our estimation procedure for each replication. First, we estimate the distribution and density functions of observed bids using (4) and (5).

Next, we compute pseudo private values $\hat{V}_{p\ell}$ corresponding to $B_{p\ell}$ as

$$\hat{V}_{p\ell}=B_{p\ell}+\frac{1}{I-1}\hat{G}(B_{p\ell}|Z_\ell)/\hat{g}(B_{p\ell}|Z_\ell)$$

for $p=1,\ldots,5$ and $\ell=1,\ldots,L$.

[9]To replicate the GPV (2000) estimator we choose the bandwidths according to the optimal rates. Thus, the order of the bandwidths is $L^{-1/9}$ for h_G and the second step bandwidth h_x and $L^{-1/10}$ for h_{gb} and h_{gx} and the second step bandwidths h_{fv} and h_{fx}. Specifically we use $h_G=1.06\hat{\sigma}_x(IL)^{-1/9}$, $h_{gx}=1.06\hat{\sigma}_x(IL)^{-1/10}$, $h_{gb}=1.06\hat{\sigma}_b(IL)^{-1/10}$ where $\hat{\sigma}_b$ and $\hat{\sigma}_x$ are as defined above. The second step bandwidths are $h_{fv}=1.06\hat{\sigma}_{\hat{v}}(n_t)^{-1/10}$,$h_{fx}=1.06\hat{\sigma}_x(n_t)^{-1/10}$ and $h_x=1.06\hat{\sigma}_x(L)^{-1/9}$, where n_t is the number of observations remaining after trimming.

In a second step we formulate the following (infeasible) moment condition

$$\mathrm{E}\left[1\!\!I(\underline{V} \leq V \leq \overline{V}) \frac{\partial \ln f(V|I,X;\theta_0)}{\partial \theta}\right] = 0.$$

Our feasible moment condition then becomes[10]

$$\mathrm{E}\left[1\!\!I(\underline{V} \leq \hat{V} \leq \overline{V}) \frac{\partial \ln f(\hat{V}|I,X;\theta_0)}{\partial \theta}\right] \approx 0.$$

Thus our sample moment condition can be expressed as follows

$$\frac{1}{IL}\sum_{\ell=1}^{L}\sum_{p=1}^{I} 1\!\!I(0.055 \leq \hat{V}_{p\ell} \leq 30) \frac{\partial \ln f(\hat{V}_{p\ell}|I,X;\hat{\theta})}{\partial \theta} = 0.$$

Using the above condition we implement an efficient GMM procedure to obtain $\hat{\theta}$. We represent our results for a fix value of X. In particular, we set X equal to its median. Figure 1 below shows the true density of private values against the two estimators we are comparing. The semiparametric estimator developed in this paper (dashed line) does a good job in matching the true density. Moreover, when comparing with the nonparametric GPV (2000) estimator, we can see that our semiparametric estimator is not subject to boundary effects.[11]

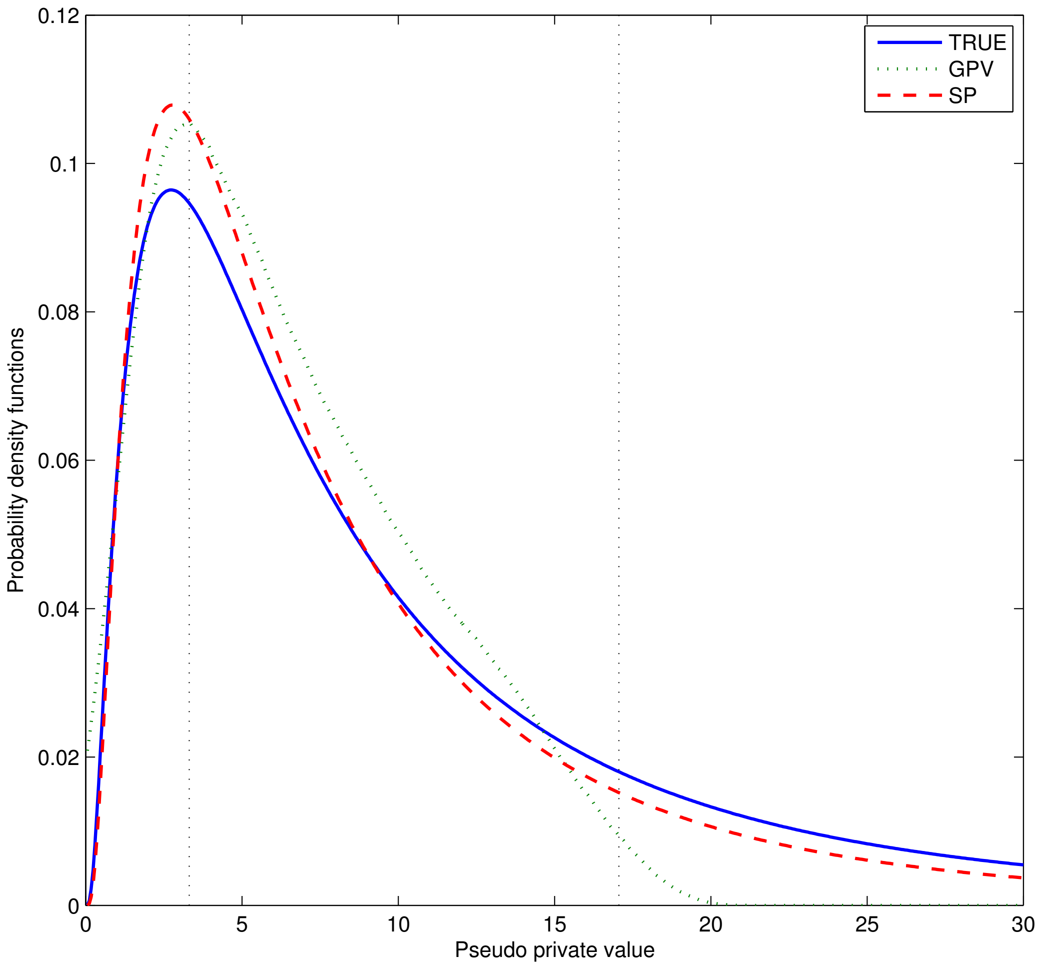
4 A More General Class of Models

In this Section we indicate how to extend our procedure to a more general class of auction models. To keep the notation as simple as possible we consider models without observed object heterogeneity. This is not restrictive since relaxing this assumption implies that the distribution and density functions have to be replaced by their conditional counterparts.

[10] This moment condition does not satisfy Assumption **A7**-(i), but we still use it because it serves two purposes. First, and as mentioned earlier, we are trying to use the same exercise as GPV(2000). Second, since even then the estimator performs well (see below) it shows that the assumption is a strong sufficient condition, and the estimator can perform well even when the condition is violated.

[11] The vertical lines in the figure correspond to the trimming we have conducted to which one h_f is added (and subtracted) to eliminate remaining boundary effects.

Figure 1: Densities of bidders' private values



Binding Reserve Price

The first natural extension of the model considered in Section 2 is the symmetric IPV first-price auction model with a binding reserve price, announced or random.

Announced Reserve Price

An announced binding reserve price ($p_0 > \underline{V}$) constitutes a screening device for participating in the auction. As pointed out by GPV (2000) the Bayesian Nash equilibrium strategy is still given by

(2) in this set up, but the number I of potential bidders becomes unobserved and typically different from the observed number, I^*, of actual bidders who have submitted a bid ($\geq p_0$). Hence the model has a new structural element, namely I, in addition to the latent distribution of bidders' private values. As shown in GPV (2000), the differential equation defining the equilibrium strategy can be rewritten as

$$V_i = \xi_0(B_i, G_0^*, F(p_0), I) = B_i + \frac{1}{I-1}\left(\frac{G_0^*(B_i)}{g_0^*(B_i)} + \frac{F(p_0)}{1-F(p_0)}\frac{1}{g_0^*(B_i)}\right),$$

for $i = 1, \ldots, I^*$ and where $G_0^*(\cdot)$ is the truncated distribution of an observed bid conditional upon the fact that the corresponding private value is grater than or equal to p_0. Provided one can estimate I and $F(p_0)$ this equation is the basis for a two step procedure analogous to that of Section 2.[12]

Random Reserve Price

In some cases, as in timber and wine auctions, the seller may decide not to announce the reserve price at the time the auction takes place. Hence, the reserve price is said to be secret or random. Since bidders do not know it when submitting their bids, this fact brings into the model a new kind of uncertainty that has to be taken into account. To present the basic equation underlying our two-step procedure in this model we need first to introduce additional notation. Let V_0 be the private value of the risk-neutral seller for the auctioned object. Moreover, we assume that V_0 is distributed according to $H(\cdot)$ defined on the same support as $F(\cdot)$ and that $H(\cdot)$ is common knowledge. Elyakime, Laffont, Loisel and Vuong (1994) have shown that in a first-price sealed bid auction $p_0 = V_0$. In addition the bidders' equilibrium strategy is the solution of a differential equation which in general cannot be solved explicitely. See Li and Perrigne (2003). However, this differential equation can be rewritten as follows

$$V_i = \xi_0(B_i, H, G_0, I) = B_i + \frac{1}{(I-1)\left(\frac{g_0(B_i)}{G_0(B_i)} + \frac{h(B_i)}{H(B_i)}\right)},$$

[12]Introduction of heterogeneity across auctioned objects can be easily implemented if $p_{0\ell}$ is an unknown deterministic function of exogenous variables. See GPV (2000), Section 4 for further details.

for $i = 1, \ldots, I$. As mentioned by Perrigne and Vuong (1999) since the reserve price is kept secret, all potential bidders submit a bid. Hence I is typically observed. The above equation can be used as the basis of a two-step procedure similar to the one described in Section 2. Namely, in a first step observed bids and reserve prices can be used to estimate nonparametrically the distribution $G_0(\cdot)$, its density $g_0(\cdot)$ as well as the distribution $H(\cdot)$ and its density $h(\cdot)$. Next, pseudo private values can be recovered using the equation above in order to define a set of moment conditions for estimating the parameter of interest θ_0 in a second step.

The Symmetric Affiliated Private Value (APV) Model

To assume independence across private value can be restrictive since one can expect some degree of affiliation or positive correlation among private values. Thus, a second natural extension of our framework is to consider the more general class of model encompassed by symmetric APV models. Affiliation means that if one bidder draws a high valuation for the auctioned object, then others bidders are likely to draw higher valuations too. Laffont and Vuong (1996) study the problem of identification and theoretical restrictions in a general framework, namely in Affiliated Value (AV) models. In particular they show that any symmetric AV model is observationally equivalent to some symmetric APV model because the utility function is not identified from observed bids only.[13] Therefore when only data on observed bids are available the result in Laffont and Vuong (1996) implies that APV models can be considered without loss of generality, provided that we have identification.

We briefly indicate here how to adapt our estimation procedure to this kind of models. We assume that all bids are observed and that the reserve price is nonbinding. Let $Y_i = \max_{j \neq i} V_j$. The differential equation defining the equilibrium strategy in the APV model can be written as follows

$$V_i = \xi_0(B_i, G_0) \equiv B_i + \frac{G_{0,b_1|B_1}(B_i|B_i)}{g_{0,b_1|B_1}(B_i|B_i)},$$

[13]Two auction models are said to be observationally equivalent given observed bids, if they lead to the same equilibrium bids distribution.

for all $V_i \in [\underline{V}, \overline{V}]$ subject to the boundary condition $s(\underline{V}) = \underline{V}$, where $G_{0,b_1|B_1}(x_1|X_1) = F_{Y_1|V_1}(s^{-1}(x_1))$, $B_1 = s(Y_1)$. This equation is again the basis for the identification result and estimation procedure.

The theoretical restrictions as shown by Li, Perrigne and Vuong (1999) indicate that the joint distribution of bids $G_0(\cdot)$ can be rationalized by a symmetric APV model if and only if (i) $G_0(\cdot)$ is symmetric and affiliated and (ii) the function $\xi_0(\cdot, G_0)$ is strictly increasing on its support. Moreover, if these two conditions are satisfied, then the joint distribution $F(\cdot)$ of private values is identified.

Regarding estimation, the equation above suggests a two-step procedure analogous to the one described in Section 2. In the first step the ratio $G_{b_1|B_1}(\cdot|\cdot)/g_{b_1|B_1}(\cdot|\cdot)$ can be estimated nonparametrically and then pseudo private values can be recovered. In the second step a GMM procedure can be implemented to estimate the parameter of interest of the underlying distribution of private values.

Asymmetric Models

Assuming that bidders are ex ante identical may constitute a limitation. Therefore, in some cases one needs to use models relaxing this assumption. However, a common feature share by asymmetric auction models is that they lead to systems of differential equations without a closed form solution. Hence, the direct approach becomes extremely difficult to implement. Nevertheless, using our indirect two-step procedure, asymmetric models can be structurally estimated while avoiding solving for the equilibrium strategy as well as of its inverse.

The Asymmetric IPV Model

Following the exposition in Perrigne and Vuong (2008) we assume that asymmetry is ex ante known to all bidders. Let $F_1(\cdot), \ldots, F_I(\cdot)$ be the private value distributions of the I bidders whose identities are observed and let $G_1(\cdot), \ldots, G_I(\cdot)$ be the corresponding bid distributions.[14] We can express the intractable system of differential equations as follows

$$V_i = B_i + \frac{1}{\sum_{j \neq i} \frac{g_j(B_i)}{G_j(B_i)}}$$

[14]In the context of procurements, Flambard and Perrigne (2007) use an asymmetric model to analyze snow removal contracts in Canada.

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

The above system of equations leads naturally to a two-step procedure similar to the one proposed in Section 2.

The Asymmetric APV Model

For simplicity we consider only two types of bidders. That is, the model assumes that the I-dimensional vector $(V_{11}, \ldots, V_{1I_1}, V_{01}, \ldots, V_{0I_0})$ is distributed as $F(\cdot)$ which is exchangeable in its first I_1 and last I_0 arguments. We can interprete this structure as follows. There is symmetry within each subgroup as bidders of the same type are assumed to be ex ante identical. Even more, since $F(\cdot)$ is affiliated, there is general positive dependence among private values. From Campo, Perrigne and Vuong (1998) the system of differential equation defining equilibrium strategies is

$$V_1 = \xi_1(B_1, G) \equiv B_1 + \frac{G_{B_1^*, B_0|B_1}(B_1, B_1|B_1)}{\partial G_{B_1^*, B_0|B_1}(B_1, B_1|B_1)/\partial T},$$

$$V_0 = \xi_0(B_0, G) \equiv B_0 + \frac{G_{B_1, B_0^*|B_0}(B_0, B_0|B_0)}{\partial G_{B_1, B_0^*|B_0}(B_0, B_0|B_0)/\partial T},$$

where $B_j^* = \max_{i \neq 1, i \in G_j} B_{ji}$, $B_j = \max_{i \in G_j} B_{ji}$, for $j = 1, 0$ and the partial derivatives with respect to T indicate the total derivative with respect to the first two arguments. Campo, Perrigne and Vuong (1998) establish that $F(\cdot, \ldots, \cdot)$ is identified. Moreover they use a nonparametric two-step procedure following GPV (2000) to estimate the model. We propose instead to use our two-step semiparametric procedure using the above system of equations to recover pseudo private values after obtaining nonparametric estimates for $G_{B_1^*, B_0|B_1}(\cdot, \cdot|\cdot)$ and $G_{B_1, B_0^*|B_0}(\cdot, \cdot|\cdot)$. Next in a second step a model for the private values is specified through a set of moment conditions.

5 Conclusions

In this paper we develop an indirect procedure to estimate first-price sealed-bid auction models, contributing in this way to the structural analysis of auction data that has been developed in the last fifteen years. Following GPV (2000) our procedure is in two steps. The difference with GPV (2000) is that our second step is implemented using a GMM procedure so that our resulting model is semiparametric. We show that our semiparametric estimator converges uniformly at the parametric $\sqrt{L}$ rate while the nonparametric estimator in GPV (2000) was shown to converge at the best possible rate according to the minimax theory which is slower than the parametric rate. Moreover, our procedure is not subject to the so-called curse of dimensionality or in other words the convergence rate is independent of the dimension of the exogenous variables. We establish consistency and asymptotic normality of our estimator.

Given the nature of our procedure it is not necessary to solve explicitly for the equilibrium strategy or its inverse. This is a valuable advantage with respect to direct methods specially when estimating models that lead to intractable first-order conditions, such as asymmetric auction models. More generally, our method extends to models which have been estimated using a nonparametric indirect procedure. In this respect, we briefly outline how this can be done in models with a binding reserve price (announced or random), affiliated private value models and asymmetric models.

