February, 2014

Keywords and Phrases. Instrumental variables, semiparametric regression, efficient estimation.

JEL Classifications. C14, C21
Appendix 2: Proof of Lemma 1, Theorems 1 and 3.

Throughout the proof, we use $c$ or $C$ to denote some fixed constants.

Lemma 1 Define

$$S_{n,j}(z) = \frac{1}{n} \sum_{i=1}^{n} K_{h}(Z_{i}^{c} - z^{c}) \left( \frac{Z_{i}^{c} - z^{c}}{h} \right)^{j} I(Z_{i}^{d} = z^{d}) g(U_{i}) w(Z_{i}^{c} - z^{c}; z), |j| = 0, 1, 2, \cdots, J,$$

where $Z_{i}, U_{i}$ are iid, $Z_{i}^{c} \in R^{l}, Z_{i}^{d} \in R^{d}, K_{h}(z^{c}) = \frac{1}{n} c_{h} K(\frac{z^{c}}{n})$, and $K(\cdot)$ is a kernel function defined on $R^{l}$. If we have

$L_{1}.$ $K(\cdot)$ is bounded with compact support and for Euclidean norm $||.||$,

$$|w^{l} K(u) - v^{l} K(v)| \leq c_{K} ||u - v||, \ for \ 0 \leq |j| \leq J.$$

$L_{2}.$ $g(u)$ is a measurable function of $u_{i}$ and $E|g(u)|^{s} < \infty \ for \ s > 2.$

$L_{3}.$ $\sup_{z \in G} \int \{|g(u)|^{s} f_{z,u}(z, u) du < \infty, f_{z,u}(z, u) < \infty, \ and \ f_{z,u}(z, u) \ is \ continuous \ around \ z^{c}.$

$L_{4}.$ $\sup_{z \in G} |w(Z_{i}^{c} - z^{c}; z)| < \infty, \forall z^{d} \in G^{d}, \ a \ compact \ subset \ of \ R^{d}, |w(Z_{i}^{c} - z^{c}; z^{c}, z^{d}) - w(Z_{i}^{c} - z_{k}^{c}; z^{d})| < c_{d} |z^{c} - z_{k}^{c}|.$

$L_{5}.$ $nh^{l} \to \infty.$

Then for $z = (z^{c}, z^{d}) \in G = G^{c} \times G^{d}, \ z^{c} \in G^{c}, \ a \ compact \ subset \ of \ R^{l},$

$$\sup_{z \in G} |S_{n,j}(z) - E(S_{n,j}(z))| = O_{P} \left( \left( \frac{nh^{l}}{ln(n)} \right)^{\frac{1}{2}} \right).$$

Proof. Let's define

$$S_{n,j}^{B}(x) = \frac{1}{n} \sum_{i=1}^{n} K_{h}(Z_{i}^{c} - z^{c}) \left( \frac{Z_{i}^{c} - z^{c}}{h} \right)^{j} I(Z_{i}^{d} = z^{d}) g(U_{i}) w(Z_{i}^{c} - z^{c}; z) I(|g(U_{i})| \leq B_{n}),$$

where $B_{1} \leq B_{2} \leq \cdots \ such \ that \ \sum_{i=1}^{\infty} B_{i}^{-s} < \infty \ for \ some \ s > 0$. Since $G^{c} \times G^{d}$ is compact, we could cover $G$ by a finite number $l_{n}$ of $l_{n}$ dimensional cubes $I_{k}$ with center $z_{k}, k = 1, 2, \cdots, l_{n}$ and length $r_{n}$. We could choose $l_{n}$ sufficiently large such that $r_{n}$ is sufficiently small and each cube $I_{k}$ corresponds to one fixed possible value of $z^{d}, \ i.e., \ z^{d} = z_{k}^{d}$ if $z \in I_{k}$. Since $G$ is compact, $l_{n}r_{n} = c,$ $c$ a constant. Suppose we let $l_{n} = \left( \frac{ln(n)}{ln(n) + c \ v_{n}} \right)^{\frac{1}{m}}$, then $r_{n} = c^{\frac{1}{m}}$. Since

$$\sup_{z \in G} |S_{n,j}^{B}(z) - E(S_{n,j}^{B}(z))| = \max_{1 \leq k \leq l_{n}} \sup_{z \in I_{k}} |S_{n,j}^{B}(z) - E(S_{n,j}^{B}(z))| = \max_{1 \leq k \leq l_{n}} \sup_{z \in I_{k}} |S_{n,j}^{B}(z^{c}, z^{d}) - E(S_{n,j}^{B}(z^{c}, z^{d}))|$$

$$\leq \max_{1 \leq k \leq l_{n}} \sup_{z \in I_{k}} |S_{n,j}^{B}(z^{c}, z_{k}^{d}) - S_{n,j}^{B}(z^{c}, z_{k}^{d})| + \max_{1 \leq k \leq l_{n}} \sup_{z \in I_{k}} |E(S_{n,j}^{B}(z^{c}, z_{k}^{d})) - E(S_{n,j}^{B}(z^{c}, z_{k}^{d}))|$$

$$I_{1} + I_{2} + I_{3}$$

The lemma is proved if we can show

(1) $I_{0} = \sup_{z \in G} |S_{n,j}(z) - E(S_{n,j}(z)) - [S_{n,j}^{B}(z) - E(S_{n,j}^{B}(z))]| = O_{a.s.}(B_{1}^{-s})$ for $B_{1}^{-s} = O\left( \left( \frac{ln(n)}{ln(n)} \right)^{\frac{1}{m}} \right),$

(2) $I_{1} = O_{a.s.}(\left( \frac{ln(n)}{ln(n)} \right)^{\frac{1}{m}}).$

(3) $I_{2} = O_{P}(\left( \frac{ln(n)}{ln(n)} \right)^{\frac{1}{m}})$.

(4) $I_{3} = O_{a.s.}(\left( \frac{ln(n)}{ln(n)} \right)^{\frac{1}{m}}).$
\[ I_{01} = \sup_{z \in G} \left| \frac{1}{n} \sum_{i=1}^{n} K_h(Z_i^z - z^c) \left( \frac{Z_i^z - z^c}{h} \right)^3 I(Z_i^d = z^d)g(U_i)w(Z_i)I(|g(U_i)| > B_n) \right| . \]

By Chebychev’s inequality, \( \sum_{i=1}^{\infty} P(|g(U_i)| > B_i) \leq \sum_{i=1}^{\infty} \frac{E(|g(U_i)|^r)}{B_i^r} < e \sum_{i=1}^{\infty} B_i^{-s} \leq \infty, \) by construction of \( B_i \) and \( L_2. \) By Borel-Cantelli Lemma, \( P(|g(U_i)| > B_i \ i.o.) = 0. \) To see this, \( P(|g(U_i)| > B_i \ i.o.) = \lim_{i \to \infty} P(\bigcup_{m=i}^{\infty} \{\omega : |g(U_m)| > B_m\}) = \lim_{i \to \infty} \sum_{m=i}^{\infty} P(\{\omega : |g(U_m)| > B_m\}) = 0 \) since \( \sum_{i=1}^{\infty} P(\{\omega : |g(U_i)| > B_i\}) < \infty. \) So \( \forall \epsilon > 0, \) there exists \( i' > \epsilon \) such that \( \forall i > i', \)

\[ P(\bigcup_{m=i}^{\infty} \{\omega : |g(U_m)| > B_m\}) < \epsilon, \quad \text{or} \quad P(\bigcap_{m=i}^{\infty} \{\omega : |g(U_m)| \leq B_m\}) > 1 - \epsilon. \]

So \( \forall m > i', \) \( P(|g(U_m)| \leq B_m) > 1 - \epsilon \) or \( |g(U_m)| \leq B_m \) for sufficiently large \( m. \) Since \( B_i \) is an increasing sequence, w.p.1, \( |g(U_m)| \leq B_n \) for \( m \geq i' \) and \( n \geq m. \)

When \( i = \{1, 2, \ldots, i'\}, \) \( P(|g(U_i)| > B_n) > 1 - \epsilon. \) To see this, \( \forall \epsilon > 0, \) and sufficiently large \( n, \)