Finally, we conducted a set of Monte Carlo simulations. The main purpose for this was to asses the performance of our estimator in finite samples relative to the nonparametric estimator proposed by GPV (2000). Our semiparametric estimator does a good job in matching the true density. Moreover, when comparing with the nonparametric GPV (2000) estimator, we can see that the estimator developed in this paper is not subject to boundary effects.

Appendix

Proofs of Asymptotic Properties

This Appendix gives the proofs of our asymptotic results (Propositions 1 and 2).

We state first two important results.

Results: Under A4 we have,

(i) $\sup_{(b,x,i)} |\hat{g}(b|x,i) - g_0(b|x,i)| = O_{as}\left(h_{1g}^R + h_{2g}^R + \sqrt{\frac{\log L}{Lh_{1g}h_{2g}}}\right)$

(ii) $\sup_{(b,x,i)} \left|\hat{G}(b|x,i) - G_0(b|x,i)\right| = O_{as}\left(h_G^{R+1} + \sqrt{\frac{\log L}{Lh_G}}\right)$

For a proof of the above results we refer the reader to Korostelev and Tsybakov (1993).

We observe that the above results imply that $\sup_{p\ell}|\hat{V}_{p\ell} - V_{p\ell}| = o_{as}(1)$.

Proof of Proposition 1: It suffices to show that $\sup_{\theta\in\Theta} \| S_L(\theta) - \hat{S}_L(\theta) \| = o_{as}(1)$. From the triangle inequality, A5-(iv) it follows that

$$\begin{aligned}
\sup_{\theta\in\Theta}\|S_L(\theta) - \hat{S}_L(\theta)\| &= \sup_{\theta\in\Theta}\left\|\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}m(V_{p\ell}, Z_\ell;\theta) - \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}m(\hat{V}_{p\ell}, Z_\ell;\theta)\right\| \\
&= \sup_{\theta\in\Theta}\left\|\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}[m(V_{p\ell}, Z_\ell;\theta) - m(\hat{V}_{p\ell}, Z_\ell;\theta)]\right\| \\
&\leq \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}\sup_{\theta\in\Theta}\left\|m(V_{p\ell}, Z_\ell;\theta) - m(\hat{V}_{p\ell}, Z_\ell;\theta)\right\| \\
&\leq \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}K_1(Z_\ell)|\hat{V}_{p\ell} - V_{p\ell}| \\
&= \{\mathrm{E}[K_1(Z)] + o_{as}(1)\}\sup_{p\ell}|\hat{V}_{p\ell} - V_{p\ell}| \\
&= o_{as}(1) \qquad\qquad (\text{A.1})
\end{aligned}$$

Where we use the fact that $\hat{V}_{p\ell}$ is a consistent estimator of $V_{p\ell}$, i.e. we make use of the 2 results stated at the beginning of this Appendix. Therefore, the desired result follows. *Q.E.D.*

Proof of Proposition 2

From the FOCs that characterize $\tilde{\theta}$ and $\hat{\theta}$ respectively, we have

$$\frac{1}{2}\frac{\partial Q_L}{\partial \theta}(\tilde{\theta}) = \frac{\partial S_L^T}{\partial \theta}(\tilde{\theta})\Omega S_L(\tilde{\theta}) = 0 \tag{A.2}$$

$$\frac{1}{2}\frac{\partial \hat{Q}_L}{\partial \theta}(\hat{\theta}) = \frac{\partial \hat{S}_L^T}{\partial \theta}(\hat{\theta})\Omega \hat{S}_L(\hat{\theta}) = 0. \tag{A.3}$$

We can use a Taylor expansion around θ_0 to obtain

$$S_L(\tilde{\theta}) = S_L(\theta_0) + \frac{\partial S_L}{\partial \theta^T}(\overline{\theta})(\tilde{\theta} - \theta_0) \tag{A.4}$$

$$\hat{S}_L(\hat{\theta}) = \hat{S}_L(\theta_0) + \frac{\partial \hat{S}_L}{\partial \theta^T}(\overline{\theta}^*)(\hat{\theta} - \theta_0), \tag{A.5}$$

where $\overline{\theta}$ and $\overline{\theta}^*$ are vectors between $\tilde{\theta}$ and θ_0, and $\hat{\theta}$ and θ_0, respectively.

Thus using (A.4) in (A.2) we get

$$\begin{aligned}
&\frac{\partial S_L^T}{\partial \theta}(\tilde{\theta})\Omega\left[S_L(\theta_0) + \frac{\partial S_L}{\partial \theta^T}(\overline{\theta})(\tilde{\theta} - \theta_0)\right] \\
&= \frac{\partial S_L^T}{\partial \theta}(\tilde{\theta})\Omega S_L(\theta_0) + \frac{\partial S_L^T}{\partial \theta}(\tilde{\theta})\Omega\frac{\partial S_L}{\partial \theta^T}(\overline{\theta})(\tilde{\theta} - \theta_0) = 0.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\sqrt{L}(\tilde{\theta} - \theta_0) &= -\left[\frac{\partial S_L^T}{\partial \theta}(\tilde{\theta})\Omega\frac{\partial S_L}{\partial \theta^T}(\overline{\theta})\right]^{-1}\frac{\partial S_L^T}{\partial \theta}(\tilde{\theta})\Omega\sqrt{L}S_L(\theta_0) \\
&= -\tilde{A}^{-1}\tilde{B}\sqrt{L}S_L(\theta_0).
\end{aligned}$$

Similarly using (A.5) in (A.3) yields

$$\begin{aligned}
\sqrt{L}(\hat{\theta} - \theta_0) &= -\left[\frac{\partial \hat{S}_L^T}{\partial \theta}(\hat{\theta})\Omega\frac{\partial \hat{S}_L}{\partial \theta^T}(\overline{\theta}^*)\right]^{-1}\frac{\partial \hat{S}_L^T}{\partial \theta}(\hat{\theta})\Omega\sqrt{L}\hat{S}_L(\theta_0) \\
&= -\hat{A}^{-1}\hat{B}\sqrt{L}\hat{S}_L(\theta_0).
\end{aligned}$$

we need to show: (i) $\tilde{B} - \hat{B} = o_{as}(1)$, (ii) $\tilde{A} - \hat{A} = o_{as}(1)$ since this and A6-(iii) imply that the difference of the inverses is also $o_{as}(1)$ and (iii) $\sqrt{L}[S_L(\theta_0) - \hat{S}_L(\theta_0)] = O_p(1)$.

The proof consists of three steps. Step 1: We prove $\tilde{B} - \hat{B} = o_{as}(1)$. The term $\tilde{B} - \hat{B}$ can be written as

$$\begin{aligned}
\tilde{B} - \hat{B} &= \frac{\partial S_L^T}{\partial \theta}(\tilde{\theta})\Omega - \frac{\partial \hat{S}_L^T}{\partial \theta}(\hat{\theta})\Omega \\
&= \left(\frac{\partial S_L^T}{\partial \theta}(\tilde{\theta}) - \frac{\partial \hat{S}_L^T}{\partial \theta}(\hat{\theta}) \right) \Omega \\
&= \left(\frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{I_\ell} \sum_{p=1}^{I_\ell} \Big(m_3^T(V_{p\ell}, Z_\ell, \tilde{\theta}) - m_3^T(\hat{V}_{p\ell}, Z_\ell; \hat{\theta}) \Big) \right) \Omega.
\end{aligned}$$

It suffices to show that the norm of the term between brackets is $o_{as}(1)$ since Ω is a positive definite matrix. Namely

$$\begin{aligned}
&\left\| \frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{I_\ell} \sum_{p=1}^{I_\ell} m_3^T(V_{p\ell}, Z_\ell, \tilde{\theta}) - m_3^T(\hat{V}_{p\ell}, Z_\ell; \hat{\theta})) \right\| \\
&= \left\| \frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{I_\ell} \sum_{p=1}^{I_\ell} \Big[(m_3^T(V_{p\ell}, Z_\ell, \tilde{\theta}) - m_3^T(\hat{V}_{p\ell}, Z_\ell; \tilde{\theta})) + m_3^T(\hat{V}_{p\ell}, Z_\ell; \tilde{\theta}) - m_3^T(\hat{V}_{p\ell}, Z_\ell; \hat{\theta})) \Big] \right\| \\
&\le \left\| \frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{I_\ell} \sum_{p=1}^{I_\ell} \Big[(m_3^T(V_{p\ell}, Z_\ell; \tilde{\theta}) - m_3^T(\hat{V}_{p\ell}, Z_\ell; \tilde{\theta})) \Big] \right\| \\
&\quad + \left\| \frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{I_\ell} \sum_{p=1}^{I_\ell} m_3^T(\hat{V}_{p\ell}, Z_\ell; \tilde{\theta}) - m_3^T(\hat{V}_{p\ell}, Z_\ell; \hat{\theta})) \right\| \\
&= C + D,
\end{aligned} \tag{A.6}$$

where the last line follows from the triangle inequality. The term C is

$$\begin{aligned}
C &= \left\| \frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{I_\ell} \sum_{p=1}^{I_\ell} \Big[(m_3^T(V_{p\ell}, Z_\ell; \tilde{\theta}) - m_3^T(\hat{V}_{p\ell}, Z_\ell; \tilde{\theta})) \Big] \right\| \\
&\le \frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{I_\ell} \sum_{p=1}^{I_\ell} \left\| \Big[m_3^T(V_{p\ell}, Z_\ell; \tilde{\theta}) - m_3^T(\hat{V}_{p\ell}, Z_\ell; \tilde{\theta}) \Big] \right\| \\
&\le \frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{I_\ell} \sum_{p=1}^{I_\ell} K_3(Z_\ell) |V_{p\ell} - \hat{V}_{p\ell}| \\
&\le \{ \mathrm{E}[K_3(Z)] + o_{as}(1) \} \sup_{p\ell} |\hat{V}_{p\ell} - V_{p\ell}| \\
&= o_{as}(1)
\end{aligned}$$

where we use A6-(i) and the fact that $\hat{V}_{p\ell}$ is uniformly consistent, i.e., we use the two results stated at the beginning of this Appendix.

We consider now the term D

$$
\begin{aligned}
D &= \left\| \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}\left[m_3^T(\hat{V}_{p\ell}, Z_\ell;\tilde{\theta}) - m_3^T(\hat{V}_{p\ell}, Z_\ell;\hat{\theta})\right]\right\| \\
&\leq \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}\left\|m_3^T(\hat{V}_{p\ell}, Z_\ell;\tilde{\theta}) - m_3^T(\hat{V}_{p\ell}, Z_\ell;\hat{\theta})\right\| \\
&\leq \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}K_4(Z_\ell)\|\tilde{\theta}-\hat{\theta}\| \\
&= \{\mathrm{E}[K_4(Z)] + o_{as}(1)\}o_{as}(1)
\end{aligned}
$$

where we have used A6-(ii) and the fact that $\tilde{\theta}$ and $\hat{\theta}$ are consistent estimators for θ_0.

Step 2: We prove $\tilde{A} - \hat{A} = o_{as}(1)$. The term $\tilde{A} - \hat{A}$ is

$$
\begin{aligned}
\tilde{A} - \hat{A} &= \left(\frac{\partial S_L^T}{\partial\theta}(\tilde{\theta})\Omega\frac{\partial S_L}{\partial\theta^T}(\overline{\theta})\right) - \left(\frac{\partial \hat{S}_L^T}{\partial\theta}(\hat{\theta})\Omega\frac{\partial \hat{S}_L}{\partial\theta^T}(\overline{\theta}^*)\right) \\
&= \left[\frac{\partial S_L^T}{\partial\theta}(\tilde{\theta})\Omega\left(\frac{\partial S_L}{\partial\theta^T}(\tilde{\theta}) + o_{as}(1)\right)\right] - \left[\frac{\partial \hat{S}_L^T}{\partial\theta}(\hat{\theta})\Omega\left(\frac{\partial \hat{S}_L}{\partial\theta^T}(\hat{\theta}) + o_{as}(1)\right)\right] \\
&= \left(\frac{\partial S_L^T}{\partial\theta}(\tilde{\theta})\Omega\frac{\partial S_L}{\partial\theta^T}(\tilde{\theta})\right) - \left(\frac{\partial \hat{S}_L^T}{\partial\theta}(\hat{\theta})\Omega\frac{\partial \hat{S}_L}{\partial\theta^T}(\hat{\theta})\right) + o_{as}(1) \\
&= \left[\left(\frac{\partial S_L^T}{\partial\theta}(\tilde{\theta}) - \frac{\partial \hat{S}_L^T}{\partial\theta}(\hat{\theta})\right)\Omega\right]\left(\frac{\partial S_L}{\partial\theta^T}(\tilde{\theta}) + \frac{\partial \hat{S}_L}{\partial\theta^T}(\hat{\theta})\right). \qquad \text{(A.7)}
\end{aligned}
$$

where the second equality comes from the following

$$
\begin{aligned}
\left\|\frac{\partial S_L}{\partial\theta^T}(\overline{\theta}) - \frac{\partial S_L}{\partial\theta^T}(\tilde{\theta})\right\| &\leq \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}\|m_3(V_{p\ell}, Z_\ell;\overline{\theta}) - m_3(V_{p\ell}, Z_\ell;\tilde{\theta})\| \\
&\leq \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}K_4(Z_\ell)\|\overline{\theta}-\tilde{\theta}\| \\
&= \{\mathrm{E}[K_4(Z)] + o_{as}(1)\}o_{as}(1) \\
&= o_{as}(1)
\end{aligned}
$$

where we use A6-(ii), the fact that $\tilde{\theta} \leq \overline{\theta} \leq \theta_0$ and that $\tilde{\theta} \xrightarrow{a.s} \theta_0$. A similar argument can be used to show that

$$\frac{\partial \hat{S}_L}{\partial \theta^T}(\overline{\theta}^*) = \frac{\partial \hat{S}_L}{\partial \theta^T}(\hat{\theta}) + o_{as}(1),$$

since $\hat{\theta} \leq \overline{\theta}^* \leq \theta_0$ and $\hat{\theta} \xrightarrow{a.s} \theta_0$.

Now, for the last line in (A.7) we observe that by Step 1, the first factor in (A.7) is $o_{as}(1)$ and the second factor can be expressed as follows

$$\begin{aligned}
\left\| \left(\frac{\partial S_L}{\partial \theta^T}(\tilde{\theta}) + \frac{\partial \hat{S}_L}{\partial \theta^T}(\hat{\theta}) \right) \right\| &= \left\| \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} [m_3(V_{p\ell}, Z_\ell; \tilde{\theta}) + m_3(\hat{V}_{p\ell}, Z_\ell; \hat{\theta})] \right\| \\
&\leq \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} \|[m_3(V_{p\ell}, Z_\ell; \tilde{\theta}) + m_3(\hat{V}_{p\ell}, Z_\ell; \hat{\theta})]\| \\
&\leq \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} \|m_3(V_{p\ell}, Z_\ell; \tilde{\theta})\| + \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} \|m_3(\hat{V}_{p\ell}, Z_\ell; \hat{\theta})\| \\
&\leq \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} \sup_{\theta\in\Theta} \|m_3(V_{p\ell}, Z_\ell; \theta)\| + \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} \|m_3(\hat{V}_{p\ell}, Z_\ell; \hat{\theta}) - m_3(\hat{V}_{p\ell}, Z_\ell; \theta_0) \\
&\quad + m_3(\hat{V}_{p\ell}, Z_\ell; \theta_0)\| \\
&\leq \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} K_5(V_{p\ell}, Z_\ell) + \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} \|m_3(\hat{V}_{p\ell}, Z_\ell; \hat{\theta}) - m_3(\hat{V}_{p\ell}, Z_\ell; \theta_0)\| \\
&\quad + \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} \|m_3(\hat{V}_{p\ell}, Z_\ell; \theta_0) - m_3(V_{p\ell}, Z_\ell; \theta_0) + m_3(V_{p\ell}, Z_\ell; \theta_0)\| \\
&\leq \{\mathrm{E}[K_5(V, Z)] + o_{as}(1)\} + \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} K_4(Z_\ell)\|\hat{\theta} - \theta_0\| \\
&\quad + \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} \|m_3(\hat{V}_{p\ell}, Z_\ell; \theta_0) - m_3(V_{p\ell}, Z_\ell; \theta_0)\| + \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} \|m_3(V_{p\ell}, Z_\ell; \theta_0)\| \\
&\leq \{\mathrm{E}[K_5(V, Z)] + o_{as}(1)\} + \{\mathrm{E}[K_4(Z)] + o_{as}(1)\}o_{as}(1) + \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} K_3(Z_\ell)|\hat{V}_{p\ell} - V_{p\ell}| \\
&\quad + \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{I_\ell}\sum_{p=1}^{I_\ell} \sup_{\theta\in\Theta} \|m_3(V_{p\ell}, Z_\ell; \theta)\| \\
&\leq \{\mathrm{E}[K_5(V, Z)] + o_{as}(1)\} + \{\mathrm{E}[K_4(Z)] + o_{as}(1)\}o_{as}(1) + \{\mathrm{E}[K_3(Z)] + o_{as}(1)\} \sup_{p\ell} |\hat{V}_{p\ell} - V_{p\ell}| \\
&\quad + \{\mathrm{E}[K_5(V, Z)] + o_{as}(1)\} \\
&= 2\{\mathrm{E}[K_5(V, Z)] + o_{as}(1)\} < \infty
\end{aligned}$$

where we use A6-(ii),(iv),(v) and the two results stated at the beginning of this Appendix.

Therefore the second factor in the last line of (A.7) converges to a finite limit and since the first factor is $o_{as}(1)$ the desired result follows.

STEP 3: We prove $\sqrt{L}(S_L(\theta_0) - \hat{S}_L(\theta_0)) = O_p(1)$.

The term $\sqrt{L}(S_L(\theta_0) - \hat{S}_L(\theta_0))$ is

$$
\begin{aligned}
B = \sqrt{L}(S_L(\theta_0) - \hat{S}_L(\theta_0)) &= \sqrt{L}\left(\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell} m(V_{p\ell}, Z_\ell; \theta_0) - \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell} m(\hat{V}_{p\ell}, Z_\ell; \theta_0)\right) \\
&= \sqrt{L}\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}\Big[m(V_{p\ell}, Z_\ell; \theta_0) - m(\hat{V}_{p\ell}, Z_\ell; \theta_0)\Big]
\end{aligned}
$$

We prove

$$
\begin{aligned}
B &= \sqrt{L}\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}\Big[m(V_{p\ell}, Z_\ell; \theta_0) - m(\hat{V}_{p\ell}, Z_\ell; \theta_0)\Big] \\
&= O_p(1) + o_{as}(1).
\end{aligned}
$$

The above expression can be rewritten as

$$
\begin{aligned}
B &= -\sqrt{L}\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}\Big[m_1(V_{p\ell}, Z_\ell; \theta_0)(\hat{V}_{p\ell} - V_{p\ell})\Big] \\
&\quad + \sqrt{L}\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}\left[m_1(V_{p\ell}, Z_\ell; \theta_0) - m_1(V^*_{p\ell}, Z_\ell; \theta_0)\right](\hat{V}_{p\ell} - V_{p\ell}) \\
&= B_1 + B_2,
\end{aligned} \tag{A.8}
$$

where the second equality comes from a Taylor expansion of order one and the following

$$
\begin{aligned}
& m(V_{p\ell}, Z_\ell; \theta_0) - m(\hat{V}_{p\ell}, Z_\ell; \theta_0) \\
&= m_1(V^*_{p\ell}, Z_\ell; \theta_0)(V_{p\ell} - \hat{V}_{p\ell}) \\
&= m_1(V_{p\ell}, Z_\ell; \theta_0)(\hat{V}_{p\ell} - V_{p\ell}) + m_1(V^*_{p\ell}, Z_\ell; \theta_0)(V_{p\ell} - \hat{V}_{p\ell}) - m_1(V_{p\ell}, Z_\ell; \theta_0)(\hat{V}_{p\ell} - V_{p\ell}) \\
&= -m_1(V_{p\ell}, Z_\ell; \theta_0)(\hat{V}_{p\ell} - V_{p\ell}) + [m_1(V_{p\ell}, Z_\ell; \theta_0) - m_1(V^*_{p\ell}, Z_\ell; \theta_0)](\hat{V}_{p\ell} - V_{p\ell})
\end{aligned}
$$

STEP 3.1: We consider B_1 in (A.8) and moreover we observe that for each i we can write

$$
\begin{aligned}
\|B_1\| &= \left\| \sqrt{L}\frac{1}{L}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i}\sum_{p=1}^{i} m_1(V_{p\ell},X_\ell,i;\theta_0)(\hat{V}_{p\ell}-V_{p\ell}) \right\| \\
&= \left\| \sqrt{L}\frac{1}{L}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i}\sum_{p=1}^{i} m_1(V_{p\ell},X_\ell,i;\theta_0)\frac{1}{i-1}\left[\frac{\hat{G}(B_{p\ell}|X_\ell,i)}{\hat{g}(B_{p\ell}|X_\ell,i)}-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\right] \right\| \\
&= \left\| \sqrt{L}\frac{1}{L}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i} m_1(V_{p\ell},X_\ell,i;\theta_0)\left\{\frac{\hat{G}(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0^2(B_{p\ell|X_\ell,i})}\right.\right. \\
&\quad \hat{g}(B_{p\ell}|X_\ell,i)+\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\frac{1}{\hat{g}(B_{p\ell}|X_\ell,i)g_0(B_{p\ell}|X_\ell,i)}\left[\hat{g}(B_{p\ell}|X_\ell,i)-g_0(B_{p\ell}|X_\ell,i)\right]^2 \\
&\quad -\frac{1}{\hat{g}(B_{p\ell}|X_\ell,i)g_0(B_{p\ell}|X_\ell,i)}\left[\hat{G}(B_{p\ell}|X_\ell,i)-G_0(B_{p\ell}|X_\ell,i)\right] \\
&\quad \left.\left. [\hat{g}(B_{p\ell}|X_\ell,i)-g_0(B_{p\ell}|X_\ell,i)]\right\}\right\| \\
&\le \left\| \sqrt{L}\frac{1}{L}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i} m_1(V_{p\ell},X_\ell,i;\theta_0)\left[\frac{\hat{G}(B_{p\ell}|X_\ell,i)}{\hat{g}(B_{p\ell}|X_\ell,i)}-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0^2(B_{p\ell|X_\ell,i})}\right.\right. \\
&\quad \left.\left.\hat{g}(B_{p\ell}|X_\ell,i)\right]\right\| + \left\| \sqrt{L}\frac{1}{L}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i} m_1(V_{p\ell},X_\ell,i;\theta_0)\right. \\
&\quad \left(\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\frac{1}{\hat{g}(B_{p\ell}|X_\ell,i)g_0(B_{p\ell}|X_\ell,i)}[\hat{g}(B_{p\ell}|X_\ell,i)-g_0(B_{p\ell}|X_\ell,i)]^2\right. \\
&\quad -\frac{1}{\hat{g}(B_{p\ell}|X_\ell,i)g_0(B_{p\ell}|X_\ell,i)}\left[\hat{G}(B_{p\ell}|X_\ell,i)-G_0(B_{p\ell}|X_\ell,i)\right] \\
&\quad \left.\left.\left[\hat{g}(B_{p\ell}|X_\ell,i)-g_0(B_{p\ell}|X_\ell,i)\right]\right)\right\| \\
&= \|B_{11}\|+\|B_{12}\| \qquad \text{(A.9)}
\end{aligned}
$$

where the third line uses the following identity

$$
\frac{\tilde{a}}{\tilde{b}}-\frac{a}{b}=\frac{\tilde{a}-\frac{a}{b}\tilde{b}}{b}+\frac{a}{b}\frac{1}{\tilde{b}b}[\tilde{b}-b]^2-\frac{1}{\tilde{b}b}[\tilde{a}-a][\tilde{b}-b].
$$

The term B_{11} can be written as

$$
\begin{aligned}
B_{11} &= \sqrt{L}\frac{1}{L}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i} m_1(V_{p\ell},X_\ell,i;\theta_0)\left[\frac{\hat{G}(B_{p\ell}|X_\ell,i)-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\hat{g}(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\right] \\
&= \sqrt{L}(R_L+\frac{L(L-1)}{L^2}U_L)
\end{aligned}
$$

$$= \quad B_{111} + B_{112}, \tag{A.10}$$

where

$$R_L = \frac{1}{L^2}\frac{L}{n_i}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i}\frac{m_1(V_{p\ell},X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}\Bigg[\omega^G_{i,R+1,j}K_{G,h_G}(0)1\!\!1(B_{p\ell}\le B_{p\ell})$$
$$-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\omega^g_{i,R,j}K_{1g,h_g}(0)K_{2g,h_g}(0)\Bigg],$$

$$U_L = \frac{1}{L(L-1)}\frac{L}{n_i}\sum_{\{\ell:I_\ell=i\}}^{L}\sum_{\{j:I_j=i,j\neq\ell\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i}\sum_{q=1}^{i}\frac{m_1(V_{p\ell},X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}$$
$$\Bigg[\omega^G_{i,R+1,j}K_{G,h_G}(X_j-X_\ell)1\!\!1(B_{qj}\le B_{p\ell})$$
$$-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\omega^g_{i,R,j}K_{1g,h_g}(X_j-X_\ell)K_{2g,h_g}(B_{qj}-B_{p\ell})\Bigg].$$

To see how to obtain the last line in (A.10), we observe that the term within brackets in the first line of (A.10) can be expressed as

$$\frac{\hat{G}(B_{p\ell}|X_\ell,i)-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\hat{g}(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)} =$$
$$= \frac{1}{g_0(B_{p\ell}|X_\ell,i)}\left[\hat{G}(B_{p\ell}|X_\ell,i)-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\hat{g}(B_{p\ell}|X_\ell,i)\right]$$
$$= \frac{1}{g_0(B_{p\ell}|X_\ell,i)}$$
$$\Bigg[\frac{1}{Lh_G}\frac{L}{n_i}\sum_{\{j:I_j=i\}}^{L}\sum_{q=1}^{i}e_1^T\left(\frac{X^T_{i,R+1}W^G_xX_{i,R+1}}{n_i}\right)^{-1}X_{R+1,j}K_G\left(\frac{X_j-X_\ell}{h_G}\right)1\!\!1(B_{qj}\le B_{p\ell})$$
$$-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}$$
$$\frac{1}{Lh_g^2}\frac{L}{n_i}\sum_{\{j:I_j=i\}}^{L}\sum_{q=1}^{i}e_1^T\left(\frac{X^T_{i,R}W^g_xX_{i,R}}{n_i}\right)^{-1}X_{R,j}K_{1g}\left(\frac{X_j-X_\ell}{h_g}\right)K_{2g}\left(\frac{B_{qj}-B_{p\ell}}{h_g}\right)\Bigg]$$
$$= \frac{1}{g_0(B_{p\ell}|X_\ell,i)}\Bigg[\frac{1}{L}\frac{L}{n_i}\sum_{\{j:I_j=i\}}^{L}\sum_{q=1}^{i}\omega^G_{i,R+1,j}K_{G,h_G}(X_j-X_\ell)1\!\!1(B_{qj}\le B_{p\ell})-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}$$
$$\frac{1}{L}\frac{L}{n_i}\sum_{\{j:I_j=i\}}^{L}\sum_{q=1}^{i}\omega^g_{i,R,j}K_{1g,h_g}(X_j-X_\ell)K_{2g,h_g}(B_{qj}-B_{p\ell})\Bigg] \tag{A.11}$$

where we have used the following notations

$$K_{G,h_G}(X_j - X_\ell) = \frac{1}{h_G} K_G \left(\frac{X_j - X_\ell}{h_G} \right), \tag{A.12}$$

$$K_{1g,h_g}(X_j - X_\ell) = \frac{1}{h_g} K_{1g} \left(\frac{X_j - X_\ell}{h_g} \right), \tag{A.13}$$

$$K_{2g,h_g}(B_{qj} - B_{p\ell}) = \frac{1}{h_g} K_{2g} \left(\frac{B_{qj} - B_{p\ell}}{h_g} \right), \tag{A.14}$$

$$\omega^G_{i,R+1,j} = e_1^T \left(\frac{X^T_{i,R+1} W^G_x X_{i,R+1}}{n_i} \right)^{-1} X_{R+1,j}, \tag{A.15}$$

$$\omega^g_{i,R,j} = e_1^T \left(\frac{X^T_{i,R} W^g_x X_{i,R}}{n_i} \right)^{-1} X_{R,j}, \tag{A.16}$$

Now using (A.11) in the first line of (A.10), we get

$$\begin{aligned} B_{11} \quad = \quad & \sqrt{L} \Bigg(\frac{1}{L^2} \frac{L}{n_i} \sum_{\{\ell: I_\ell = i\}}^{L} \sum_{\{j: I_j = i\}}^{L} \frac{1}{i(i-1)} \sum_{p=1}^{i} \sum_{q=1}^{i} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \\ & \Bigg[\omega^G_{i,R+1,j} K_{G,h_G}(X_j - X_\ell) 1\!\!1(B_{qj} \le B_{p\ell}) \\ & - \frac{G_0(B_{p\ell}|X_\ell, i)}{g_0(B_{p\ell}|X_\ell, i)} \omega^g_{i,R,j} K_{1g,h_g}(X_j - X_\ell) K_{2g,h_g}(B_{qj} - B_{p\ell}) \Bigg] \Bigg). \end{aligned} \tag{A.17}$$

The term between parenthesis in (A.17) can be decomposed as follows. Namely,

1) Diagonal terms ($\ell = j, p = q$)

$$\begin{aligned} R_L \quad = \quad & \frac{1}{L^2} \frac{L}{n_i} \sum_{\{\ell: I_\ell = i\}}^{L} \frac{1}{i(i-1)} \sum_{p=1}^{i} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \Bigg[\omega^G_{i,R+1,j} K_{G,h_G}(0) 1\!\!1(B_{p\ell} \le B_{p\ell}) \\ & - \frac{G_0(B_{p\ell}|X_\ell, i)}{g_0(B_{p\ell}|X_\ell, i)} \omega^g_{i,R,j} K_{1g,h_g}(0) K_{2g,h_g}(0) \Bigg], \end{aligned} \tag{A.18}$$

2) Off-diagonal terms ($\ell \neq j$)

$$\frac{L(L-1)}{L^2}U_L = \frac{1}{L^2}\frac{L}{n_i}\sum_{\{\ell:I_\ell=i\}}^{L}\sum_{\{j:I_j=i,j\neq\ell\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i}\sum_{q=1}^{i}\frac{m_1(V_{p\ell},X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}$$
$$\Bigg[\omega^G_{i,R+1,j}K_{G,h_G}(X_j-X_\ell)\mathbb{1}(B_{qj}\leq B_{p\ell})$$
$$-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\omega^g_{i,R,j}K_{1g,h_g}(X_j-X_\ell)K_{2g,h_g}(B_{qj}-B_{p\ell})\Bigg]. \tag{A.19}$$

From (A.18) and (A.19) we have the expression in the last line of (A.10). It remains to show that $B_{11}=B_{111}+B_{112}=o_{as}(1)$. We consider first $B_{111}=\sqrt{L}R_L$ in (A.10). Specifically,

$$\begin{aligned}
B_{111} &= \sqrt{L}\|R_L\| = \sqrt{L}\Bigg\|\Bigg(\frac{1}{L^2}\frac{L}{n_i}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i}\frac{m_1(V_{p\ell},X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}\Bigg[\omega^G_{i,R+1,j}K_{G,h_G}(0)\\
&\quad -\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\omega^g_{i,R,j}K_{1g,h_g}(0)K_{2g,h_g}(0)\Bigg]\Bigg)\Bigg\|\\
&= \sqrt{L}\Bigg\|\frac{1}{L}\frac{L}{n_i}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i}\frac{m_1(V_{p\ell},X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}\Bigg[\omega^G_{i,R+1,j}\frac{K_{G,h_G}(0)}{L}-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\\
&\quad \omega^g_{i,R,j}\frac{K_{1g,h_g}(0)K_{2g,h_g}(0)}{L}\Bigg]\Bigg\|\\
&\leq \Bigg(\frac{1}{L}\frac{L}{n_i}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i}\|m_1(V_{p\ell},X_\ell,i;\theta_0)\|^2\Bigg)^{\frac{1}{2}}\\
&\quad \sqrt{L}\Bigg(\frac{1}{L}\frac{L}{n_i}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i}\frac{1}{g_0(B_{p\ell}|X_\ell,i)^2}\Bigg[\omega^G_{i,R+1,j}\frac{K_{G,h_G}(0)}{L}-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\\
&\quad \omega^g_{i,R,j}\frac{K_{1g,h_g}(0)K_{2g,h_g}(0)}{L}\Bigg]^2\Bigg)^{\frac{1}{2}}\\
&= CD,
\end{aligned}$$

where the inequality comes from Cauchy-Schwartz. First we show that $C^2<\infty$. Using A6-(vi), $0<(1/(i-1))<1$ for each $i\in\mathcal{I}$ and $L/n_i=L/(iL_i)<\infty$ we get

$$C^2 = \frac{1}{L}\frac{L}{n_i}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i}\|m_1(V_{p\ell},X_\ell,i;\theta_0)\|^2$$

$$
\begin{aligned}
&\leq \frac{L}{n_i}\frac{1}{L}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i}\sum_{p=1}^{i}\frac{1}{i(i-1)}\sup_{\theta\in\Theta}\|m_1(V_{p\ell},X_\ell,i;\theta)\|^2 \\
&< \frac{1}{L}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i}\sum_{p=1}^{i}K_7(V_{p\ell},X_\ell,i)^2 \\
&= \mathrm{E}[K_7(V,X,i)^2]+o_{as}(1)<\infty.
\end{aligned}
$$