\( P(|g(U_i)| > B_n) < \frac{E(|g(U_i)|^r)}{B_n^r} < \frac{1}{B_n^r} < \epsilon, \) since \( E|g(U_i)|^r < \infty \) and \( B_i \) is an increasing sequence. So in all, \( \forall \epsilon > 0, \) and for \( n \) sufficiently large, we have \( I(|g(U_i)| > B_n) = 0 \) w.p.1. So \( I_{01} = 0 \) a.s.

\[ I_{02} = \sup_{z \in G} \left\{ EK_h(\sum_{c} Z_i^c - z^c) \left( \frac{Z_i^c - z^c}{h} \right)^3 I(\sum_{c} Z_i^c = z^c)g(U_i)w(Z_i)I(|g(U_i)| > B_n) \right\} \]

By variable of change,

\[ I_{02} = \sup_{z \in G} \left\{ \sum_{c} K(\sum_{c} Z_i^c = z^c) \left( \frac{Z_i^c - z^c}{h} \right)^3 I(\sum_{c} Z_i^c = z^c)g(U_i)w(Z_i)I(|g(U_i)| > B_n) \right\} \]

where to obtain the first inequality we use \( L_4, \) the second we use \( L_1 \) and Hölder’s inequality, the third and fourth we use \( L_3. \) The last line above we use Chebychev’s inequality again and \( L_2. \)

\[ S_{n,j}^{B}(z^c, z_k^c) = S_{n,j}^{B}(z_k^c) \]

\[ = \frac{1}{n^{1+r}} \sum_{h=1}^{n} \sum_{i=1}^{n} \left\{ K_h(Z_i^c - z^c) \left( \frac{Z_i^c - z^c}{h} \right)^3 w(Z_i^c - z^c, z_k^c) \right\} \]

\[ -\left( \frac{Z_i^c - z^c}{h} \right)^3 w(Z_i^c - z^c, z_k^c)I(\sum_{c} Z_i^c = Z_k^c)g(U_i)I(|g(U_i)| \leq B_n) \]

\[ \leq \frac{1}{n^{1+r}} \sum_{h=1}^{n} \left\{ \left( \frac{Z_i^c - z^c}{h} \right)^3 - K_h(Z_i^c - z^c) \left( \frac{Z_i^c - z^c}{h} \right)^3 \right\} \frac{w(Z_i^c - z^c, z_k^c)}{} \]

\[ \leq \frac{1}{n^{1+r}} \sum_{h=1}^{n} \left\{ \frac{1}{h} \sum_{i=1}^{n} [K(Z_i^c - z^c) \left( \frac{Z_i^c - z^c}{h} \right)^3] w(Z_i^c - z^c, z_k^c)I(\sum_{c} Z_i^c = Z_k^c)g(U_i)I(|g(U_i)| \leq B_n) \right\} \]

\[ \leq \frac{1}{n^{1+r}} \sum_{h=1}^{n} \left\{ c[\frac{Z_i^c - z^c}{h}] + c[Z_i^c - z^c] \right\} |g(U_i)| I_1 \text{ and } L_4, \]

since \( z \in I_k \) for some \( k, \|Z_i^c - z^c\| \leq c r_n \) and with \( L_4, \)

\[ I_1 \leq \frac{c}{n^{1+r}} \sum_{i=1}^{n} |g(U_i)|, \] by \( L_2 \) and Kolmogorov’s Theorem,

\[ \frac{1}{n^{1+r}} \sum_{i=1}^{n} |g(U_i)|^{a.s.} E|g(U_i)| < \infty. \]

So \( I_1 \leq \frac{c}{n^{1+r}} \sum_{i=1}^{n} |g(U_i)|^{a.s.} E|g(U_i)| < \infty. \)
We could show (4) \(I_3 = O_{a.s.}\left(\frac{(\ln n)^{\frac{1}{2}}}{\sqrt{n}}\right)^{\frac{1}{2}}\) similarly.

(3) It is sufficient to show \(\exists\) a constant \(\Delta > 0\) and \(N > 0\) such that \(\forall \varepsilon > 0\) and \(n > N\),
\[
P\left(\frac{(\ln n)^{\frac{1}{2}}}{\sqrt{n}} I_2 \geq \Delta\right) < \varepsilon.
\]

Let \(\epsilon_n = \frac{(\ln n)^{\frac{1}{2}}}{\sqrt{n}} \Delta\), then \(P(I_2 \geq \epsilon_n) \leq \sum_{k=1}^{n} P(|S_{n,j}^B(z_k) - E S_{n,j}^B(z_k)| \geq \epsilon_n).
\]
\[
|S_{n,j}^B(z_k) - E S_{n,j}^B(z_k)| = \left| \frac{1}{n} \sum_{i=1}^{n} W_{in} \right|.
\]
\[
= \left| \frac{1}{n} \sum_{i=1}^{n} |\frac{n}{n} K(Z_t^i - z_k^i) \left( \frac{Z_t^i - z_k^i}{2} \right)| I(Z_t^i = z_k^i) g(U_i) w(Z_t^i - z_k^i; z_k) I(|g(U_i)| \leq B_n) \right|
\]
\[
- \left| \frac{1}{n} E K(Z_t^i - z_k^i) \left( \frac{Z_t^i - z_k^i}{2} \right)^2 I(Z_t^i = z_k^i) g(U_i) w(Z_t^i - z_k^i; z_k) I(|g(U_i)| \leq B_n) \right|.
\]

Since \(E W_{in} = 0\), \(|W_{in}| \leq 2e^{\frac{B_n}{2}}\) by \(L_1\) and \(L_4\), and \(\{W_{in}\}_{i=1}^{n}\) is an independent sequence, by Bernstein’s inequality,
\[
P(|S_{n,j}^B(z_k) - E S_{n,j}^B(z_k)| \geq \epsilon_n) \leq 2 \exp \left( -\frac{n \epsilon_n^2}{2E n + 2B_n \epsilon_n} \right),
\]
where \(\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} V(W_{in}) = E W_{in}^2 = I_{21} - I_{22}^2\)
\[
= \frac{1}{n} \sum_{i=1}^{n} K Z_t^i - z_k^i)^2 \left( \frac{Z_t^i - z_k^i}{2} \right)^2 I(Z_t^i = z_k^i) g(U_i) w(Z_t^i - z_k^i; z_k) I(|g(U_i)| \leq B_n) \right|
\]
\[
- \left| \frac{1}{n} E K(Z_t^i - z_k^i) \left( \frac{Z_t^i - z_k^i}{2} \right)^2 I(Z_t^i = z_k^i) g(U_i) w(Z_t^i - z_k^i; z_k) I(|g(U_i)| \leq B_n) \right|^2.
\]

\(I_{22} = \sum_{i=1}^{n} |f Z_t^i - z_k^i | K |f Z_t^i - z_k^i | I(|g(U_i)| \leq B_n) f Z_t^i - z_k^i | h \Psi_i, Z_t^i, U_i | d \psi_i d U_i \)
\[
\leq e \int |K |f Z_t^i - z_k^i | g(U_i) | h \Psi_i, Z_t^i, U_i | d \psi_i d U_i \rightarrow e \int |K |f Z_t^i - z_k^i | g(U_i) | h \Psi_i, Z_t^i, U_i | d \psi_i d U_i \rightarrow 0.
\]

Similarly \(h \epsilon_n I_{21} = O(1)\). So \(2h \epsilon_n \sigma^2 < \infty\). If \(B_n \epsilon_n \rightarrow \infty\), then \(C_n = 2h \epsilon_n \sigma^2 + 2B_n \epsilon_n < \infty\), then
\[
P(I_2 \geq \epsilon_n) \leq 2 \exp \left( -\frac{2n \epsilon_n^2}{2h \epsilon_n \sigma^2 + 2B_n \epsilon_n} \right) = 2 \exp \left( -\frac{2n \epsilon_n^2}{2h \epsilon_n \sigma^2 + 2B_n \epsilon_n} \right) \rightarrow 0.
\]

Above is true since \(C_n < \infty\), if we let \(\Delta^2 \geq C_n(1 + \epsilon_i)\), then
\[
2 \exp \left( -\frac{2n \epsilon_n^2}{2h \epsilon_n \sigma^2 + 2B_n \epsilon_n} \right) \leq 2 \exp \left( -\frac{2n \epsilon_n^2}{2h \epsilon_n \sigma^2 + 2B_n \epsilon_n} \right) \rightarrow 0
\]
by \(L_5\).