It remains to consider the D term above. Namely,

$$
\begin{aligned}
D &\leq \sqrt{L}\left(\frac{1}{L}\frac{L}{n_i}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i}\frac{1}{g_0(B_{p\ell}|X_\ell,i)^2}\left[\omega^G_{i,R+1,j}\frac{K_G(0)}{Lh_G}-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\right.\right. \\
&\qquad \left.\left.\omega^g_{i,R,j}\frac{K_{1g,h_g}(0)K_{2g,h_g}(0)}{Lh_g^2}\right]^2\right)^{\frac{1}{2}} \\
&= \sqrt{L}\left(\frac{1}{L}\frac{L}{n_i}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i(i-1)}\sum_{p=1}^{i}\frac{1}{g_0(B_{p\ell}|X_\ell,i)^2}\left[O_p(1)O_p\left(\frac{1}{Lh_G}\right)-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\right.\right. \\
&\qquad \left.\left.O_p(1)O_p\left(\frac{1}{Lh_g^2}\right)\right]^2\right)^{\frac{1}{2}} \\
&< \sqrt{L}\kappa_1\left[O_p\left(\frac{1}{Lh_G}\right)-\kappa_2 O_p\left(\frac{1}{Lh_g^2}\right)\right] \\
&= \kappa_1\left[O_p\left(\frac{1}{\sqrt{L}h_G}\right)-\kappa_2 O_p\left(\frac{1}{\sqrt{L}h_g^2}\right)\right] \\
&= \kappa_1[o_p(1)-\kappa_2 o_p(1)] \\
&= o_p(1),
\end{aligned}
$$

where after the first equality we use (A.12)- (A.16). The second line follows from observing that

$$
\begin{aligned}
\omega^G_{i,R+1,j} &= e_1^T\left(\frac{X^T_{i,R+1}W^G_x X_{i,R+1}}{n_i}\right)^{-1}X_{R+1,j} \\
&= e_1^T\left[\frac{1}{n_i h_G}\sum_{s=1}^{n_i}\mathbf{x}_s^T\mathbf{x}_s K_G\left(\frac{x_s-x_j}{h_g}\right)\right]^{-1}e_1 \\
&= O_p(1)
\end{aligned}
$$

and similarly for $\omega^g_{i,R,j}$.

The third line uses the fact that densities are bounded away from zero and $0<(1/i(i-1))<1$ for all i. The last line follows from Assumption A4.AN. Thus, $B_{111}=CD=o(1)$ as desired.

Let $Y_{p\ell} = (B_{p\ell}, X_\ell)$ and for each i define ,

$r_L(Y_{p\ell}, i) = E[p_L((Y_{p\ell}, i), (Y_{qj}, i))|(Y_{p\ell}, i)]$, where $p_L(\cdot, \cdot)$ is a symmetric function,

$\theta_L = E[r_L(Y_{p\ell}, i)] = E[p_L((Y_{p\ell}, i), (Y_{qj}, i))]$,

$\hat{U}_L = \theta_L + \frac{2}{L} \sum_{\{\ell: I_\ell = i\}}^{L} \frac{1}{i} \sum_{p=1}^{i} [r_L(Y_{p\ell}, i) - \theta_L]$.

Next, we consider B_{112} in (A.10)

$$\begin{aligned} B_{112} &= \frac{L(L-1)}{L^2} \sqrt{L} U_L \\ &= \frac{L(L-1)}{L^2} \sqrt{L}(U_L - \hat{U}_L) + \frac{L(L-1)}{L^2} \sqrt{L} \hat{U}_L \\ &= B_{1121} + B_{1122}, \end{aligned} \tag{A.20}$$

where U_L can be written as a U-statistic. Namely,

$$\begin{aligned} U_L &= \frac{1}{L(L-1)} \frac{L}{n_i} \sum_{\{\ell: I_\ell = i\}}^{L} \sum_{\{j: I_j = i, j \neq \ell\}}^{L} \frac{1}{i(i-1)} \sum_{p=1}^{i} \sum_{q=1}^{i} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \\ &\quad \Bigg[\omega^G_{i,R+1,j} K_{G,h_G}(X_j - X_\ell) 1\!\!1(B_{qj} \leq B_{p\ell}) \\ &\quad - \frac{G_0(B_{p\ell}|X_\ell, i)}{g_0(B_{p\ell}|X_\ell, i)} \omega^g_{i,R,j} K_{1g,h_g}(X_j - X_\ell) K_{2g,h_g}(B_{qj} - B_{p\ell}) \Bigg] \\ &= \frac{1}{L(L-1)} \sum_{\{\ell: I_\ell = i\}}^{L-1} \sum_{\{j: I_j = i, j = \ell+1\}}^{L} \frac{1}{i} \sum_{p=1}^{i} \sum_{q=1}^{i} \Bigg\{ \frac{L}{n_i(i-1)} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \\ &\quad \Bigg[\omega^G_{i,R+1,j} K_{G,h_G}(X_j - X_\ell) 1\!\!1(B_{qj} \leq B_{p\ell}) \\ &\quad - \frac{G_0(B_{p\ell}|X_\ell, i)}{g_0(B_{p\ell}|X_\ell, i)} \omega^g_{i,R,j} K_{1g,h_g}(X_j - X_\ell) K_{2g,h_g}(B_{qj} - B_{p\ell}) \Bigg] \Bigg\} \\ &= \frac{2}{L(L-1)} \sum_{\{\ell: I_\ell = i\}}^{L-1} \sum_{\{j: I_j = i, j = \ell+1\}}^{L} \frac{1}{i} \sum_{p=1}^{i} \sum_{q=1}^{i} \Bigg[\frac{m_1(V_{p\ell}, X_\ell, i; \theta_0) K^{**}(B_{p\ell}, B_{qj}, X_\ell, X_j, i)}{2} \\ &\quad + \frac{m_1(V_{qj}, X_j, i; \theta_0) K^{**}(B_{qj}, B_{p\ell}, X_j, X_\ell, i)}{2} \Bigg] \\ &= \begin{pmatrix} L \\ 2 \end{pmatrix}^{-1} \sum_{\{\ell: I_\ell = i\}}^{L-1} \sum_{\{j: I_j = i, j = \ell+1\}}^{L} \frac{1}{i} \sum_{p=1}^{i} \sum_{q=1}^{i} p_L\left((B_{p\ell}, X_\ell, i), (B_{qj}, X_j, i)\right). \end{aligned}$$

We prove $B_{1121} = \sqrt{L}(U_L - \hat{U}_L) = o_p(1)$. By Lemma 3.1 in Powell, Stock and Stoker (1989) it suffices to show that $\mathrm{E}[\|p_L((Y_{p\ell}, i), (Y_{qj}, i))\|^2] = o(L)$. We will show that $\mathrm{E}[\|p_L((Y_{p\ell}, i), (Y_{qj}, i))\|^2 | i] = o(L)$ since it implies the aforementioned condition.

$$
\begin{aligned}
&\mathrm{E}[\|p_L((Y_{p\ell}, i), (Y_{qj}, i))\|^2 | i] = \int \|p_L((Y_{p\ell}, i), (Y_{qj}, i))\|^2 g_0(Y_{p\ell}|i) g_0(Y_{qj}|i) dY_{p\ell} dY_{qj} \\
=\;& \frac{1}{4} \int \Bigg\| \frac{L}{n_i(i-1)} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \left[\frac{1}{h_G} \omega^G_{i,R+1,j} K_G\left(\frac{X_j - X_\ell}{h_G}\right) 1\!\!1(B_{qj} \le B_{p\ell}) \right. \\
& \left. - \frac{G_0(B_{p\ell}|X_\ell, i)}{g_0(B_{p\ell}|X_\ell, i)} \frac{1}{h_g^2} \omega^g_{i,R,j} K_{1g}\left(\frac{X_j - X_\ell}{h_g}\right) K_{2g}\left(\frac{B_{qj} - B_{p\ell}}{h_g}\right) \right] \\
& + \frac{L}{n_i(i-1)} \frac{m_1(V_{qj}, X_j, i; \theta_0)}{g_0(B_{qj}|X_j, i)} \left[\frac{1}{h_G} \omega^G_{i,R+1,j} K_G\left(\frac{X_\ell - X_j}{h_G}\right) 1\!\!1(B_{p\ell} \le B_{qj}) \right. \\
& \left. - \frac{G_0(B_{qj}|X_j, i)}{g_0(B_{qj}|X_j, i)} \frac{1}{h_g^2} \omega^g_{i,R,j} K_{1g}\left(\frac{X_\ell - X_j}{h_g}\right) K_{2g}\left(\frac{B_{p\ell} - B_{qj}}{h_g}\right) \right] \Bigg\|^2 \\
& g_0(Y_{p\ell}|i) g_0(Y_{qj}|i) dY_{p\ell} dY_{qj} \\
=\;& \frac{1}{4} \int \Bigg\| \left[\frac{L}{n_i(i-1)} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \frac{1}{h_G} \omega^G_{i,R+1,j} K_G\left(\frac{X_j - X_\ell}{h_G}\right) 1\!\!1(B_{qj} \le B_{p\ell}) \right. \\
& \left. + \frac{L}{n_i(i-1)} \frac{m_1(V_{qj}, X_j, i; \theta_0)}{g_0(B_{qj}|X_j, i)} \frac{1}{h_G} \omega^G_{i,R+1,j} K_G\left(\frac{X_\ell - X_j}{h_G}\right) 1\!\!1(B_{p\ell} \le B_{qj}) \right] \\
& + \left[\frac{L}{n_i(i-1)} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \frac{G_0(B_{p\ell}|X_\ell, i)}{g_0(B_{p\ell}|X_\ell, i)} \frac{1}{h_g^2} \omega^g_{i,R,j} K_{1g}\left(\frac{X_j - X_\ell}{h_g}\right) \right. \\
& K_{2g}\left(\frac{B_{qj} - B_{p\ell}}{h_g}\right) + \frac{L}{n_i(i-1)} \frac{m_1(V_{qj}, X_j, i; \theta_0)}{g_0(B_{qj}|X_j, i)} \frac{G_0(B_{qj}|X_j, i)}{g_0(B_{qj}|X_j, i)} \frac{1}{h_g^2} \omega^g_{i,R,j} K_{1g}\left(\frac{X_\ell - X_j}{h_g}\right) \\
& \left. K_{2g}\left(\frac{B_{p\ell} - B_{qj}}{h_g}\right) \right] \Bigg\|^2 g_0(Y_{p\ell}|i) g_0(Y_{qj}|i) dY_{p\ell} dY_{qj} \\
=\;& \frac{1}{4} \int \Bigg\| \frac{1}{h_G} \frac{L}{n_i} \left[\frac{1}{(i-1)} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \omega^G_{i,R+1,j} K_G\left(\frac{X_j - X_\ell}{h_G}\right) 1\!\!1(B_{qj} \le B_{p\ell}) \right. \\
& \left. + \frac{1}{(i-1)} \frac{m_1(V_{qj}, X_j, i; \theta_0)}{g_0(B_{qj}|X_j, i)} \omega^G_{i,R+1,j} K_G\left(\frac{X_\ell - X_j}{h_G}\right) 1\!\!1(B_{p\ell} \le B_{qj}) \right] \\
& + \frac{1}{h_g^2} \frac{L}{n_i} \left[\frac{1}{(i-1)} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \frac{G_0(B_{p\ell}|X_\ell, i)}{g_0(B_{p\ell}|X_\ell, i)} \omega^g_{i,R,j} K_{1g}\left(\frac{X_j - X_\ell}{h_g}\right) K_{2g}\left(\frac{B_{qj} - B_{p\ell}}{h_g}\right) \right. \\
& \left. + \frac{1}{i-1} \frac{m_1(V_{qj}, X_j, i; \theta_0)}{g_0(B_{qj}|X_j, i)} \frac{G_0(B_{qj}|X_j, i)}{g_0(B_{qj}|X_j, i)} \omega^g_{i,R,j} K_{1g}\left(\frac{X_\ell - X_j}{h_g}\right) K_{2g}\left(\frac{B_{p\ell} - B_{qj}}{h_g}\right) \right] \Bigg\|^2 \\
& g_0(Y_{p\ell}|i) g_0(Y_{qj}|i) dY_{p\ell} dY_{qj} \\
\le\;& \frac{1}{2} \int \Bigg\{ \Bigg\| \frac{L}{n_i} \frac{1}{h_G} \left[\frac{1}{(i-1)} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \omega^G_{i,R+1,j} K_G\left(\frac{X_j - X_\ell}{h_G}\right) 1\!\!1(B_{qj} \le B_{p\ell}) \right.
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{(i-1)}\frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)}\omega^G_{i,R+1,j}K_G\left(\frac{X_\ell-X_j}{h_G}\right)1\!\!1(B_{p\ell}\le B_{qj})\Bigg]\Bigg\|^2\\
&+\Bigg\|\frac{L}{n_i}\frac{1}{h_g^2}\Bigg[\frac{1}{(i-1)}\frac{m_1(V_{p\ell},X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\omega^g_{i,R,j}K_{1g}\left(\frac{X_j-X_\ell}{h_g}\right)K_{2g}\left(\frac{B_{qj}-B_{p\ell}}{h_g}\right)\\
&+\frac{1}{(i-1)}\frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)}\frac{G_0(B_{qj}|X_j,i)}{g_0(B_{qj}|X_j,i)}\omega^g_{i,R,j}K_{1g}\left(\frac{X_\ell-X_j}{h_g}\right)K_{2g}\left(\frac{B_{p\ell}-B_{qj}}{h_g}\right)\Bigg]\Bigg\|^2\Bigg\}\\
&g_0(Y_{p\ell}|i)g_0(Y_{qj}|i)dY_{p\ell}dY_{qj}\\
=\;& C+D \qquad\qquad (A.21)
\end{aligned}
$$

where the inequality comes from using $(a+b)^2 \le 2(a^2+b^2)$. Therefore we need to show that both C and D are $o(L)$.

We consider first the C term in (A.21), and note that we can write $V_{p\ell} = \xi(B_{p\ell}, X_\ell, i)$. It gives

$$
\begin{aligned}
C =\;& \frac{1}{2h_G^2}\int\left(\frac{L}{n_i}\frac{1}{(i-1)}\right)^2\Bigg\|\frac{m_1(\xi(B_{p\ell},X_\ell,i),X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}\omega^G_{i,R+1,j}K_G\left(\frac{X_j-X_\ell}{h_G}\right)1\!\!1(B_{qj}\le B_{p\ell})\\
&+\frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)}\omega^G_{i,R+1,j}K_G\left(\frac{X_\ell-X_j}{h_G}\right)1\!\!1(B_{p\ell}\le B_{qj})\Bigg\|^2 g_0(Y_{p\ell}|i)g_0(Y_{qj}|i)dY_{p\ell}dY_{qj}\\
=\;& \frac{1}{2}\int\left(\frac{L}{n_i}\frac{1}{(i-1)}\right)^2\Bigg\|\frac{m_1(\xi(uh_G+Y_{qj},i),u_2h_G+X_j,i;\theta_0)}{g_0(u_1h_G+B_{qj}|u_2h_G+X_j,i)}\omega^G_{i,R+1,j}K_G(-u_2)\\
&1\!\!1(B_{qj}\le u_1h_G+B_{qj})+\frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)}\omega^G_{i,R+1,j}K_G(u_2)1\!\!1(u_1h_G+B_{qj}\le B_{qj})\Bigg\|^2\\
&g_0(uh_G+Y_{qj}|i)g_0(Y_{qj}|i)dudY_{qj}\\
\le\;& \int\Bigg[\Bigg\|\frac{m_1(\xi(uh_G+Y_{qj},i),u_2h_G+X_j,i;\theta_0)}{g_0(u_1h_G+B_{qj}|u_2h_G+X_j,i)}\omega^G_{i,R+1,j}K_G(-u_2)1\!\!1(B_{qj}\le u_1h_G+B_{qj})\Bigg\|^2\\
&+\Bigg\|\frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)}\omega^G_{i,R+1,j}K_G(u_2)1\!\!1(u_1h_G+B_{qj}\le B_{qj})\Bigg\|^2\Bigg]\\
&g_0(uh_G+Y_{qj}|i)g_0(Y_{qj}|i)dudY_{qj}\\
=\;& C_1+C_2,
\end{aligned}
$$

where we have used the change of variable

$$u=\frac{Y_{p\ell}-Y_{qj}}{h_G}=\left(\frac{B_{p\ell}-B_{qj}}{h_G},\frac{X_\ell-X_j}{h_G}\right)=(u_1,u_2),$$

and the inequality comes from using $(a+b)^2 \le 2(a^2+b^2)$ and $\left(\frac{L}{n_i}\frac{1}{(i-1)}\right)^2 < \infty$.

Next we consider C_1

$$\begin{aligned} C_1 &= \int \left\| \frac{m_1(\xi(uh_G+Y_{qj},i),u_2h_G+X_j,i;\theta_0)}{g_0(u_1h_G+B_{qj}|u_2h_G+X_j,i)} \omega^G_{i,R+1,j} K_G(-u_2) 1\!\!1(B_{qj} \le u_1h_G+B_{qj}) \right\|^2 \\ & g_0(uh_G+Y_{qj}|i) g_0(Y_{qj}|i) du dY_{qj} \\ &= \int \left\| \frac{m_1(\xi(uh_G+Y_{qj},i),u_2h_G+X_j,i;\theta_0)}{g_0(u_1h_G+B_{qj}|u_2h_G+X_j,i)} \right\|^2 \left\| \omega^G_{i,R+1,j} K_G(-u_2) 1\!\!1(B_{qj} \le u_1h_G+B_{qj}) \right\|^2 \\ & g_0(uh_G+Y_{qj}|i) g_0(Y_{qj}|i) du dY_{qj} \\ &= \int \left\| m_1(\xi(uh_G+Y_{qj},i),u_2h_G+X_j,i;\theta_0) \right\|^2 \left\| \omega^G_{i,R+1,j} K_G(-u_2) 1\!\!1(B_{qj} \le u_1h_G+B_{qj}) \right\|^2 \\ & \frac{g_0(uh_G+Y_{qj}|i) g(Y_{qj}|i)}{g_0(u_1h_G+B_{qj}|u_2h_G+X_j,i)^2} du dY_{qj} \\ &\le \int \| m_1(\xi(uh_G+Y_{qj},i),u_2h_G+X_j,i;\theta_0) \|^2 \left\| \omega^G_{i,R+1,j} K_G(-u_2) 1\!\!1(B_{qj} \le u_1h_G+B_{qj}) \right\|^2 \\ & g_0(Y_{qj}|i) du dY_{qj}, \end{aligned}$$

where the last inequality comes from the assumption that densities are bounded.

By the Lebesgue Dominated Convergence (LDC) Theorem and A6-(vi), the above integral converges to

$$\int \left\| O_p(1) K_G(-u_2) \right\|^2 du \int \| m_1(V_{qj},X_j,i;\theta_0) \|^2 g_0(Y_{qj}|i) dY_{qj} < \infty.$$

Hence, $C_1 = o(L)$ as $L \to \infty$

A similar argument can be used to show that $C_2 = o(L)$ as $L \to \infty$. Therefore, $C = C_1 + C_2 = o(L)$.

Next we consider the D term in (A.21). Namely

$$\begin{aligned} D &= \frac{1}{2} \int \left\| \frac{L}{n_i} \frac{1}{h_g^2} \left[\frac{1}{(i-1)} \frac{m_1(V_{p\ell},X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)} \frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)} \omega^g_{i,R,j} K_{1g}\left(\frac{X_j-X_\ell}{h_g}\right) \right.\right. \\ & K_{2g}\left(\frac{B_{qj}-B_{p\ell}}{h_g}\right) + \frac{1}{(i-1)} \frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)} \frac{G_0(B_{qj}|X_j,i)}{g_0(B_{qj}|X_j,i)} \omega^g_{i,R,j} \\ & \left.\left. K_{1g}\left(\frac{X_\ell-X_j}{h_g}\right) K_{2g}\left(\frac{B_{p\ell}-B_{qj}}{h_g}\right) \right] \right\|^2 g_0(Y_{p\ell}|i) g_0(Y_{qj}|i) dY_{p\ell} dY_{qj} \\ &= \frac{1}{2h_g^4} \left(\frac{L}{n_i} \frac{1}{(i-1)} \right)^2 \int \left\| \frac{m_1(\xi(B_{p\ell},X_\ell,i),X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)} \frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)} \omega^g_{i,R,j} \right. \\ & K_{1g}\left(\frac{X_j-X_\ell}{h_g}\right) K_{2g}\left(\frac{B_{qj}-B_{p\ell}}{h_g}\right) \\ & \left. + \frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)} \frac{G_0(B_{qj}|X_j,i)}{g_0(B_{qj}|X_j,i)} \omega^g_{i,R,j} K_{1g}\left(\frac{X_\ell-X_j}{h_g}\right) K_{2g}\left(\frac{B_{p\ell}-B_{qj}}{h_g}\right) \right\|^2 \end{aligned}$$

$$
\begin{aligned}
& g_0(Y_{p\ell}|i)g_0(Y_{qj}|i)dY_{p\ell}dY_{qj} \\
\leq\ & \frac{1}{2h_g^2}\int \Bigg\| \frac{m_1(\xi(uh_G+Y_{qj},i),u_2h_G+X_j,i;\theta_0)}{g_0(u_1h_G+B_{qj}|u_2h_G+X_j,i)}\frac{G_0(u_1h_g+B_{qj}|u_2h_g+X_j,i)}{g_0(u_1h_g+B_{qj}|u_2h_g+X_j,i)} \\
& \omega^g_{i,R,j}K_{1g}(-u_2)K_{2g}(-u_1) \\
& +\frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)}\frac{G_0(B_{qj}|X_j,i)}{g_0(B_{qj}|X_j,i)}\omega^g_{i,R,j}K_{1g}(u_2)K_{2g}(u_1)\Bigg\|^2 g_0(uh_g+Y_{qj}|i)g_0(Y_{qj}|i)dudY_{qj} \\
\leq\ & \frac{1}{h_g^2}\int \Bigg[\Bigg\| \frac{m_1(\xi(uh_G+Y_{qj},i),u_2h_G+X_j,i;\theta_0)}{g_0(u_1h_G+B_{qj}|u_2h_G+X_j,i)}\frac{G_0(u_1h_g+B_{qj}|u_2h_g+X_j,i)}{g_0(u_1h_g+B_{qj}|u_2h_g+X_j,i)}\omega^g_{i,R,j} \\
& K_{1g}(-u_2)K_{2g}(-u_1)\Bigg\|^2 + \Bigg\|\frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)}\frac{G_0(B_{qj}|X_j,i)}{g_0(B_{qj}|X_j,i)}\omega^g_{i,R,j}K_{1g}(u_2)K_{2g}(u_1)\Bigg\|^2\Bigg] \\
& g_0(uh_g+Y_{qj}|i)g_0(Y_{qj}|i)dudY_{qj} \\
=\ & D_1+D_2,
\end{aligned}
$$

where we have used the change of variable

$$u=\frac{Y_{p\ell}-Y_{qj}}{h_g}=\left(\frac{B_{p\ell}-B_{qj}}{h_g},\frac{X_\ell-X_j}{h_g}\right)=(u_1,u_2),$$

the first inequality comes from the fact that $\left(\frac{L}{n_i}\frac{1}{(i-1)}\right)^2<\infty$, and the second inequality comes from using $(a+b)^2\leq 2(a^2+b^2)$.

We consider first D_1. Specifically,

$$
\begin{aligned}
D_1 &= \frac{1}{h_g^2}\int \Big\|m_1(\xi(uh_g+Y_{qj},i),u_2h_g+X_j,i;\theta_0)G_0(u_1h_g+B_{qj}|u_2h_g+X_j,i)\Big\|^2 \\
& \Big\|\omega^g_{i,R,j}K_{1g}(-u_2)K_{2g}(-u_1)\Big\|^2\frac{g_0(uh_g+Y_{qj}|i)g_0(Y_{qj}|i)}{g_0(u_1h_g+B_{qj}|u_2h_g+X_j,i)^4}dudY_{qj} \\
&\leq \frac{1}{h_g^2}\int \|m_1(\xi(uh_g+Y_{qj},i),u_2h_g+X_j,i;\theta_0)\|^2\Big\|\omega^g_{i,R,j}K_{1g}(-u_2)K_{2g}(-u_1)\Big\|^2 \\
& g_0(Y_{qj}|i)dudY_{qj}
\end{aligned}
$$

where the inequality uses the fact that $G(\cdot|\cdot,i)$ is bounded and that densities are bounded from above.

By the LDC Theorem and A6-(vi) the above integral converges to

$$\int \|O_p(1)K_{1g}(-u_2)K_{2g}(-u_1)\|^2du\int\|m_1(V_{qj},X_j,i;\theta_0)\|^2g_0(Y_{qj}|i)dY_{qj}<\infty.$$

Hence, $D_1 = o(L)$ if and only if $Lh_g^2 \to \infty$ as implied by A4.AN-(ii) since

$$Lh_g^2 = \sqrt{L}\sqrt{L}h_g^2 \to \infty.$$

A similar argument can be used to show that $D_2 = o(L)$. That is, $D_2 = o(L)$ if and only if $Lh_g^2 \to \infty$, as implied by A4.AN-(ii).

Therefore, $C + D = C_1 + C_2 + D_1 + D_2 = o(L)$ and the desired result follows, i.e by Lemma 3.1 in Powell, Stock and Stoker (1989) $\sqrt{L}(U_L - \hat{U}_L) = o_p(1)$.

Next we consider the second term in (A.20)

$$\begin{aligned}
B_{1122} &= \frac{L(L-1)}{L^2}\sqrt{L}\hat{U}_L \\
&= \frac{L(L-1)}{L^2}\sqrt{L}\left\{\theta_L + \frac{2}{L}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i}\sum_{p=1}^{i}[r_L(Y_{p\ell},i)-\theta_L]\right\} \\
&= \frac{L(L-1)}{L^2}\sqrt{L}E[p_L((Y_{p\ell},i),(Y_{qj},i))] + \frac{L(L-1)}{L^2}\sqrt{L}\frac{2}{L}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i}\sum_{p=1}^{i}[r_L(Y_{p\ell},i)-\theta_L].
\end{aligned}$$

By the Central Limit Theorem (CLT), the second term above is $O_p(1)$ (See Lemma A1 below). Therefore, it remains to show that the first term is $o_{as}(1)$. We consider the expectation in the first term above. Namely

$$\begin{aligned}
&\mathrm{E}[p_L((Y_{p\ell},i),(Y_{qj},i))] \\
&= \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)}\int\left\{\frac{m_1(V_{p\ell},X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}\left[\omega_{i,R+1,j}^{G}K_{G,h_G}(X_j-X_\ell)\mathbb{1}(B_{qj}\le B_{p\ell})\right.\right. \\
&\quad \left.-\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\omega_{i,R,j}^{g}K_{1g,h_g}(X_j-X_\ell)K_{2g,h_g}(B_{qj}-B_{p\ell})\right] \\
&\quad +\frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)}\left[\omega_{i,R+1,j}^{G}K_{G,h_G}(X_\ell-X_j)\mathbb{1}(B_{p\ell}\le B_{qj})\right. \\
&\quad \left.\left.-\frac{G_0(B_{qj}|X_j,i)}{g_0(B_{qj}|X_j,i)}\omega_{i,R,j}^{g}K_{1g,h_g}(X_\ell-X_j)K_{2g,h_g}(B_{p\ell}-B_{qj})\right]\right\} \\
&\quad g_0(B_{p\ell},X_\ell,i)g_0(B_{qj},X_j,i)dY_{p\ell}dY_{qj} \\
&= \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)}\int\frac{1}{h_G}\left[\frac{m_1(V_{p\ell},X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}\omega_{i,R+1,j}^{G}K_G\left(\frac{X_j-X_\ell}{h_G}\right)\mathbb{1}(B_{qj}\le B_{p\ell})\right. \\
&\quad \left.+\frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)}\omega_{i,R+1,j}^{G}K_G\left(\frac{X_\ell-X_j}{h_G}\right)\mathbb{1}(B_{p\ell}\le B_{qj})\right]
\end{aligned}$$

$$
\begin{aligned}
& g_0(B_{p\ell},X_\ell,i)g_0(B_{qj},X_j,i)dY_{p\ell}dY_{qj} \\
& -\frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)}\int \frac{1}{h_g^2}\left[\frac{m_1(V_{p\ell},X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}\frac{G_0(B_{p\ell}|X_\ell,i)}{g_0(B_{p\ell}|X_\ell,i)}\omega^g_{i,R,j}K_{1g}\left(\frac{X_j-X_\ell}{h_g}\right)\right. \\
& K_{2g}\left(\frac{B_{qj}-B_{p\ell}}{h_g}\right)+\frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)}\frac{G_0(B_{qj}|X_j,i)}{g_0(B_{qj}|X_j,i)}\omega^g_{i,R,j}K_{1g}\left(\frac{X_\ell-X_j}{h_g}\right) \\
& \left. K_{2g}\left(\frac{B_{p\ell}-B_{qj}}{h_g}\right)\right]g_0(B_{p\ell},X_\ell,i)g_0(B_{qj},X_j,i)dY_{p\ell}dY_{qj} \\
= \; & A_1-A_2. \qquad \text{(A.22)}
\end{aligned}
$$

We consider first A_1.

$$
\begin{aligned}
\|A_1\| \; = \; & \left\|\frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)}\int \frac{1}{h_G}\left[\frac{m_1(\xi(B_{p\ell},X_\ell,i),X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}\omega^G_{i,R+1,j}K_G\left(\frac{X_j-X_\ell}{h_G}\right)\right.\right. \\
& \left. 1\!\!1(B_{qj}\le B_{p\ell})+\frac{m_1(V_{qj},X_j,i;\theta_0)}{g_0(B_{qj}|X_j,i)}\omega^G_{i,R+1,j}K_G\left(\frac{X_\ell-X_j}{h_G}\right)1\!\!1(B_{p\ell}\le B_{qj})\right] \\
& \left. g_0(B_{p\ell},X_\ell,i)g_0(B_{qj},X_j,i)dY_{p\ell}dY_{qj}\right\| \\
\le \; & \|A_{11}\|+\|A_{12}\|
\end{aligned}
$$

It is enough to show that $\|A_{11}\| = o_{as}(1/\sqrt{L})$ since the same argument can be used to show that $\|A_{12}\| = o_{as}(1/\sqrt{L})$. We observe the following

$$
\begin{aligned}
\|A_{11}\| \; = \; & \left\|\frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)}\int \frac{1}{h_G}\left[\frac{m_1(\xi(B_{p\ell},X_\ell,i),X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}\omega^G_{i,R+1,j}K_G\left(\frac{X_j-X_\ell}{h_G}\right)\right.\right. \\
& \left.\left. 1\!\!1(B_{qj}\le B_{p\ell})\right]g_0(B_{p\ell},X_\ell,i)g_0(B_{qj},X_j,i)dY_{p\ell}dY_{qj}\right\| \\
\le \; & \left\|\sum_i\int \frac{1}{h_G}\frac{m_1(\xi(B_{p\ell},X_\ell,i),X_\ell,i;\theta_0)}{g_0(B_{p\ell}|X_\ell,i)}\omega^G_{i,R+1,j}K_G\left(\frac{X_j-X_\ell}{h_G}\right)1\!\!1(B_{qj}\le B_{p\ell})\right. \\
& \left. g_0(B_{p\ell},X_\ell,i)g_0(B_{qj},X_j,i)dY_{p\ell}dY_{qj}\right\| \\
\le \; & \left\|\sum_i\int \frac{1}{h_G}m_1(\xi(B_{p\ell},X_\ell,i),X_\ell,i;\theta_0)\omega^G_{i,R+1,j}K_G\left(\frac{X_j-X_\ell}{h_G}\right)1\!\!1(B_{qj}\le B_{p\ell})\right. \\
& \left. g_0(B_{p\ell},X_\ell,i)dY_{p\ell}dY_{qj}\right\| \\
= \; & h_G\left\|\sum_i\int m_1(\xi(uh_G+Y_{qj},i),u_2h_G+X_j,i;\theta_0)\omega^G_{i,R+1,j}K_G(-u_2)\right.
\end{aligned}
$$

$$
\begin{aligned}
& \mathbb{1}(B_{qj} \leq u_1 h_G + B_{qj}) g_0(uh_G + Y_{qj}, i) du dY_{qj} \Bigg\| \\
\leq \quad & h_G \Bigg\| \sum_i \int m_1(\xi(uh_G + Y_{qj}, i), u_2 h_G + X_j, i; \theta_0) K_G(-u_2) \\
& \mathbb{1}(B_{qj} \leq u_1 h_G + B_{qj}) g_0(uh_G + Y_{qj}, i) du dY_{qj} \Bigg\| \qquad \text{(A.23)}
\end{aligned}
$$

where we have used the fact that densities are bounded and the change of variable $u = \frac{Y_{p\ell} - Y_{qj}}{h_G} = \left(\frac{B_{p\ell} - B_{qj}}{h_G}, \frac{X_\ell - X_j}{h_G} \right) = (u_1, u_2)$. The last inequality comes from observing that $\omega^G_{i,R+1,j} = O_p(1)$.