If we let \(B_n = n^{-\delta + \delta/2}\) for \(s > 2\) and \(\delta > 0\), then \(B_n \epsilon_n \rightarrow \infty\) for sufficiently large \(s\). To see this, \(B_n \epsilon_n = n^{-\delta + \delta/2} \frac{(\ln n)^{\frac{1}{2}}}{\sqrt{n}} \Delta\). By \(L_5\), we could let \((h \epsilon_n)^{-\frac{1}{2}} = n^{-\delta + \delta/2}\), for \(\frac{1}{2} > \delta \geq 0\), then \(B_n \epsilon_n = n^{-\delta + \delta/2} \frac{(\ln n)^{\frac{1}{2}}}{\sqrt{n}} \Delta\). If we let \(s > \frac{1}{2} + (\delta - \delta_1)^{-1}\), then \(B_n \epsilon_n \rightarrow 0\).

It is easy to see that for \(B_n = n^{-\delta + \delta/2}\), we easily have \(\sum_{i=1}^{\infty} B_i^{-s} < \infty\). Furthermore \(B_n^{-s} < n^{-\frac{s}{2}}\), so \(B_n^{-s} = O \left( \frac{(\ln n)^{\frac{1}{2}}}{\sqrt{n}} \right)^{\frac{s}{2}}\).

**Theorem 1:** Proof. Note \(Y_t - \hat{E}(Y|X_t) = (X_t - \hat{E}(X|Z_t)) + m(Z_t) - \hat{E}(m(z_t)|Z_t) + \epsilon_t - \hat{E}(\epsilon|Z_t)\),
so we could write
\[
\hat{\beta} - \beta = \left[ \frac{n}{n} \hat{W} X - (E W_t X) - (E W_t X)^{-1} \right]
\]
\[
\times \frac{1}{n} \hat{W} (m - \hat{E}(m(z_t)|Z_t) + \hat{E}(\epsilon|Z_t)) = \left[ \frac{n}{n} \hat{W} (m - \hat{E}(m(z_t)|Z_t) + \hat{E}(\epsilon|Z_t)) \right]^C.
\]

where \(\hat{m} = \{m(Z_t)\}_{t=1}^{n}, \hat{E}(m(z_t)|Z_t) = \{ \hat{E}(m(z_t)|Z_t) \}_{t=1}^{n}, \hat{\epsilon} = \{ \epsilon_t \}_{t=1}^{n}, \hat{E}(\epsilon|Z_t) = \{ \hat{E}(\epsilon|Z_t) \}_{t=1}^{n}.
\]

Let’s denote \(E X_t Z_t = \hat{g}(Z_t)\) and \(E X_t Z_t = \hat{g}_1(Z_t)\), then \(W_{t,k} = \hat{g}(Z_t) - g_k(Z_t) + g_{1,k}(Z_t) - g_{1,k}(Z_t) + g_k(Z_t) - g_{1,k}(Z_t)\). The \((i,j)\)th element of \(\frac{n}{n} \hat{W} X\) is
\[
\frac{1}{n} \sum_{i=1}^{n} \hat{W}_{i,t} X_{i,j} = \frac{1}{n} \sum_{i=1}^{n} \hat{W}_{i,t} (X_{i,j} - \hat{g}_j(Z_t)) + \frac{1}{n} \sum_{i=1}^{n} W_{i,t} \hat{W}_{i,j}.
\]

3
\[
\frac{1}{n} \sum_{i=1}^{n} \tilde{W}_{i,j} \tilde{W}_{i,j} = \frac{1}{n} \sum_{i=1}^{n} [g_i(Z_t) - \hat{g}_i(Z_t)] [g_j(Z_t) - \hat{g}_j(Z_t)] + \frac{1}{n} \sum_{i=1}^{n} [\hat{g}_i(Z_t) - g_i(Z_t)] [g_j(Z_t) - \hat{g}_j(Z_t)] + \frac{n}{n} \sum_{i=1}^{n} [\hat{g}_i(Z_t) - g_i(Z_t)] [\hat{g}_j(Z_t) - g_j(Z_t)]
\]

Similarly, for \( k = 1, 2, \ldots, K \), the \( k \)th element of \( C \) is

\[
C_k = \frac{1}{n} \sum_{i=1}^{n} \tilde{W}_{i,k} (m(Z_t) - \hat{E}(m(Z_t))) = \frac{1}{n} \sum_{i=1}^{n} [g_k(Z_t) - \hat{g}_k(Z_t)] [m(Z_t) - \hat{E}(m(Z_t))]
\]

We show below (1) \( A_i = o_p(1), \ i = 1, \ldots, 8, 10, 11, 12, \)

\[
A_9 = E[g_i(Z_t) - g_{i,j}(Z_t)] [g_j(Z_t) - g_{j,i}(Z_t)] = o_p(1),
\]

so together we have \( \frac{1}{n} \hat{W}' \hat{X} \) and \( E W' W = o_p(1) \). By A1(3) and Slutsky’ Theorem, \( (\frac{1}{n} \hat{W}' \hat{X})^{-1} - (E W' W)^{-1} = o_p(1). \)

(2) Denote \( \ln z = (\frac{\ln(n)}{n h_{1e}^*})^{-1/2} + h_{1e}^*, \ln z = (\frac{\ln(n)}{n h_{1e}^*})^{1/2} + h^*. \) We show that

\[
C_{1K} = O_p(\ln z * \ln z), \ C_{2K} = O_p(\ln z^2), \ C_{3K} = o_p(n^{-1/2}), \ C_{4K} = o_p(n^{-1/2}) + O_p(\ln z * (\frac{\ln(n)}{n h_{1e}^*})^{1/2}),
\]

\[
C_{5K} = o_p(n^{-1/2}) + O_p(\ln z * (\frac{\ln(n)}{n h_{1e}^*})^{1/2}).
\]

For \( C_6 = [C_{61}, C_{62}, \ldots, C_{6K}]' \), \( \sqrt{n} C_6 \) \( \stackrel{d}{\rightarrow} N(0, \Phi_0) \), where \( \Phi_0 \) is defined in Theorem 1. With A5, \( nh_{1e}^* \rightarrow \infty, nh_{1e}^{2/3} \rightarrow \infty, nh_{4e}^* \rightarrow 0 \) and \( nh_{4e}^{1/3} \rightarrow 0 \). It implies that \( C_{1K} + \ldots + C_{5K} = o_p(n^{-1/2}) + O_p(\ln z^2 + \ln z^2) = o_p(n^{-1/2}). \)

Combining results in (1) and (2) and using A1(3), we conclude

\[
\sqrt{n} \hat{b} - \beta \frac{d}{d} N(0, (E W' W)^{-1} \Phi_0 (E W' W)^{-1}).
\]

(1) (a) Define \( \hat{f}_1(z_{10}) = \frac{1}{n h_{1e}^*} \sum_{i=1}^{n} K_{11}(Z_{t_i} - z_{10}) \). We first show \( \sup_{z_{10} \in G_1} |\hat{f}_1(z_{10}) - f_1(z_{10})| = O_p(\ln z_{10}) \). We apply Lemma 1 with \( S_n,0(z_{10}) = \hat{f}_1(z_{10}) \), so

\[
\sup_{z_{10} \in G_1} |\hat{f}_1(z_{10}) - E\hat{f}_1(z_{10})| = O_p(\frac{n h_{1e}^*}{\ln(n)} - \frac{1}{2}).
\]

Condition \( L_1 \) is satisfied with A3, \( L_2 \) is satisfied since \( g(u) = 1, L_3 \) is true with A2(1) and (2), \( L_4 \) is satisfied since \( w(z) = 1 \). Since the data are iid in A1(1),

4
\( \tilde{E}(z_{10}) = \int \frac{1}{h_{1c}} K_{1}(\tilde{Z}_{i1}^{t} - \tilde{z}_{10}^{c}) f_{i}(Z_{i1}^{c}, z_{10}^{d}) dZ_{i1}^{c}, \) with \( \Psi_{i} = \tilde{Z}_{i1}^{t} - \tilde{z}_{10}^{c} \)

\[ = \int K_{1}(\Psi_{i}) f_{i}(z_{10}^{d}, h_{1c}) + \frac{1}{h_{1c}} \Psi_{i} d\Psi_{i}, \text{ with } A2(1) \]

\[ = \int K_{1}(\Psi_{i}) f_{i}(z_{10}^{d}, h_{1c}) + \frac{1}{h_{1c}} \Psi_{i} d\Psi_{i}, \text{ with } A2(1) \]

\[ = f_{1}(z_{10}) + O(h_{1}^{s}) \text{ uniformly } \forall z_{10} \in G_{1} \text{ by A3, A2(1) and Dominated Convergence Theorem, where } z_{10}^{c} \text{ is between } z_{10}^{d} \text{ and } Z_{i1}. \]

So for any \( y_{i} \), we have

\[ \text{Sup}_{\tilde{z}_{10} \in G} |\tilde{f}(z_{10}) - f_{i}(z_{10})| = O(p(ln z)) \text{ with } A2(1), (5) \text{ and A3.} \]

(c) Define \( S_{1n}(z_{10}) = S_{1n0}(z_{10}) \), where \( s_{1n0}(z_{10}) = \sum_{i=1}^{n} K_{1}(Z_{i1}^{c} - z_{10}^{c}) \)

\[ = \sum_{i=1}^{n} K_{1}(Z_{i1}^{c} - z_{10}^{c}) \]

Let \( W_{1} = \sum_{i=1}^{n} |K_{1}(Z_{i1}^{c} - z_{10}^{c})| \)

\[ = (1 + 0_{1c}) S_{1n}^{1}(z_{10}) \]

Then \( \tilde{E}(z_{10}) = \sum_{i=1}^{n} W_{1}(z_{10}^{d}, z_{10}) A_{1}. \)

With (1)(a), we have \( s_{1n0}(z_{10}) = f_{1}(z_{10}) + O(p(ln z_{10})). \)

So for any \( y_{i} \), we have

\[ \text{Sup}_{\tilde{z}_{10} \in G} |\tilde{f}(z_{10}) - f_{i}(z_{10})| = O(p(ln z)) \text{ with } A2(1), (5) \text{ and A3.} \]

(c) Define \( S_{1n}(z_{10}) = \sum_{i=1}^{n} K_{1}(Z_{i1}^{c} - z_{10}^{c}) \)

\[ = \sum_{i=1}^{n} K_{1}(Z_{i1}^{c} - z_{10}^{c}) \]

Then \( \tilde{E}(z_{10}) = \sum_{i=1}^{n} W_{1}(z_{10}^{d}, z_{10}) A_{1}. \)

With (1)(a), we have \( s_{1n0}(z_{10}) = f_{1}(z_{10}) + O(p(ln z_{10})). \)

So for any \( y_{i} \), we have

\[ \text{Sup}_{\tilde{z}_{10} \in G} |\tilde{f}(z_{10}) - f_{i}(z_{10})| = O(p(ln z)) \text{ with } A2(1), (5) \text{ and A3.} \]
\[(10_r) S_{n}^{-1}(z_0) - \left(\frac{1}{m} S_{n}^{-1} 0_{r} \right) = [O_p(nZ) \ O_p(h)] \text{ and we expect } \]
\[\frac{1}{n h^2} \sum_{i=1}^{n} K_i(z_i - z_0)(z_i - z_0) g_i^* = O_p(\frac{1}{n h^2} \sum_{i=1}^{n} K_i(z_i - z_0) g_i) .] \]

(d) We show \[\sup_{z_0 \in C} |g_{1,j}(z_0) - g_{1,j}(z_1)| = O_p(h n.)\]

Since \[X_{i,j} = e_{i,j} + g_{1,j}(Z_{1i}), \text{ let } g_{1,j}^{(1)}(z_1) = \frac{\partial}{\partial z_1} g_{1,j}(z_1), \ g_{1,j}^{(2)}(z_1) = \frac{\partial^2}{\partial z_1^2} g_{1,j}(z_1), \]
\[g_{1,j}(z_1) - g_{1,j}(z_1) = \hat{E}(X_{0,0}, j|z_0) - g_{1,j}(z_0) \]
\[= \frac{1}{n h^2 (z_0)} \sum_{i=1}^{n} K_1(z_i - z_0)(e_{i,j} + 1(z_i - z_0) g_{1,j}^{(2)}(z_0) = \frac{\partial}{\partial z_1} g_{1,j}(z_0), \]
\[\text{where } z_0^{*} = \lambda Z_{1i} + (1 - \lambda) Z_{1i} \text{ for some } \lambda \in (0, 1). \text{ We apply Lemma 1 and use assumptions A2(1)-(3), A3, A4(1)} \text{ to obtain the claimed result in the same fashion as in (1)(a).}

(e) We show \[\sup_{z_0 \in C} |g_{2,j}(z_0) - g_{2,j}(z_0)| = O_p(h n.)\]

Since \[X_{i,j} = e_{i,j} + g_{2,j}(Z_{1i}), \text{ let } g_{2,j}^{(1)}(z_1) = \frac{\partial}{\partial z_1} g_{2,j}(z_1), \ g_{2,j}^{(2)}(z_1) = \frac{\partial^2}{\partial z_1^2} g_{2,j}(z_1), \]
\[g_{2,j}(z_1) - g_{2,j}(z_1) = \hat{E}(X_{0,0}, j|z_0) - g_{2,j}(z_0) \]
\[= \frac{1}{n h^2 (z_0)} \sum_{i=1}^{n} K_1(z_i - z_0)(e_{i,j} + 1(z_i - z_0) g_{2,j}^{(2)}(z_0) \text{ and similarly, we use terms in (d) and (e) to show terms A2, A4 and A5 are } o_p(1). \]

A3 \[\leq \sup_{z_0 \in C} |g_i(z_0) - g_i(z_0)| = \sup_{z_0 \in C} |g_i(z_0) - g_i(z_1)| = o_p(1) \] by Khinchin’s theorem, \[\frac{1}{n} \sum_{i=1}^{n} |g_i(Z_k) - g_i(Z_1)| \leq o_p(1) \text{ since } Z_k \text{ is iid, provided } E|g_j(Z_k) - g_j(Z_1)| < \infty. \]

By Khinchin’s theorem, \[A_0 = \sup_{z_0 \in C} E|g_j(Z_k) - g_j(Z_1)| = \sup_{z_0 \in C} E|g_j(Z_k) - g_j(Z_1)| \leq E|g_j(Z_k) - g_j(Z_1)| = e_j(1) < \infty. \] Since \[E|g_j(Z_k) - g_j(Z_1)| \leq E|g_j(Z_k) - g_j(Z_1)| = E|g_j(Z_k) - g_j(Z_1)| \leq E|g_j(Z_k) - g_j(Z_1)| = E|g_j(Z_k) - g_j(Z_1)| \leq \infty \text{ by Cauchy-Schwarz inequality and A4(1).}

Since \[E|e_{i,j} = E|X_{i,j} - g_{1,j}(Z_{1i}) < \infty, A_{10} \text{ and A11 are } o_p(1) \text{ with similar argument.}

Since \[E|g_i(Z_k) - g_i(Z_1)| = E|g_i(Z_k) - g_i(Z_1)| \leq E|g_i(Z_k) - g_i(Z_1)| = E|g_i(Z_k) - g_i(Z_1)| \leq \infty \text{ by Cauchy-Schwarz inequality and A4(1).}

(2) (a) \[\sup_{z_0 \in G} |\hat{E}(m_1(z_0))| \leq O_p(h_2^{1/2} n^{1/2} \lambda u_{10}^{1/2} + h_1^{1/2} + m_1^{1/2}). \]

\[\hat{E}(m_1(z_0)) = m(Z_1)|Z_1 - m(Z_1)| \]
\[= \frac{1}{n h^2 (z_0)} \sum_{i=1}^{n} K_1(Z_i - Z_1)|Z_i - m(Z_1)| - (Z_i - Z_1)m^{(1)}(Z_1)|Z_i - Z_1)\]
\[= \frac{1}{n h^2 (z_0)} \sum_{i=1}^{n} K_1(Z_i - Z_1)|Z_i - Z_1|m^{(2)}(Z_i - Z_1) (Z_i - Z_1)^{1/2} \]
\[+ E(\frac{1}{n h^2 (z_0)} \sum_{i=1}^{n} K_1(Z_i - Z_1)|Z_i - Z_1|m^{(2)} (Z_i - Z_1) (Z_i - Z_1)^{1/2} ) \]
\[= [V F M_n(Z_1) + h_1^{1/2} D F M_n(Z_1)] \text{ for } \lambda \in (0, 1). \text{ E}_i(\cdot) \text{ refers to the conditional expectation given } Z_i. \text{ The claim is shown with (2)(b) and (c).} \]
where the second to last equality uses the fact that
\[ C \] follows from Lemma 1 and assumption A2(7).