We consider now the expectation inside the norm of the term above evaluated at $h_G = 0$, namely

$$
\begin{aligned}
& \sum_i \int m_1(\xi(Y_{qj}, i), X_j, i; \theta_0) g_0(Y_{qj}, i) dY_{qj} \\
= \quad & \sum_i \int \left[\int m_1(\xi(Y_{qj}, i), X_j, i; \theta_0) g_0(B_{qj}|X_j, i) dB_{qj} \right] f_m(X_j, i) dX_j \\
= \quad & \mathrm{E}[m_1(V, X, i; \theta_0)]
\end{aligned}
$$

where we use A3-(ii) and the Law of Iterated Expectations. The last line in the expression above follows from observing that the integral inside can be solved by using twice integration by parts as follows

$$
\begin{aligned}
& \int_{\underline{B}(X_j)}^{\overline{B}(X_j,i)} m_1(\xi(B_{qj}, X_j, i), X_j, i; \theta_0) g_0(B_{qj}|X_j, i) dB_{qj} \\
= \quad & m_1(\xi(\overline{B}(X_j, i), X_j, i), X_j, i; \theta_0) G_0(\overline{B}(X_j, i)|X_j, i) - m_1(\xi(\underline{B}(X_j), X_j, i), X_j, i; \theta_0) \\
& G_0(\underline{B}(X_j)|X_j, i) - \int_{\underline{B}(X_j)}^{\overline{B}(X_j,i)} m_{11}(\xi(B_{qj}, X_j, i), X_j, i; \theta_0) G_0(B_{qj}|X_j, i) dB_{qj} \\
= \quad & m_1(\overline{V}, X_j, i; \theta_0) - \int_{\underline{B}(X_j)}^{\overline{B}(X_j,i)} m_{11}(\xi(B_{qj}, X_j, i), X_j, i; \theta_0) G_0(B_{qj}|X_j, i) dB_{qj} \\
= \quad & m_1(\overline{V}, X_j, i; \theta_0) - m_1(\xi(\overline{B}(X_j, i), X_j, i), X_j, i; \theta_0) G_0(\overline{B}(X_j, i)|X_j, i) \\
& + m_1(\xi(\underline{B}(X_j), X_j, i), Z_j; \theta_0) G_0(\underline{B}(X_j)|X_j, i) \\
& + \int_{\underline{B}(X_j)}^{\overline{B}(X_j,i)} m_1(\xi(B_{qj}, X_j, i), X_j, i; \theta_0) g_0(B_{qj}|X_j, i) dB_{qj} \\
= \quad & m_1(\overline{V}, X_j, i; \theta_0) - m_1(\overline{V}, X_j, i; \theta_0) \\
& + \int_{\underline{B}(X_j)}^{\overline{B}(X_j,i)} m_1(\xi(B_{qj}, X_j, i), X_j, i; \theta_0) g_0(B_{qj}|X_j, i) dB_{qj}
\end{aligned}
$$

$$
\begin{aligned}
&= \int_{\underline{V}(X_j,i)}^{\overline{V}(X_j,i)} m_1(V_{qj}, X_j, i; \theta_0) \frac{g_0(B_{qj}|X_j,i)}{\xi_1(B_{qj}, X_j, i)} dV_{qj} \\
&= \mathrm{E}[m_1(V, X, i; \theta_0)|X, i]
\end{aligned}
$$

where the fifth equality uses $G_0(B_{qj}|X_j,i) = F(\xi(B_{qj}, X_j, i)|X_j, i)$, so that $g_0(B_{qj}|X_j,i) = f(V_{qj}|X_j,i)\xi_1(B_{qj}, X_j, i)$.

Therefore at $h_G = 0$ the integral inside the norm in (A.23) vanishes by A7-(i). Thus, we can apply a Taylor expansion of order $R+1$ in the RHS of (A.23) around h_G to obtain

$$
\begin{aligned}
\|A_{11}\| &\leq h_G \sum_i \left\| d_1 h_G + d_2 \frac{h_G^2}{2} + \ldots + d_R \frac{h_G^R}{R!} + O(h_G^{R+1}) \right\| \\
&= \sum_i \left\| d_1 h_G^2 + d_2 \frac{h_G^3}{2} + \ldots + d_R \frac{h_G^{R+1}}{R!} + O(h_G^{R+2}) \right\|
\end{aligned}
$$

We note that the remainder term vanishes, i.e. $\sqrt{L} h_G^{R+2} = o(1)$, and also that $\sqrt{L} h_G^{R+1} = o(1)$, by A4.AN-(i). The remaining $R-1$ terms also vanish by A3-(iii), i.e, since the kernels are of order $R-1$. To see this observe that the kth coordinate of d_ρ, $\rho = 1, \ldots, R-1$ is

$$
\begin{aligned}
d_{k_\rho} &= \frac{\partial^\rho}{\partial h_G^\rho} \int [H_k(uh_G + \overline{Y}) - H_k(uh_G + \underline{Y})] K_G(-u_2) du|_{h_G=0} \\
&= \sum_{k_1,\ldots,k_\rho=1}^{2} \int (u_{k_1} \ldots u_{k_\rho}) K_G(-u_2) \frac{\partial^\rho}{\partial Y_{k_1} \ldots \partial Y_{k_\rho}} H_k(\overline{Y}) du \\
&\quad - \sum_{k_1,\ldots,k_\rho=1}^{2} \int (u_{k_1} \ldots u_{k_\rho}) K_G(-u_2) \frac{\partial^\rho}{\partial Y_{k_1} \ldots \partial Y_{k_\rho}} H_k(\underline{Y}) du \\
&= 0,
\end{aligned}
$$

where $dH_k/dY(y) = m_{1,k}(\xi(y,i), x, i; \theta_0) g_0(y,i)$. The third equality uses A3-(iii), that is since $K_G(\cdot)$ is a higher order kernel, all moments of order strictly smaller than $R-1$ vanish.

It remains to consider now A_2 in (A.22). Namely

$$
\begin{aligned}
\|A_2\| = &\Bigg\| \frac{1}{2} \frac{L}{n_i} \sum_i \frac{1}{(i-1)} \int \frac{1}{h_g^2} \Bigg[\frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \frac{G_0(B_{p\ell}|X_\ell, i)}{g_0(B_{p\ell}|X_\ell, i)} \\
&\omega_{i,R,j}^g K_{1g}\left(\frac{X_j - X_\ell}{h_g}\right) K_{2g}\left(\frac{B_{qj} - B_{p\ell}}{h_g}\right) \\
&+ \frac{m_1(V_{qj}, X_j, i; \theta_0)}{g_0(B_{qj}|X_j, i)} \frac{G_0(B_{qj}|X_j, i)}{g_0(B_{qj}|X_j, i)} \omega_{i,R,j}^g K_{1g}\left(\frac{X_\ell - X_j}{h_g}\right) K_{2g}\left(\frac{B_{p\ell} - B_{qj}}{h_g}\right) \Bigg]
\end{aligned}
$$

$$
\begin{aligned}
& g_0(B_{p\ell}, X_\ell, i) g_0(B_{qj}, X_j, i) dY_{p\ell} dY_{qj} \Bigg\| \\
\leq \quad & \|A_{21}\| + \|A_{21}\|
\end{aligned}
$$

We show only that $A_{21} = o_{as}(1/\sqrt{L})$ since a similar argument can be used to show that $A_{22} = o_{as}(1/\sqrt{L})$. We observe the following

$$
\begin{aligned}
\|A_{21}\| \quad = \quad & \Bigg\| \frac{1}{2} \frac{L}{n_i} \sum_i \frac{1}{(i-1)} \int \frac{1}{h_g^2} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \frac{G_0(B_{p\ell}|X_\ell, i)}{g_0(B_{p\ell}|X_\ell, i)} \\
& \omega^g_{i,R,j} K_{1g} \left(\frac{X_j - X_\ell}{h_g} \right) K_{2g} \left(\frac{B_{qj} - B_{p\ell}}{h_g} \right) g_0(B_{p\ell}, X_\ell, i) g_0(B_{qj}, X_j, i) dY_{p\ell} dY_{qj} \Bigg\| \\
\leq \quad & \Bigg\| \sum_i \int \frac{1}{h_g^2} \frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)} \frac{G_0(B_{p\ell}|X_\ell, i)}{g_0(B_{p\ell}|X_\ell, i)} \omega^g_{i,R,j} K_{1g} \left(\frac{X_j - X_\ell}{h_g} \right) \\
& K_{2g} \left(\frac{B_{qj} - B_{p\ell}}{h_g} \right) g_0(B_{p\ell}, X_\ell, i) g_0(B_{qj}, X_j, i) dY_{p\ell} dY_{qj} \Bigg\| \\
\leq \quad & \Bigg\| \sum_i \int \frac{1}{h_g^2} m_1(\xi(B_{p\ell}, X_\ell, i), X_\ell, i; \theta_0) \omega^g_{i,R,j} K_{1g} \left(\frac{X_j - X_\ell}{h_g} \right) K_{2g} \left(\frac{B_{qj} - B_{p\ell}}{h_g} \right) \\
& g_0(B_{p\ell}, X_\ell, i) g_0(B_{qj}, X_j, i) dY_{p\ell} dY_{qj} \Bigg\| \\
= \quad & \Bigg\| \sum_i \int m_1(\xi(uh_g + Y_{qj}, i), u_2 + X_j, i; \theta_0) \omega^g_{i,R,j} K_{1g,h_g}(-u_2) K_{2g,h_g}(-u_1) \\
& g_0(uh_g + Y_{qj}, i) du dY_{qj} \Bigg\| \\
\leq \quad & \Bigg\| \sum_i \int m_1(\xi(uh_g + Y_{qj}, i), u_2 + X_j, i; \theta_0) K_{1g,h_g}(-u_2) K_{2g,h_g}(-u_1) \\
& g_0(uh_g + Y_{qj}, i) du dY_{qj} \Bigg\|,
\end{aligned}
$$

where we have used that $(1/2)(L/n_i)1/(i-1) \leq \infty$ and also that densities are bounded. The last equality uses the change of variable $u = (Y_{p\ell} - Y_{qj})/h_g$ and the last inequality comes from observing that $\omega^g_{i,R,j} = O_p(1)$.

We observe that A_{21} can be expanded as a Taylor series of order R in the bandwidth h_g. Moreover, $A_{21} = 0|_{h_g=0}$ by A7-(i) as we have already shown above for $A_{11} = 0|_{hG=0}$.

Therefore, by A7-(i) we can apply a Taylor expansion around h_g to obtain

$$\|A_{21}\| \leq \sum_{I_j} \left\| c_1 h_g + c_2 \frac{h_g^2}{2} + \ldots + c_{R-1} \frac{h_g^{R-1}}{(R-1)!} + O(h_g^R) \right\|.$$

We note that the remainder term vanishes, i.e. $\sqrt{L} h_g^R = o(1)$ by A4.AN-(ii). The remaining $R-1$ terms also vanish by A3-(iii). To see this observe that the kth coordinate of c_ρ, $\rho = 1, \ldots, R-1$ is

$$\begin{aligned}
c_{k_\rho} &= \frac{\partial^\rho}{\partial h_g^\rho} \int [H_k(uh_g + \overline{Y}) - H_k(uh_g + \underline{Y})] K_{1g}(-u_2) K_{2g}(-u_1) du|_{h_g=0} \\
&= \sum_{k_1,\ldots,k_\rho=1}^{2} \int (u_{k_1} \ldots u_{k_\rho}) K_{1g}(-u_2) K_{2g}(-u_1) \frac{\partial^\rho}{\partial Y_{k_1} \ldots \partial Y_{k_\rho}} H_k(\overline{Y}) du \\
&\quad - \sum_{k_1,\ldots,k_\rho=1}^{2} \int (u_{k_1} \ldots u_{k_\rho}) K_{1g}(-u_2) K_{2g}(-u_1) \frac{\partial^\rho}{\partial Y_{k_1} \ldots \partial Y_{k_\rho}} H_k(\underline{Y}) du \\
&= 0,
\end{aligned}$$

where $dH_k/dY(y) = m_{1,k}(\xi(y,i), x, i; \theta_0) g_0(y,i)$. The third equality uses A3-(iii), that is since $K_{1g}(\cdot)$ and $K_{2g}(\cdot)$ are higher order kernels, all moments of order strictly smaller than $R-1$ vanish.

We consider next B_{12} in (A.9)

$$\begin{aligned}
\|B_{12}\| &= \Bigg\| \frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{I_\ell(I_\ell - 1)} \sum_{p=1}^{I_\ell} m_1(V_{p\ell, Z_\ell;\theta_0}) \sqrt{L} \Bigg[\frac{1}{\hat{g}(B_{p\ell}|Z_\ell) g_0(B_{p\ell}|Z_\ell)} \Bigg(\frac{G_0(B_{p\ell}|Z_\ell)}{g_0(B_{p\ell}|Z_\ell)} \\
&\quad [\hat{g}(B_{p\ell}|Z_\ell) - g_0(B_{p\ell}|Z_\ell)]^2 - \Big[\hat{G}(B_{p\ell}|Z_\ell) - G_0(B_{p\ell}|Z_\ell)\Big] [\hat{g}(B_{p\ell}|Z_\ell) - g_0(B_{p\ell}|Z_\ell)] \Bigg) \Bigg] \Bigg\| \\
&\leq \left(\frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{I_\ell(I_\ell - 1)} \sum_{p=1}^{I_\ell} \|m_1(V_{p\ell}, Z_\ell; \theta_0)\|^2 \right)^{\frac{1}{2}} \\
&\quad \Bigg\{ \frac{1}{L} \sum_{\ell=1}^{L} \frac{1}{I_\ell(I_\ell - 1)} \sum_{p=1}^{I_\ell} L \Bigg[\frac{1}{\hat{g}(B_{p\ell}|Z_\ell) g_0(B_{p\ell}|Z_\ell)} \Bigg(\frac{G_0(B_{p\ell}|Z_\ell)}{g_0(B_{p\ell}|Z_\ell)} [\hat{g}(B_{p\ell}|Z_\ell) - g_0(B_{p\ell}|Z_\ell)] \\
&\quad - \Big[\hat{G}(B_{p\ell}|Z_\ell) - G_0(B_{p\ell}|Z_\ell)\Big] \Big[\hat{g}(B_{p\ell}|Z_\ell) - g_0(B_{p\ell}|Z_\ell)\Big] \Bigg) \Bigg]^2 \Bigg\}^{\frac{1}{2}} \\
&= B_{121} B_{122},
\end{aligned}$$

where the inequality comes from Cauchy-Schwartz.

First we show that $B_{121}^2 < \infty$.

$$
\begin{aligned}
B_{121}^2 &= \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell(I_\ell-1)}\sum_{p=1}^{I_\ell}\|m_1(V_{p\ell},Z_\ell;\theta_0)\|^2\\
&\le \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell(I_\ell-1)}\sum_{p=1}^{I_\ell}\sup_{\theta\in\Theta}\|m_1(V_{p\ell},Z_\ell;\theta)\|^2\\
&\le \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell(I_\ell-1)}\sum_{p=1}^{I_\ell}K_7(V_{p\ell},Z_\ell)^2\\
&\le \frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}K_7(V_{p\ell},Z_\ell)^2\\
&= \mathrm{E}[K_7(V,Z)^2]+o_{as}(1)<\infty
\end{aligned}
$$

where the second inequality comes from A6-(vi) and the fact that $0 < 1/(I_\ell - 1) \le 1$.