(d) With Lemma 1, A1(2), A2(1)-(3), A3 and A4(2), we have sup
\[ \sup_{Z_{1t} \in G_1} |\hat{E}| = o_p \left( \frac{n^{1/4}}{\ln n} \right) \] uniformly over \( G_1 \).

\[ C_{1k} \leq \sup_{Z_{1t} \in G_1} |g_k(Z_t) - g_k(Z_t)| + m(Z_{1t}) = o_p(\ln n * ln z_1) \]

\[ C_{2k} \leq \sup_{Z_{1t} \in G_1} |g_1,k(Z_{1t}) - g_1,k(Z_{1t})| + m(Z_{1t}) = o_p(\ln z_1^2) \]

\[ C_{3k} = \frac{1}{n} \sum_{t \in i} \left[ \frac{1}{n} \sum_{t \in i} \frac{W_{t,k}}{h_1^2} \left( K_{1t}(Z_{1t} - Z_{1t}) \right) ^2 \right] 
\]

\[ = o_p(n^{-1/2}h_1^4) \]

Note \( DFM_n(Z_{1t}) \) depends only on \( Z_{1t} \) and is bounded.

\[ EW_{t,k}DFM_n(Z_{1t}) = E[E[g_k(Z_t) - g_1,k(Z_{1t}) | Z_{1t}]]DFM_n(Z_{1t}) = 0, \]

and \( E(W_{t,k}DFM_n(Z_{1t})) < \infty \) by A4, so \( C_{31k} = o_p(n^{-1/2}h_1^4) \).

\[ C_{32k} = \frac{1}{n} \sum_{t \in i} \left[ \frac{1}{n} \sum_{t \in i} \frac{W_{t,k}}{h_1^2} \left( K_{1t}(Z_{1t} - Z_{1t}) \right) ^2 \right] 
\]

\[ = o_p(n^{-1/2}h_1^4) \]

where the second to last equality uses the fact that \( \phi_1 \) is symmetric in \( (Z_t, Z_{1t}) \) and the last equality uses the \( H^2 \)-decomposition for U-statistics with sample size dependent kernel (Theorem 1 in Yao and Martins-Filho (2013)).

Since \( E(W_{t,k}Z_{1t}) = 0, E(\hat{\phi}_{1t} \mid Z_t) = 0 \) as well.

\[ E(\hat{\phi}_{1t}) \leq CE_{1t} \]

\[ \leq CE_{1t} \left[ \frac{1}{h_1^2} \left( K_{1t}(Z_{1t} - Z_{1t}) \right) ^2 \right] 
\]

\[ C_{4k} = \frac{1}{n} \sum_{t \in i} \left[ g_k(Z_t) - g_k(Z_t) \right] \epsilon_t + o_p(\ln n * (\frac{\ln n}{h_1^2} )^{1/2}) \]

so \( C_{32k} = o_p(n^{-1/2}h_1^{-1/2}h_1^4) \).

\[ C_{41k} = \frac{1}{n} \sum_{t \in i} \left[ \frac{1}{h_1^2} \left( K_{1t}(Z_{1t} - Z_{1t}) \right) ^2 \right] \epsilon_k + \frac{1}{n^2} \sum_{t \in i} \epsilon_k K_{1t}(Z_{1t} - Z_{1t}) \left[ \frac{1}{h_1^2} \left( K_{1t}(Z_{1t} - Z_{1t}) \right) ^2 \right] \]

\[ C_{41k} + C_{412k} \]

Following \( C_{32k} \), \( \phi_1 = \psi_1 + \psi_1 \) when \( t \neq i \).
Finally with the Cramer-Rao device, we obtain

**Theorem 3:**

The conclusion of Theorem 3 follows from (1) and (3).

When $t = i$, $C_{411k} = \frac{K_i}{n h_i} + \frac{1}{n} \sum_{t} \frac{\sigma_{\phi(t)}}{\sqrt{\mathbb{E} \hat{\phi}_{ni}(t)}} = o_p(n^{-1/2})$.

Similarly, for $\phi_{2nti} = \psi_{2nti} + \psi_{2nti}$ when $t \neq i$.

$C_{412k} = \frac{1}{n} \sum_{t} E(\phi_{2nti}(Z_t, \epsilon_t) - \frac{1}{n} E\phi_{2nti} + O_p(n^{-1}(E\phi_{2nti}^2)^{1/2})) = o_p(n^{-1/2}).

Since $E(\epsilon_t|Z_t) = 0$, $E(\epsilon_{ti}|Z_t) = 0$, $E\phi_{2nti}^2 \leq CE\psi_{2nti}^2 = O(h^{-1/2})$ by assumption A4, so $o_p(n^{-1}(E\phi_{2nti}^2)^{1/2}) = o_p(n^{-1/2})$. Moreover, $E(\phi_{2nti}(Z_t, \epsilon_t) = E(\psi_{2nti}(Z_t, \epsilon_t)) = \frac{1}{n} \mathbb{E}(\mathbb{E}\hat{\phi}_{nti}(Z_t) \mid \tau) = \frac{1}{n} \sum_{t} E(\phi_{nti}(Z_t, \epsilon_t)) = \frac{1}{n} \sum_{t} E(\psi_{nti}(Z_t, \epsilon_t)) = o_p(n^{-1/2}).$

So $C_{412k} = o_p(n^{-1/2})$ and $C_{41k} = o_p(n^{-1/2})$.

$C_{5k} = \frac{1}{n} \sum_{t} E\phi_{ki}(Z_t) = \frac{1}{n} \sum_{t} E\phi_{ki}(Z_t) = o_p(n^{-1/2}).$

$C_{51k} = \frac{1}{n} \sum_{t} E\phi_{ki}(Z_t) = o_p(n^{-1/2}).$

$C_{6k} = \frac{1}{n} \sum_{t} E\phi_{ki}(Z_t) = o_p(n^{-1/2}).$

Since $\hat{E}(\epsilon_t|Z_t) = \frac{1}{n} \sum_{t} \epsilon_t = 0$, $E(\epsilon_{ti}|Z_t) = 0$, $E\phi_{2nti}^2 \leq CE\psi_{2nti}^2 = O(h^{-1/2})$ by assumption A4, so $o_p(n^{-1}(E\phi_{2nti}^2)^{1/2}) = o_p(n^{-1/2})$. Moreover, $E(\phi_{2nti}(Z_t, \epsilon_t) = E(\psi_{2nti}(Z_t, \epsilon_t)) = \frac{1}{n} \mathbb{E}(\mathbb{E}\hat{\phi}_{nti}(Z_t) \mid \tau) = \frac{1}{n} \sum_{t} E(\phi_{nti}(Z_t, \epsilon_t)) = \frac{1}{n} \sum_{t} E(\psi_{nti}(Z_t, \epsilon_t)) = o_p(n^{-1/2}).$

So $C_{6k} = o_p(n^{-1/2})$ and $C_{41k} = o_p(n^{-1/2})$.

Finally with the Cramer-Rao device, we obtain

$\sqrt{n} C_{6k} \to N(0, \sigma^2)$.

**Theorem 3: Proof.**

Let’s define the infeasible estimator

$\hat{\beta} = (\hat{\Omega}^{-1}(\hat{Z}_1)X)^{-1}(\hat{\Omega}^{-1}(\hat{Z}_1)(Y - \hat{E}(Y|\hat{Z}_1)))$, where the true $\sigma^2(Z_t)$ is known in $\Omega^{-1}(Z_1)$.

In the following we show

1. $\sqrt{n}(\hat{\beta} - \beta) \to N(0, (E\sigma^2(Z_t)W_t^{-1}W_t^{-1})^{-1})$.

2. $\sup_{Z_t \in [0]} |\hat{\sigma}^2(Z_t) - \sigma^2(Z_t)| = O_p((\frac{1}{n} h_i^2)^{1/2}) + O_p(n^{-1/2})$.

Result (2) might be of use by itself. Here repeated use of (2) enables us to obtain

3. $\sqrt{n}(\hat{\beta} - \beta) \to o_p(1)$.

The conclusion of Theorem 3 follows from (1) and (3).

1. $\hat{\beta} - \beta = [(\frac{1}{n} \hat{\Omega}^{-1}(\hat{Z}_1)X)^{-1} - (\frac{1}{n} \Omega^{-1}(\hat{Z}_1)X)^{-1}]W_t^{-1} + (\frac{1}{n} \hat{\Omega}^{-1}(\hat{Z}_1)X)^{-1}

(a) The $(i,j)$th element of $\frac{1}{n} \hat{\Omega}^{-1}(\hat{Z}_1)X$ is
\[
\frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma^2(Z_{1t})} \hat{W}_{t,i} X_{t,i} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma^2(Z_{1t})} \hat{W}_{t,i} \hat{X}_{t,i} + \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma^2(Z_{1t})} \hat{W}_{t,i} \hat{W}_{t,i}.
\]

Since \( \sigma^2(Z_{1t}) \) is iid, \( \sum_{i=1}^{n} \frac{1}{\sigma^2(Z_{1t})} \hat{W}_{t,i} \hat{X}_{t,i} \) is \( \sigma^2(Z_{1t}) \) by A6(1), and we follow the proof of Theorem 1 to obtain

\[
A_1 = o_p(1), \quad i = 1, 2, \ldots, 8, 10, 11, 12,
\]

provided \( \sum_{i=1}^{n} \frac{1}{\sigma^2(Z_{1t})} \hat{W}_{t,i} \hat{W}_{t,i} \rightarrow E \frac{1}{\sigma^2(Z_{1t})} \hat{W}_{t,i} \hat{W}_{t,i} = o_p(1). \)

So together we have \( \frac{1}{n} \hat{W} \sigma^2(Z_{1t})^{-1} \hat{W} \rightarrow E \frac{1}{\sigma^2(Z_{1t})} \hat{W} \hat{W} = o_p(1). \) By A6(2) and Slutsky' Theorem, \( (\frac{1}{n} \hat{W} \sigma^2(Z_{1t})^{-1} \hat{W} \hat{W})^{-1} \rightarrow E \frac{1}{\sigma^2(Z_{1t})} \hat{W} \hat{W}^{-1} = o_p(1). \)

(b) Similarly for \( k = 1, 2, \ldots, K, \) the \( k \)th element of \( C \) is

\[
C_k = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma^2(Z_{1t})} \hat{W}_{t,k} \sigma^2(Z_{1t}) \rightarrow E(m(z_{1t}) \sigma^2 | Z_{1t}) + E(e | Z_{1t})
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma^2(Z_{1t})} [g_k(Z_{1t}) - g_k(Z_{1t}) | m(Z_{1t}) - E(m(z_{1t}) | Z_{1t}) | m(Z_{1t}) - E(m(z_{1t}) | Z_{1t})]
\]

\[
+ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma^2(Z_{1t})} [g_k(Z_{1t}) - g_k(Z_{1t}) | m(Z_{1t}) - E(m(z_{1t}) | Z_{1t}) | m(Z_{1t}) - E(m(z_{1t}) | Z_{1t})]
\]

\[
+ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma^2(Z_{1t})} [g_k(Z_{1t}) - g_k(Z_{1t}) | m(Z_{1t}) - E(m(z_{1t}) | Z_{1t}) | m(Z_{1t}) - E(m(z_{1t}) | Z_{1t})]
\]

\[
= C_k + C_{2k} + \cdots + C_{6k}
\]

Since \( \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma^2(Z_{1t})} \rightarrow E \frac{1}{\sigma^2(Z_{1t})} \) \( g_k(Z_{1t}) - g_k(Z_{1t}) | m(Z_{1t}) - E(m(z_{1t}) | Z_{1t}) \) \( \rightarrow E \frac{1}{\sigma^2(Z_{1t})} \), we follow proof of Theorem 1 to obtain

\[
C_{ik} = o_p(n^{-1/2}) \quad \text{for} \quad i = 1, 2, 3, 4, 5
\]

with the additional assumption A6(1).

\[
C_{6k} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma^2(Z_{1t})} \hat{W}_{t,k} \epsilon_t
\]

Since \( E[\frac{1}{\sigma^2(Z_{1t})} | W_{t,k} \epsilon_t] = 0, E[\frac{1}{\sigma^2(Z_{1t})} | W_{t,k} \epsilon_t]^2 = E[\frac{1}{\sigma^2(Z_{1t})} (g_k(Z_{1t}) - g_k(Z_{1t}))^2] \), by Central Limit Theorem (Lindeberg-Lévy), with assumption A6, we have
\[ \sqrt{n}C_{61k} \xrightarrow{d} N(0, E \frac{W_{1k}^2}{\sigma^2(Z_{1i})}). \]

Finally with the Cramer-Rao device, for \( C_6 = [C_{61}, C_{62}, \ldots, C_{6k}]' \), we obtain

\[ \sqrt{n}C_6 \xrightarrow{d} N(0, E \frac{1}{\sigma^2(Z_{1i})} W_{1i}' W_{1i}). \]

So combine results in (a) and (b), we obtain the claim in (1).

(2) (a) We first note since \( \hat{\beta} - \beta = O_p(n^{-\frac{1}{2}}), \)

\[ \tilde{\epsilon}_t = m(Z_{1t}) - \hat{E}(m(z_1)|Z_{1t}) + \epsilon_t - \hat{E}(\epsilon|Z_{1t}) + (X_t - \hat{E}(X|Z_{1t}))\beta \]

\[ \tilde{\epsilon}_t^2 = (m(Z_{1t}) - \hat{E}(m(z_1)|Z_{1t}))^2 + (\epsilon_t - \hat{E}(\epsilon|Z_{1t}))^2 + ((X_t - \hat{E}(X|Z_{1t}))(\beta - \hat{\beta}) \]

\[ + 2(m(Z_{1t}) - \hat{E}(m(z_1)|Z_{1t})) (\epsilon_t - \hat{E}(\epsilon|Z_{1t})) \]

\[ + 2(\epsilon_t - \hat{E}(\epsilon|Z_{1t}))(X_t - \hat{E}(X|Z_{1t}))(\beta - \hat{\beta}) \]

\[ = I_1 + \cdots + I_6. \]

So \( \hat{\sigma}^2(Z_{1t}) = \hat{E}(\tilde{\epsilon}^2|Z_{1t}) = \hat{E}(I_1|Z_{1t}) + \cdots + \hat{E}(I_6|Z_{1t}). \)

\[ \hat{E}(I_1|Z_{1t}) = \frac{1}{nh_{i1}^2 f_1(Z_{1i})} \sum_{i=1}^n K_1(\frac{Z_{1i}^* - Z_{1i}^0}{h_1}) I(Z_{1i}^d = Z_{1i}^d)(m(Z_{1i}) - \hat{E}(m(z_1)|Z_{1i})) = O_p(1) \]

\[ \leq \sup_{Z_{1i} \in G_1} |m(Z_{1i}) - \hat{E}(m(z_1)|Z_{1i})|^2 [\frac{1}{f_1(Z_{1i})} (1 + O_p(h_1))] \frac{1}{nh_{i1}^2} \sum_{i=1}^n |K_1(\frac{Z_{1i}^* - Z_{1i}^0}{h_1})| I(Z_{1i}^d = Z_{1i}^d). \]

We notice that \( |K_1(\cdot)| \) satisfies the Lipschitz condition given assumption A3. So with assumption A6(3), we apply Lemma 1 and obtain \( \sup_{Z_{1i} \in G_1} |\hat{E}(I_1|Z_{1i})| = O_p((\frac{nh_{i1}^2}{mn})^{\frac{3}{2}}). \)

\[ E_{I_{11}} \rightarrow f_1(Z_{1t}) \int |K_1(\psi)|\psi \, d\psi < \infty \text{ uniformly in } Z_{1i} \in G_1. \]