Next we show that $B_{122} = o(1)$.

$$
\begin{aligned}
B_{122} = &\left\{\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell(I_\ell-1)}\sum_{p=1}^{I_\ell}L\left[\frac{1}{\hat{g}(B_{p\ell}|Z_\ell)g_0(B_{p\ell}|Z_\ell)}\left(\frac{G_0(B_{p\ell}|Z_\ell)}{g_0(B_{p\ell}|Z_\ell)}\left[\hat{g}(B_{p\ell}|Z_\ell)-g_0(B_{p\ell}|Z_\ell)\right]^2\right.\right.\right.\\
&\left.\left.\left.-\left[\hat{G}(B_{p\ell}|Z_\ell)-G_0(B_{p\ell}|Z_\ell)\right]\left[\hat{g}(B_{p\ell}|Z_\ell)-g_0(B_{p\ell}|Z_\ell)\right]\right)\right]^2\right\}^{\frac{1}{2}}\\
&< \sqrt{L}\left\{\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell(I_\ell-1)}\sum_{p=1}^{I_\ell}\left[\kappa_1\left(\kappa_2 O\left(\frac{1}{{r_g}^2}\right)-O\left(\frac{1}{r_G}\right)O\left(\frac{1}{r_g}\right)\right)\right]^2\right\}^{\frac{1}{2}}\\
&\le \sqrt{L}\kappa_1\left[\kappa_2 O\left(\frac{1}{{r_g}^2}\right)-O\left(\frac{1}{r_G}\right)O\left(\frac{1}{r_g}\right)\right]\\
&= \kappa_1\left[\kappa_2 O\left(\frac{\sqrt{L}}{{r_g}^2}\right)-O\left(\frac{\sqrt{L}}{r_G r_g}\right)\right]=o(1),
\end{aligned}
$$

where we have used

- $\left|\dfrac{1}{\hat{g}(B_{p\ell}|Z_\ell)g_0(B_{p\ell}|Z_\ell)}\right| < \kappa_1 < \infty$,

  since densities are bounded away from zero and $\hat{g}(B_{p\ell}|Z_\ell) \xrightarrow{a.s} g_0(B_{p\ell}|Z_\ell)$
- $\left|\dfrac{G_0(B_{p\ell}|Z_\ell)}{g_0(B_{p\ell}|Z_\ell)}\right| < \kappa_2 < \infty$,
  since $g_0(\cdot,\cdot)$ is bounded away from zero.

- $$O\left(\frac{1}{r_g^2}\right) = \left|\hat{g}(B_{p\ell}|Z_\ell) - g_0(B_{p\ell}|Z_\ell)\right|^2 = O\left(h_{1g}^{2R} + h_{2g}^{2R} + \frac{\log L}{Lh_{1g}h_{2g}}\right)$$

$$\begin{aligned} O\left(\frac{1}{r_G}\right)O\left(\frac{1}{r_g}\right) &= \left|\hat{G}(B_{p\ell}|Z_\ell) - G_0(B_{p\ell}|Z_\ell)\right|\left|\hat{g}(B_{p\ell}|Z_\ell) - g_0(B_{p\ell}|Z_\ell)\right| \\ &= O\left(h_G^{R+1} + \sqrt{\frac{\log L}{Lh_G}}\right)O\left(h_{1g}^R + h_{2g}^R + \sqrt{\frac{\log L}{Lh_{1g}h_{2g}}}\right) \end{aligned}$$

- For each ℓ, $0 < 1/(I_\ell - 1) \leq 1$.

Therefore, the second term in (A.9) is $o(1)$, i.e. $B_{12} = o(1)$

Step 3.B2

We consider B_2 in (A.8)

$$\begin{aligned} \|B_2\| &\leq \sqrt{L}\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}\|m_1(V_{p\ell}, Z_\ell;\theta_0) - m_1(V_{p\ell}^*, Z_\ell;\theta_0)\||\hat{V}_{p\ell} - V_{p\ell}| \\ &\leq \sqrt{L}\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}K_6(Z_\ell)|V_{p\ell} - V_{p\ell}^*||\hat{V}_{p\ell} - V_{p\ell}| \\ &\leq \sqrt{L}\frac{1}{L}\sum_{\ell=1}^{L}\frac{1}{I_\ell}\sum_{p=1}^{I_\ell}K_6(Z_\ell)(\hat{V}_{p\ell} - V_{p\ell})^2 \\ &\leq \sqrt{L}\sup_{p,\ell}(\hat{V}_{p\ell} - V_{p\ell})^2\frac{1}{L}\sum_{\ell=1}^{L}K_6(Z_\ell) \\ &= \sqrt{L}O_{as}\left(\frac{1}{r^2}\right)O_{as}(1) \\ &= O_{as}\left(\frac{\sqrt{L}}{r^2}\right)O_{as}(1) \\ &= O_{as}\left(\frac{L^{1/4}}{r}\right)O_{as}(1) = o_{as}(1), \end{aligned}$$

where the second inequality comes from A6-(v), the third uses the fact that $\hat{V}_{p\ell} \leq V_{p\ell}^* \leq V_{p\ell}$. The last equality follows from A4.AN.

Therefore, the desired result follows. *Q.E.D.*

LEMMA A1: *Let $\hat{\theta}$ be defined as in (7). Then, under A1-A3, A4.AN and A5-A.7, we have*

$$
\begin{aligned}
\sqrt{L}(\hat{\theta}-\theta_0) &= \sqrt{L}(\tilde{\theta}-\theta_0) - \frac{L(L-1)}{L^2}\frac{2}{\sqrt{L}}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i}\sum_{p=1}^{i}\Bigg\{\sum_i \frac{1}{i(i-1)}N(Y_{p\ell},i)f_m^{-1}(X_\ell,i) \\
& g_0(Y_{p\ell},i) + \mathrm{E}\left[\sum_i \frac{1}{i(i-1)}N(Y_{p\ell},i)f_m^{-1}(X_\ell,i)g_0(Y_{p\ell},i)\right]\Bigg\} + o_p(1) \\
&= -\frac{1}{\sqrt{L}}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i}\sum_{p=1}^{i}\Bigg\{(C^T\Omega C)^{-1}C\Omega m(V_{p\ell},X_\ell,i;\theta_0) + 2\frac{L(L-1)}{L^2} \\
& \sum_i \frac{1}{i(i-1)}N(Y_{p\ell},i)f_m^{-1}(X_\ell,i)g_0(Y_{p\ell},i) \\
& -\mathrm{E}\left[\sum_i \frac{1}{i(i-1)}N(Y_{p\ell},i)f_m^{-1}(X_\ell,i)g_0(Y_{p\ell},i)\right]\Bigg\} + o_p(1)
\end{aligned}
$$

where $C = \mathrm{E}\left[\partial m(V,X,i;\theta_0)/\partial\theta\right]$, and

$N(Y_{p\ell},i) = [m_1(V_{p\ell},X_\ell,i;\theta_0)/g_0(B_{p\ell}|X_\ell,i)^2]G_0(B_{p\ell}|X_\ell,i).$

Proof

From Proposition 2 we have

$$\sqrt{L}(\hat{\theta}-\theta_0) - \sqrt{L}(\tilde{\theta}-\theta_0) = \frac{L(L-1)}{L^2}\frac{2}{\sqrt{L}}\sum_{\{\ell:I_\ell=i\}}^{L}\frac{1}{i}\sum_{p=1}^{i}[r_L(Y_{p\ell},i)-\theta_L]$$

where

$Y_{p\ell} = (B_{p\ell}, X_\ell)$

$r_L(Y_{p\ell},i) = \mathrm{E}[p_L((Y_{p\ell},i),(Y_{qj},i))|(Y_{p\ell},i)]$

$\theta_L = \mathrm{E}[r_L(Y_{p\ell},i)] = \mathrm{E}[p_L((Y_{p\ell},i),(Y_{qj},i))].$

First we show that $r_L(Y_{p\ell},i) = -\sum_i \frac{1}{i(i-1)}N(Y_{p\ell},i)f_m^{-1}(X_\ell,i)g_0(Y_{p\ell},i) + t_L(Y_{p\ell},i).$

We observe that

$$
\begin{aligned}
& r_L(Y_{p\ell}, i) = \mathrm{E}[p_L((Y_{p\ell}, i), (Y_{qj}, i))|(Y_{p\ell}, i)] \\
= & \left\{ \begin{array}{ll} \int p_L((B_{p\ell}, X_\ell, i), (B_{qj}, X_j, i)) g_0(B_{qj}, X_j, i) dY_{qj} & \text{if } \ell \neq j \\ \int p_L((B_{pj}, X_j, i), (B_{qj}, X_j, i)) g_0((B_{pj}, X_j, i), (B_{qj}, X_j, i)|(B_{pj}, X_j, i)) dY_{qj} & \text{if } \ell = j. \end{array} \right.
\end{aligned}
$$

We consider first the case $\ell \neq j$.

$$
\begin{aligned}
& r_L(Y_{p\ell}, i) = \\
= \quad & \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)}\Bigg\{\int \Bigg[\frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)}\frac{1}{h_G}\omega^G_{i,R+1,j} K_G\left(\frac{X_j - X_\ell}{h_G}\right) 1\!\!1(B_{qj} \leq B_{p\ell}) \\
& + \frac{m_1(V_{qj}, X_j, i; \theta_0)}{g_0(B_{qj}|X_j, i)}\frac{1}{h_G}\omega^G_{i,R+1,j} K_G\left(\frac{X_\ell - X_j}{h_G}\right) 1\!\!1(B_{p\ell} \leq B_{qj})\Bigg] g_0(Y_{qj}, i) dY_{qj}\Bigg\} \\
& -\frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)}\Bigg\{\int \Bigg[\frac{m_1(V_{p\ell}, X_\ell, i; \theta_0)}{g_0(B_{p\ell}|X_\ell, i)}\frac{G_0(B_{p\ell}|X_\ell, i)}{g_0(B_{p\ell}|X_\ell, i)}\frac{1}{h_g^2}\omega^g_{i,R,j} K_{1g}\left(\frac{X_j - X_\ell}{h_g}\right) \\
& K_{2g}\left(\frac{B_{qj} - B_{p\ell}}{h_g}\right) + \frac{m_1(V_{qj}, X_j, i; \theta_0)}{g_0(B_{qj}|X_j, i)}\frac{G_0(B_{qj}|X_j, i)}{g_0(B_{qj}|X_j, i)}\frac{1}{h_g^2}\omega^g_{i,R,j} K_{1g}\left(\frac{X_\ell - X_j}{h_g}\right) \\
& K_{2g}\left(\frac{B_{p\ell} - B_{qj}}{h_g}\right)\Bigg] g_0(Y_{qj}, i) dY_{qj}\Bigg\} \\
= \quad & \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)}\Bigg\{\int \Bigg[M(Y_{p\ell}, i)\frac{1}{h_G}\omega^G_{i,R+1,j} K_G\left(\frac{X_j - X_\ell}{h_G}\right) 1\!\!1(B_{qj} \leq B_{p\ell}) \\
& + M(Y_{qj}, i)\frac{1}{h_G}\omega^G_{i,R+1,j} K_G\left(\frac{X_\ell - X_j}{h_G}\right) 1\!\!1(B_{p\ell} \leq B_{qj})\Bigg] g_0(Y_{qj}, i) dY_{qj}\Bigg\} \\
& -\frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)}\Bigg\{\int \Bigg[N(Y_{p\ell}, i)\frac{1}{h_g^2}\omega^g_{i,R,j} K_{1g}\left(\frac{X_j - X_\ell}{h_g}\right) K_{2g}\left(\frac{B_{qj} - B_{p\ell}}{h_g}\right) \\
& + N(Y_{qj}, i)\frac{1}{h_g^2}\omega^g_{i,R,j} K_{1g}\left(\frac{X_\ell - X_j}{h_g}\right) K_{2g}\left(\frac{B_{p\ell} - B_{qj}}{h_g}\right)\Bigg] g_0(Y_{qj}, i) dY_{qj}\Bigg\} \\
= \quad & \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)}\int h_G\Big[M(Y_{p\ell}, i)\omega^G_{i,R+1,j} K_G(u_2) 1\!\!1(B_{p\ell} \leq u_1 h_G + B_{p\ell}) + M(uh_G + Y_{p\ell}, i) \\
& \omega^G_{i,R+1,j} K_G(-u_2) 1\!\!1(u_1 h_G + B_{p\ell} \leq B_{p\ell})\Big] g_0(uh_G + Y_{p\ell}, i) du \\
& -\frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)}\int \Big[N(Y_{p\ell}, i)\omega^g_{i,R,j} K_{1g}(u_2) K_{2g}(u_1) + N(uh_g + Y_{p\ell}, i) \\
& \omega^g_{i,R,j} K_{1g}(-u_2) K_{2g}(-u_1)\Big] g_0(uh_g + Y_{p\ell}, i) du.
\end{aligned}
$$

We note that as $h = (h_G, h_g) \to 0$ we have

$$\begin{aligned}
r_L(Y_{p\ell}, i) \quad \to \quad & -\frac{1}{2}\frac{1}{i}\sum_i \frac{1}{(i-1)} \int \Big[N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) K_{1g}(u)] K_{1g}(u_2) K_{2g}(u_1) \\
& + N(Y_{p\ell}, I_j) f_m^{-1}(X_\ell, i) K_{1g}(-u_2) K_{2g}(-u_1)\Big] g_0(Y_{p\ell}, i) du \\
& = -\frac{1}{i}\sum_i \frac{1}{(i-1)} N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i)
\end{aligned}$$

where we have used the following

$$\begin{aligned}
\omega_{i,R,j}^g \quad &= \quad e_1^T \left[\frac{1}{n_i h_g}\sum_{s=1}^{n_i} \mathbf{x_s}\mathbf{x_s}^T K_{1g}\left(\frac{X_s - X_\ell}{h_g}\right)\right]^{-1} [1 \quad (X_j - X_\ell) \ldots (X_j - X_\ell)^{R-1}]^T \\
&= \quad e_1^T \left[\frac{1}{n_i h_g}\sum_{s=1}^{n_i} \mathbf{x_s}\mathbf{x_s}^T K_{1g}\left(\frac{X_s - X_\ell}{h_g}\right)\right]^{-1} [1 \quad (-u_2 h_g) \ldots (-u_2 h_g)^{R-1}]^T \\
&\xrightarrow{p} \quad e_1^T \left[\mathrm{E}\left(\mathbf{x_s}\mathbf{x_s}^{\mathrm{T}} \mathrm{K_{1g}}\left(\frac{\mathrm{X_s} - \mathrm{X_\ell}}{\mathrm{h_g}}\right)\right)\right]^{-1} e_1 = f_m^{-1}(X_\ell, i)
\end{aligned}$$

therefore we define

$$r_L(Y_{p\ell}, I_j) = -\frac{1}{i}\sum_i \frac{1}{(i-1)} N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i) + t_L(Y_{p\ell}, i)$$

We consider now the reminder term $t_L(Y_{p\ell}, i)$

$$\begin{aligned}
t_L(Y_{p\ell}, i) \quad &= \quad r_L(Y_{p\ell}, i) + \frac{1}{i}\sum_i \frac{1}{(i-1)} N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i) \\
&= \quad \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)} \int h_G \Big[M(Y_{p\ell}, i)\omega_{i,R+1,j}^G K_G(u_2) \mathbb{1}(B_{p\ell} \le u_1 h_G + B_{p\ell}) \\
& \quad + M(u h_G + Y_{p\ell}, i)\omega_{i,R+1,j}^G K_G(-u_2) \mathbb{1}(u_1 h_G + B_{p\ell} \le B_{p\ell})\Big] g_0(u h_G + Y_{p\ell}, i) du \\
& \quad -\frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)} \int \Big[N(Y_{p\ell}, i)\omega_{i,R,j}^g K_{1g}(u_2) K_{2g}(u_1) + N(u h_g + Y_{p\ell}, i) \\
& \quad \omega_{i,R,j}^g K_{1g}(-u_2) K_{2g}(-u_1)\Big] g_0(u h_g + Y_{p\ell}, i) du \\
& \quad + \frac{1}{i}\sum_i \frac{1}{(i-1)} N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i)
\end{aligned}$$