So \( I_{11} = O_p(1) \) uniformly. With result in (2)(a) in Theorem 1, we conclude

\[ \sup_{Z_{1i} \in G_1} |\hat{E}(I_3|Z_{1i})| = O_p(h_{i1}^2 (\frac{nh_{i1}^2}{mn})^{\frac{3}{2}}) + O(h_{i1}^*) = o_p(n^{-\frac{1}{2}}) \text{ with assumption A5}. \]

\[ \hat{E}(I_3|Z_{1i}) = [(1 + O_p(h_1)) \frac{1}{f_1(Z_{1i})}] \frac{1}{nh_{i1}^2} \sum_{i=1}^n K_1(\frac{Z_{1i}^* - Z_{1i}^0}{h_1}) I(Z_{1i}^d = Z_{1i}^d) \sum_{k=1}^K \sum_{k' = 1}^K (X_{i,k} - \hat{g}_{1,k}(Z_{1i})). \]

\[ \times (X_{i,k'} - \hat{g}_{1,k'}(Z_{1i})) (\beta_k - \hat{\beta}_k)(\beta_{k'} - \hat{\beta}_{k'}). \]

\[ = O_p(n^{-1})(1 + O_p(h_1)) \frac{1}{f_1(Z_{1i})} \sum_{k=1}^K \sum_{k' = 1}^K K_1(\frac{Z_{1i}^* - Z_{1i}^0}{h_1}) I(Z_{1i}^d = Z_{1i}^d) \]

\[ \times |e_{1,k} e_{1,k'} + (g_{1,k}(Z_{1i}) - \hat{g}_{1,k}(Z_{1i})) e_{1,k'} + (g_{1,k'}(Z_{1i}) - \hat{g}_{1,k'}(Z_{1i})) e_{1,k} + (g_{1,k}(Z_{1i}) - \hat{g}_{1,k}(Z_{1i})) (g_{1,k'}(Z_{1i}) - \hat{g}_{1,k'}(Z_{1i}))|. \]

\[ = O_p(n^{-1})(1 + O_p(h_1)) \frac{1}{f_1(Z_{1i})} \sum_{k=1}^K \sum_{k' = 1}^K I_{31} = O_p(n^{-1})(1 + O_p(h_1)) \frac{1}{f_1(Z_{1i})} \sum_{k=1}^K \sum_{k' = 1}^K I_{311} + \cdots + I_{314}. \]

\( I_{314} = o_p(1) \) uniformly in \( Z_{1i} \in G_1 \) with result on \( I_{11} \) above and result (1)(d) in Theorem 1.

\[ I_{312} \leq \sup_{Z_{1i} \in G_1} |(g_{1,k}(Z_{1i}) - \hat{g}_{1,k}(Z_{1i}))| \frac{1}{nh_{i1}^2} \sum_{i=1}^n |K_1(\frac{Z_{1i}^* - Z_{1i}^0}{h_1})| I(Z_{1i}^d = Z_{1i}^d) e_{1,k,i}. \]
Since with assumption A2(3), A4(1) and A6(3), we apply Lemma 1 to have \(\sup_{Z_{it} \in G_1} |I_{312}| = O_p(1)\).

So \(I_{312} = o_p(1)\) uniformly in \(Z_{1t} \in G_1\). Similarly, \(I_{313} = o_p(1)\) uniformly in \(Z_{1t} \in G_1\).

\[
I_{311} = \frac{1}{nh^1_{it}} \sum_{i=1}^{n} K_1(\frac{Z_{1i} - Z_{c1}}{h_1})I(Z_{1i} = Z_{1t})e_{i,k}e_{1,k'}.
\]

Given assumptions A4(1) and A6(3), we have similarly \(\sup_{Z_{it} \in G_1} |I_{311}| = O_p(1)\).

So in all, we have \(I_{31} = O_p(1)\) uniformly in \(Z_{1t} \in G_1\) and \(\hat{E}(I_{31}|Z_{1t}) = O_p(n^{-1/2})\) uniformly in \(Z_{1t} \in G_1\) with results (2)(a) in Theorem 1. With assumptions A4(2), A4(4) and A6(3), we apply Lemma 1 to have \(\hat{E}(I_{32}|Z_{1t}) = O_p(n^{-1/2})\).

\[
\hat{E}(I_{4}|Z_{1t}) = \left[ (1 + O_p(h_1)) \frac{1}{I_{4}(Z_{1t})} \right] \frac{1}{nh^1_{it}} \sum_{i=1}^{n} K_1(\frac{Z_{1i} - Z_{c1}}{h_1})I(Z_{1i} = Z_{1t})\]
\[
\times (\epsilon_i - \hat{E}(\epsilon_i|Z_{1i})) \leq \left[ O_p(h_1^2(n^{-1/2} - 1)) + O_p(h_1^2) \right] \frac{1}{nh^1_{it}} \sum_{i=1}^{n} K_1(\frac{Z_{1i} - Z_{c1}}{h_1})I(Z_{1i} = Z_{1t})|X_i| + o_p(n^{-1/2}).
\]

We obtain easily that \(\sup_{Z_{it} \in G_1} |I_{52}| = o_p(1)\) from result on \(I_{11}\) and \(\sup_{Z_{it} \in G_1} |I_{51}| = O_p(1)\) as argued for term \(I_{312}\). So we conclude \(\sup_{Z_{it} \in G_1} |\hat{E}(I_{5}|Z_{1t})| = o_p(n^{-1/2}).\)

\[
\hat{E}(I_{6}|Z_{1t}) = \left[ (1 + O_p(h_1)) \frac{1}{I_{6}(Z_{1t})} \right] \frac{1}{nh^1_{it}} \sum_{i=1}^{n} K_1(\frac{Z_{1i} - Z_{c1}}{h_1})I(Z_{1i} = Z_{1t})\]
\[
\times (\hat{X}_{i,k} - g_{1,k}(Z_{1i}))(\beta_k - \hat{\beta}_k) \leq O_p(n^{-1/2}) \sum_{k=1}^{K} \left[ \frac{1}{nh^1_{it}} \sum_{i=1}^{n} K_1(\frac{Z_{1i} - Z_{c1}}{h_1})I(Z_{1i} = Z_{1t}) |X_i| \right] + o_p(n^{-1/2}).
\]

It is easy to see that \(\sup_{Z_{it} \in G_1} |I_{64}| = o_p(1)\) as in term \(I_{11}\). Similarly we have \(\sup_{Z_{it} \in G_1} |I_{62}| = o_p(1)\) and \(\sup_{Z_{it} \in G_1} |I_{63}| = o_p(1)\) with results on \(I_{41}\) and \(I_{312}\).
With assumption A6(3), \( E(\|\varepsilon_i\|^2 + \delta_i X_i, k\|^2 + \delta_i |Z_{i1}|) \leq \left| E(\|\varepsilon_i\|^4 + 2\delta_i E(|X_i, k|^{4 + 2\delta_i}|Z_{i1}|) \right|^\frac{3}{2} < \infty \), so we use A6(3) and apply Lemma 1 to obtain \( \sup_{Z_{i1} \in G_1} I_{611} = O_p(1) \).

With result (2)(d) in Theorem 1, we obtain uniformly for \( Z_{i1} \in G_1 \),

\[
\hat{E}(I_2|Z_{i1}) = \left[ (1 + O_p(h_i)) \frac{1}{f_i(Z_{i1})} \right] \frac{1}{nh_i} \sum_{i=1}^{n} K_i \left( \frac{Z_{i1} - Z_{i1}^c}{h_i} \right) (Z_{i1}^d - Z_{i1}^c) \{ \partial \varepsilon_i^2 - 2\varepsilon_i \hat{E}(\varepsilon_i|Z_{i1}) \} + \hat{E}(\varepsilon_i|Z_{i1}) \right] \}
\]

(b) So we have from above

\[
\hat{E}(I_2|Z_{i1}) = \left[ (1 + O_p(h_i)) \frac{1}{f_i(Z_{i1})} \right] \frac{1}{nh_i} \sum_{i=1}^{n} K_i \left( \frac{Z_{i1}^d - Z_{i1}^c}{h_i} \right) (Z_{i1}^d - Z_{i1}^c) \varepsilon_i^2 + O_p(n^{-\frac{1}{2}}) + O_p \left( \frac{(nh_i)^c}{(nh_i)^c} \right) + O_p(h_i^c).
\]

With A6(3) and A4(4), we apply Lemma 1 to obtain \( \sup_{Z_{i1} \in G_1} |I - EI| = O_p \left( \frac{(nh_i)^c}{(nh_i)^c} \right) + O_p(h_i^c) \). With a change of variable and using A6(1) and A2(1),

\[
EI = \int K_i(\psi) \sigma^2(Z_{i1} + h_i \psi, Z_{i1}^d) f_i(Z_{i1} + h_i \psi, Z_{i1}^d) d\psi
\]

\[
= \int K_i(\psi) \sigma^2(Z_{i1}) + \sum_{i=1}^{n} \frac{\partial}{\partial \varepsilon_i} \sigma^2(Z_{i1}) \sum_{j=1}^{n} \frac{\partial}{\partial \varepsilon_j} \sigma^2(Z_{i1}) \frac{1}{nh_i} \sum_{i=1}^{n} K_i(\psi) f_i(Z_{i1}) + \sum_{i=1}^{n} \frac{\partial}{\partial \varepsilon_i} \sigma^2(Z_{i1}) \int f_i(Z_{i1}^d) h_i^c \sum_{i=1}^{n} K_i(\psi) f_i(Z_{i1}) d\psi
\]

The claim in (2) above follows from (a) and (b).