Now, using $\int K_{1g}(u_2)K_{2g}(u_1)du = 1$ we can write

$$\begin{aligned}
&\frac{1}{i}\sum_i \frac{1}{(i-1)} N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i) \\
&= \frac{1}{2}\frac{1}{i}\sum_i \frac{1}{(i-1)} \int K_{1g}(u_2)K_{2g}(u_1) N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i) du \\
&\quad + \frac{1}{2}\frac{1}{i}\sum_i \frac{1}{(i-1)} \int K_{1g}(-u_2)K_{2g}(-u_1) N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i) du
\end{aligned}$$

therefore we can write the reminder term as follows

$$\begin{aligned}
t_L(Y_{p\ell}, I_j) &= \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)} \int h_G \Big[M(Y_{p\ell}, i)\omega^G_{i,R+1,j} K_G(u_2) 1\!\!1(B_{p\ell} \le u_1 h_G + B_{p\ell}) \\
&\quad + M(uh_G + Y_{p\ell}, i)\omega^G_{i,R+1,j} K_G(-u_2) 1\!\!1(u_1 h_G + B_{p\ell} \le B_{p\ell})\Big] g_0(uh_G + Y_{p\ell}, i) du \\
&\quad - \frac{1}{2}\sum_i \frac{1}{(i-1)} \int \frac{L}{iL_i} N(Y_{p\ell}, i)\omega^g_{i,R,j} K_{1g}(u_2)K_{2g}(u_1) g_0(uh_g + Y_{p\ell}, i) du \\
&\quad \frac{1}{2}\sum_i \frac{1}{(i-1)} \int \frac{L}{iL_i} N(uh_g + Y_{p\ell}, i)\omega^g_{i,R,j} K_{1g}(-u_2)K_{2g}(-u_1) g_0(uh_g + Y_{p\ell}, i) du \\
&\quad + \frac{1}{2}\sum_i \frac{1}{(i-1)} \int \frac{1}{i} N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) K_{1g}(u_2)K_{2g}(u_1) g_0(Y_{p\ell}, i) du \\
&\quad + \frac{1}{2}\sum_i \frac{1}{(i-1)} \int \frac{1}{i} N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) K_{1g}(-u_2)K_{2g}(-u_1) g_0(Y_{p\ell}, i) du \\
&= \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)} \int h_G \Big[M(Y_{p\ell}, i)\omega^G_{i,R+1,j} K_G(u_2) 1\!\!1(B_{p\ell} \le u_1 h_G + B_{p\ell}) \\
&\quad + M(uh_G + Y_{p\ell}, i)\omega^G_{i,R+1,j} K_G(-u_2) 1\!\!1(u_1 h_G + B_{p\ell} \le B_{p\ell})\Big] g_0(uh_G + Y_{p\ell}, i) du \\
&\quad - \frac{1}{2}\frac{1}{i}\sum_i \frac{1}{(i-1)} \int N(Y_{p\ell}, i) K_{1g}(u_2)K_{2g}(u_1) \\
&\quad \left[\frac{L}{L_i}\omega^g_{i,R,j} g_0(uh_g + Y_{p\ell}, i) - f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i)\right] du \\
&\quad - \frac{1}{2}\frac{1}{i}\sum_i \frac{1}{(i-1)} \int K_{1g}(-u_2)K_{2g}(-u_1) \\
&\quad \left[\frac{L}{L_i}\omega^g_{i,R,j} N(uh_g + Y_{p\ell}, i) g_0(uh_g + Y_{p\ell}, i) - f_m^{-1}(X_\ell, i) N(Y_{p\ell}, i) g_0(Y_{p\ell}, i)\right] du \\
&= \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)} \int h_G \Big[M(Y_{p\ell}, i)\omega^G_{i,R+1,j} K_G(u_2) 1\!\!1(B_{p\ell} \le u_1 h_G + B_{p\ell}) \\
&\quad + M(uh_G + Y_{p\ell}, i)\omega^G_{i,R+1,j} K_G(-u_2) 1\!\!1(u_1 h_G + B_{p\ell} \le B_{p\ell})\Big] g_0(uh_G + Y_{p\ell}, i) du \\
&\quad - \frac{1}{2}\frac{1}{i}\sum_i \frac{1}{(i-1)} \int N(Y_{p\ell}, i) K_{1g}(u_2)K_{2g}(u_1) \left[f_m^{-1}(X_\ell, i) + o_{as}(1)\right]
\end{aligned}$$

$$
\begin{aligned}
&[g_0(uh_g + Y_{p\ell}, i) - g_0(Y_{p\ell}, i)]du \\
&-\frac{1}{2}\frac{1}{i}\sum_i \frac{1}{(i-1)} \int K_{1g}(-u_2)K_{2g}(-u_1)\left[f_m^{-1}(X_\ell, i) + o_{as}(1)\right] \\
&[N(uh_g + Y_{p\ell}, i)g_0(uh_g + Y_{p\ell}, i) \\
&-N(Y_{p\ell}, i)g_0(Y_{p\ell}, i)]du
\end{aligned}
$$

thus using the above expression we have

$$
\begin{aligned}
&\frac{2}{\sqrt{L}} \sum_{\{\ell : I_\ell = i\}}^{L} \frac{1}{i} \sum_{p=1}^{i} [r_L(Y_{p\ell}, i) - \theta_L] = \\
= \quad & \frac{2}{\sqrt{L}} \sum_{\{\ell : I_\ell = i\}}^{L} \frac{1}{i} \sum_{p=1}^{i} \Bigg\{ - \frac{1}{i}\sum_i \frac{1}{(i-1)} N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i) \\
& + \mathrm{E}\Big[\frac{1}{i}\sum_i \frac{1}{(i-1)} N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i)\Big] + t_L(Y_{p\ell}, i) - \mathrm{E}[t_L(Y_{p\ell}, i)] \Bigg\} \\
= \quad & -\frac{2}{\sqrt{L}} \sum_{\{\ell : I_\ell = i\}}^{L} \frac{1}{i} \sum_{p=1}^{i} \Bigg\{ \frac{1}{i}\sum_i \frac{1}{(i-1)} N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i) \\
& - \mathrm{E}\Big[\frac{1}{i}\sum_i \frac{1}{(i-1)} N(Y_{p\ell}, i) f_m^{-1}(X_\ell, i) g_0(Y_{p\ell}, i)\Big] \Bigg\} \\
& + \frac{2}{\sqrt{L}} \sum_{\{\ell : I_\ell = i\}}^{L} \frac{1}{i} \sum_{p=1}^{i} \Big[t_L(Y_{p\ell}, i) - \mathrm{E}[t_L(Y_{p\ell}, i)] \Big]
\end{aligned}
$$

We denote the second term above by T_L and we observe that $\mathrm{E}[T_L] = 0$. We now show that $\mathrm{var}[T_L] = o_{as}(1)$.

$$
\begin{aligned}
\mathrm{var}[T_L] \quad &= \quad 4\frac{L_i}{L} \mathrm{var}\left[\frac{1}{i}\sum_{p=1}^{i} t_1(Y_{p1}, i)\right] \\
&= \quad 4\frac{L_i}{L} \mathrm{E}\left\{ \mathrm{var}\left[\frac{1}{i}\sum_{p=1}^{i} t_1(Y_{p1}, i)\middle| i\right]\right\} + 4\frac{L_i}{L} \mathrm{var}\left\{ \mathrm{E}\left[\frac{1}{i}\sum_{p=1}^{i} t_1(Y_{p1}, i)\middle| i\right]\right\} \\
&= \quad 4\frac{L_i}{L} \mathrm{E}\left\{ \frac{1}{i}\mathrm{var}\left[t_1(Y_{p1}, i)\middle| i\right]\right\} + 4\frac{L_i}{L} \mathrm{var}\Big\{ \mathrm{E}\left[t_1(Y_{p1}, i)|i\right]\Big\} \\
&= \quad A + B \qquad\qquad \text{(A.24)}
\end{aligned}
$$

We consider first the kth coordinate of the conditional variance inside the A term above, namely

$$\begin{aligned} \text{var}\left[t_{1_k}(Y_{p1},i)\Big| i\right] &\leq \text{E}\left[t_{1_k}(Y_{p1},i)^2\Big| i\right] \\ &\leq O\left(h_G^2\right)+O\left(h_g^{2(R-1)}\right) \end{aligned}$$

where the last inequality comes from observing that

$$\begin{aligned} t_{1_k}(Y_{p1},i) &= \frac{1}{2}\frac{L}{iL_i}\sum_i \frac{1}{(i-1)}\int h_G\Big[M(Y_{p1},i)\omega^G_{i,R+1,j}K_G(u_2)1\!\!1(B_{p1}\leq u_1h_G+B_{p1}) \\ &+M(uh_G+Y_{p1},i)\omega^G_{i,R+1,j}K_G(-u_2)1\!\!1(u_1h_G+B_{p1}\leq B_{p1})\Big]g_0(uh_G+Y_{p1},i)du \\ &-\frac{1}{2}\frac{1}{i}\sum_i \frac{1}{(i-1)}\int N(Y_{p1},i)K_{1g}(u_2)K_{2g}(u_1)f_m^{-1}(X_1,i) \\ &[g_0(uh_g+Y_{p1},i)-g_0(Y_{p1},i)]du \\ &-\frac{1}{2}\frac{1}{i}\sum_i \frac{1}{(i-1)}\int K_{1g}(-u_2)K_{2g}(-u_1)f_m^{-1}(X_1,i)[N(uh_g+Y_{p1},i) \\ &g_0(uh_g+Y_{p1},i)-N(Y_{p1},i)g_0(Y_{p1},i)]du+o_{as}(1) \\ &= a+b+c+o_{as}(1) \end{aligned}$$

therefore, twice application of $(a+b)^2\leq 2(a^2+b^2)$ yields $(a+b+c)^2\leq\kappa(a^2+b^2+c^2)$, thus

$$\begin{aligned} \text{E}\left[t_{1_k}(Y_{p1},i)^2\Big| i\right] &= \text{E}[(a+b+c)^2|i]+o_{as}(1) \\ &\leq 4\text{E}[a^2+b^2+c^2|i]+o_{as}(1) \\ &= O\left(h_G^2\right)+O\left(h_g^{2(R-1)}\right)+O\left(h_g^{2(R-1)}\right) \end{aligned}$$

where the order of the last two terms after the last equality above follows from an $(R-1)$th Taylor Expansion around Y_{p1} and the fact that kernels are of order $R-1$ by A.3-(iii).

We consider now the B term in (A.24) and more precisely we consider the following

$$\begin{aligned} \frac{B}{4} &= \frac{L_i}{L}\text{var}\Big\{\text{E}\left[t_1(Y_{p1},i)|i\right]\Big\} \\ &\leq \text{E}\Big\{\text{E}\left[t_1(Y_{p1},i)|i\right]^2\Big\} \\ &\leq \text{E}\Big\{\text{E}\left[t_1(Y_{p1},i)^2|i\right]\Big\} \\ &\leq O\left(h_G^2\right)+O\left(h_g^{2(R-1)}\right)+O\left(h_g^{2(R-1)}\right) \end{aligned}$$

where the last inequality follows from the same argument used above. Hence, by Chebyshev Inequality $T_L = o_p(1)$.

We consider now the case $\ell = j$ and observe the following

$$\begin{aligned}
& r_L(Y_{pj}, i) \\
= \;& \mathrm{E}[p_L((B_{pj}, X_j, i), (B_{qj}, X_j, i))|(B_{pj}, X_j, i)] \\
= \;& \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)} \int \frac{1}{h_G}\Big[M(Y_{pj}, i)\omega^G_{i,R+1,j} K_G(0) \mathbb{1}(B_{qj} \le B_{pj}) \\
& + M(Y_{qj}, i)\omega^G_{i,R+1,j} K_G(0) \mathbb{1}(B_{pj} \le B_{qj})\Big] g_0((Y_{pj}, i), (Y_{qj}, i)|(Y_{pj}, i)) dY_{qj} \\
& - \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)} \int \frac{1}{h_g^2}\Big[N(Y_{pj}, i)\omega^g_{i,R,j} K_{1g}(0) K_{2g}\left(\frac{B_{qj} - B_{pj}}{h_g}\right) \\
& + N(Y_{qj}, i)\omega^g_{i,R,j} K_{1g}(0) K_{2g}\left(\frac{B_{pj} - B_{qj}}{h_g}\right)\Big] g_0((Y_{pj}, i), (Y_{qj}, i)|(Y_{pj}, i)) dY_{qj}
\end{aligned}$$

We now use the change of variable $u = (Y_{qj} - Y_{pj})/h_G$ and $\tilde{u} = (Y_{qj} - Y_{pj})/h_g$ to obtain

$$\begin{aligned}
& r_L(Y_{pj}, i) \\
= \;& \frac{1}{2}\frac{L}{n_i}\sum_i \frac{1}{(i-1)} \int h_G\Big[M(Y_{pj}, i)\omega^G_{i,R+1,j} K_G(0) \mathbb{1}(u_1 h_G + B_{pj} \le B_{pj}) + M(uh_G + Y_{pj}, i) \\
& \omega^G_{i,R+1,j} K_G(0) \mathbb{1}(B_{pj} \le u_1 h_G + B_{pj})\Big] g_0((Y_{pj}, i), (uh_G + Y_{pj}, i)|(Y_{pj}, i)) du \\
& - \frac{1}{2}\frac{L}{iL_i}\sum_i \frac{1}{(i-1)} \int \Big[N(Y_{pj}, i)\omega^g_{i,R,j} K_{1g}(0) K_{2g}(\tilde{u}_1) + N(\tilde{u} h_g + Y_{pj}, i) \\
& \omega^g_{i,R,j} K_{1g}(0) K_{2g}(-\tilde{u}_1)\Big] g_0((Y_{pj}, i), (uh_g + Y_{pj}, i)|(Y_{pj}, i)) d\tilde{u}
\end{aligned}$$

next, we observe that as $h = (h_G, h_g) \to 0$ we have

$$\begin{aligned}
r_L(Y_{pj}, i) \;\to\; & -\frac{1}{2}\frac{1}{i}\sum_i \frac{1}{(i-1)} \int \Big[N(Y_{pj}, i) f_m^{-1}(X_j, i) K_{1g}(0) K_{2g}(\tilde{u}_1) \\
& + N(Y_{pj}, i) f_m^{-1}(X_j, i) K_{1g}(0) K_{2g}(-\tilde{u}_1)\Big] g_0((Y_{pj}, i), (Y_{pj}, i)|(Y_{pj}, i)) d\tilde{u} \\
= \;& -\frac{1}{i}\sum_i \frac{1}{(i-1)} N(Y_{pj}, i) f_m^{-1}(X_j, i) g_0(Y_{pj}, i)
\end{aligned}$$

as before we define

$$r_L(Y_{pj}, i) = -\frac{1}{i}\sum_{i} \frac{1}{(i-1)} N(Y_{pj}, i) f_m^{-1}(X_j, i) g_0(Y_{pj}, i) + t_L(Y_{pj}, i)$$

The rest of the proof is analogous to the one for the case $\ell \neq j$. Therefore the desired result follows. *Q.E.D.*

References

Ai, C. and **X. Chen** (2003): "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions," *Econometrica*, 71, 1795-1843.

Campo, S., **E. Guerre**, **I. Perrigne** and **Q. Vuong** (2006): "Semiparametric Estimation of First-Price Auctions with Risk Averse Bidders," *Review of Economic Studies* ,78, 1, 112-147.

Campo, S., **I. Perrigne** and **Q. Vuong** (1998): "Asymmetry and Joint Bidding in OCS Wildcat Auctions," mimeo, University of Southern California.

Donald, S. and **H. Paarsch** (1993): "Piecewise Pseudo-Maximum Likelihood Estimation in Empirical Models of Auctions," *International Economic Review*, 34, 121-148.

Elyakime, B., **J.J Laffont**, **P. Loisel**, and **Q. Vuong** (1994): "First-Price Sealed-Bid Auctions with Secret Reservation Prices," *Annales d'Economie et de Statistiques*, 34, 115-141.

Fan, J. and **I. Gijbels** (1996): *Local Polynomial Modelling and Its Applications*, Chapman & Hall.

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

Hardle, W. (1991): *Smoothing Techniques with Implementation in S*, Springer Verlag.

Honoré, B.E. and **E. Kryriazidou** (2000): "Panel Data Discrete Choice Models with Lagged Dependent Variable," *Econometrica*, 68, 839-874.

Horowitz, J.L. (2000): "A Smoothed Maximum Score Estimator for the Binary Response Model," *Econometrica*, 60, 505-531.

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

Korostelev, A.P. and **A.B. Tsybakov** (1993): "Minimax Theory of Image Reconstruction," *Lecture Notes in Statistics*, 82.

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

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

Laffont, J.J., **T. Li** and **Q. Vuong** (1999): "Identification of First-Price Auctions with a Large Trader," mimeo, University of Southern California.

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

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

Lavergne, P. and **Q. Vuong** (1996): "Nonparametric Selection of Regressors: The Nonnested Case," *Econometrica*, 64, 207-219.

Li, T. and **I. Perrigne** (2003): "Timber Sale Auctions with Random Reserve Prices," *Review of Economics and Statistics*, 85, 189-200.

Li, T., **I. Perrigne** and **Q. Vuong** (1999): "Structural Estimation of Affiliated Private Values with an Application to OCS Auctions," mimeo, University of Southern California.

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

McAfee, R.P. and **J. McMillan** (1987): "Auction and Bidding," *Journal of Economic Literature,* 25, 699-738.

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

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

Paarsch, H.J. and **H. Hong** (2006): *An Introduction to the Structural Econometrics of Auction Data*, MIT Press.

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

Perrigne, I. and **Q. Vuong** (2008): "Auctions: Empirics," in *The New Palgrave Dictionary of Economics*, Second Edition, S. Durlauf and L. Blume, eds., Palgrave McMillan.

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

Powell, J.L., **J.H. Stock** and **T. M. Stoker** (1989): "Semiparametric Estimation of Index Coefficients," *Econometrica*, 57, 1403-1430.

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

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

Wilson, R. (1992): "Strategic analysis of auctions," in R.J. Aumann and S. Hart, eds. *Handbook of Game Theory with Economic Applications, Volume I*, New York: North Holland.