Note \( \hat{E}(I_1|Z_{i1}) = O_p(n^{-1/2}) \) for \( i = 1, 3, 5 \),

\[
\hat{E}(I_2|Z_{i1}) = \frac{1}{nh_i^c f_i(Z_{i1})} \sum_{i=1}^{n} K_i(f_i(Z_{i1} - Z_{i1})) \{ \sigma^2(Z_{i1}) + \sigma^2(Z_{i1}) \} (1 + O_p(1) + O_p(n^{-1/2})),
\]

\[
\hat{E}(I_4|Z_{i1}) = \frac{1}{nh_i^c f_i(Z_{i1})} \sum_{i=1}^{n} K_i(f_i(Z_{i1} - Z_{i1})) \{ \sigma^2(Z_{i1}) + \sigma^2(Z_{i1}) \} (\hat{E}(m(z_1)|Z_{i1}) |_{Z_{i1}} \varepsilon_i |_{Z_{i1}} + O_p(n^{-1/2})),
\]

\[
\hat{E}(I_6|Z_{i1}) = \frac{1}{nh_i^c f_i(Z_{i1})} \sum_{i=1}^{n} K_i(f_i(Z_{i1} - Z_{i1})) \{ \sigma^2(Z_{i1}) + \sigma^2(Z_{i1}) \} (1 + O_p(1) + O_p(n^{-1/2})),
\]

\[
\hat{E}(I_7|Z_{i1}) = \frac{1}{nh_i^c f_i(Z_{i1})} \sum_{i=1}^{n} K_i(f_i(Z_{i1} - Z_{i1})) \{ \sigma^2(Z_{i1}) + \sigma^2(Z_{i1}) \} (1 + O_p(1) + O_p(n^{-1/2})).
\]

We use this alternative expression in (3).}
(3) \[ \sqrt{n}(\hat{\beta} - \beta^H) = \sqrt{n}(\frac{1}{n}(\hat{W}'\Omega^{-1}(\hat{Z}_i)^{-1} - (W'\hat{\Omega}^{-1}(\hat{Z}_i)^{-1})W'\Omega^{-1}(\hat{Z}_i)(Y - \hat{E}(Y|\hat{Z}_i))) \\
+ (W'\Omega^{-1}(\hat{Z}_i)^{-1}W')^{-1}(\hat{W}'\hat{\Omega}^{-1}(\hat{Z}_i)^{-1} - \Omega^{-1}(\hat{Z}_i))(Y - \hat{E}(Y|\hat{Z}_i))) \]

So we show

(a) \[ \frac{1}{n}\hat{W}'\Omega^{-1}(\hat{Z}_i)^{-1} - (\frac{1}{n}\hat{W}'\hat{\Omega}^{-1}(\hat{Z}_i)^{-1}) \rightarrow P \]

(b) \[ \sqrt{n}(\frac{1}{n}(\hat{W}'\hat{\Omega}^{-1}(\hat{Z}_i)^{-1}W')^{-1})^{-1}E_{\sigma^2(Z_{1i})}W'W_i \]

Since in (1) we have \[ \sqrt{n}(\frac{1}{n}(\hat{W}'\Omega^{-1}(\hat{Z}_i)^{-1} - (W'\hat{\Omega}^{-1}(\hat{Z}_i)^{-1})W'\Omega^{-1}(\hat{Z}_i))(Y - \hat{E}(Y|\hat{Z}_i))) = O_p(1) \]

We first note \[ \sup_{Z_{1i} \in G_1} \left| \frac{1}{\sigma^2(Z_{1i})} - \frac{1}{\hat{\sigma}^2(Z_{1i})} \right| \leq \left[ \inf_{Z_{1i} \in G_1} \sigma^2(Z_{1i}) \right] \inf_{Z_{1i} \in G_1} \hat{\sigma}^2(Z_{1i}) |\hat{\sigma}^2(Z_{1i}) - \sigma^2(Z_{1i})| \]

With result (2) and A6(1), for large n, \[ \inf_{Z_{1i} \in G_1} \hat{\sigma}^2(Z_{1i}) > 0 \]

so we have

(a \ '): \[ \frac{1}{n}\hat{W}'\Omega^{-1}(\hat{Z}_i)^{-1} \rightarrow P \]

If (a \ ') then \[ \frac{1}{n}\hat{W}'\Omega^{-1}(\hat{Z}_i)^{-1} \rightarrow P \frac{1}{\sigma^2(Z_{1i})} \rightarrow P \frac{1}{\hat{\sigma}^2(Z_{1i})} \]

The \((i,j)th\) element in \[ \frac{1}{n}\hat{W}'\Omega^{-1}(\hat{Z}_i)^{-1} \rightarrow P \]

\[ A_1 + \cdots + A_9 \]

\[ A_{10} - A_1 - A_4 - A_12 - A_7. \]
Since $\text{sup}_{Z_t \in G_1} |\frac{1}{\sigma(Z_t)} - \frac{1}{\sigma^2(Z_t)}| = o_p(1)$, we follow Theorem 1 (1) to have $A_i = o_p(1)$ for $i = 1, \cdots, 12$. So we have the claim in (a)’ and (a).

(b) The $k$th element in $\frac{1}{n} \hat{W}'[\Omega_1^{-1}(\hat{Z}_1) - \hat{\Omega}_1^{-1}(\hat{Z}_1)](Y - \hat{E}(Y|\hat{Z}_1))$ is

$$C_k = \frac{1}{n} \sum_t \left[ \frac{1}{\sigma(Z_t)} - \frac{1}{\sigma^2(Z_t)} \right] [\hat{g}_k(Z_t) - g_k(Z_t)] [m(Z_t) - \hat{E}(m(Z_t)|Z_t)]$$
$$+ \frac{1}{n} \sum_t \left[ \frac{1}{\sigma(Z_t)} - \frac{1}{\sigma^2(Z_t)} \right] [\hat{g}_1,k(Z_t) - \hat{g}_1,k(Z_t)] [m(Z_t) - \hat{E}(m(Z_t)|Z_t)]$$
$$+ \frac{1}{n} \sum_t \left[ \frac{1}{\sigma(Z_t)} - \frac{1}{\sigma^2(Z_t)} \right] [\hat{g}_k(Z_t) - g_k(Z_t)] [m(Z_t) - \hat{E}(m(Z_t)|Z_t)]$$
$$+ \frac{1}{n} \sum_t \left[ \frac{1}{\sigma(Z_t)} - \frac{1}{\sigma^2(Z_t)} \right] W_{t,k} \epsilon_t$$
$$+ \frac{1}{n} \sum_t \left[ \frac{1}{\sigma(Z_t)} - \frac{1}{\sigma^2(Z_t)} \right] \hat{W}_{t,k} \hat{E}(\epsilon|Z_t)$$

$$= C_{1k} + \cdots + C_{5k}.$$

With $\text{sup}_{Z_t \in G_1} |\frac{1}{\sigma(Z_t)} - \frac{1}{\sigma^2(Z_t)}| = O_p(nz_1)$, and Theorem 1 proof (1)(d), (e), (2)(d), we easily have $\sqrt{n}C_{ik} = o_p(1)$ for $i = 1, 2, 3, 5$.

$$C_{1k} = \frac{1}{n} \sum_t \left[ \frac{1}{\sigma(Z_t)} - \frac{1}{\sigma^2(Z_t)} \right] [\hat{g}_k(Z_t) - g_k(Z_t) + g_k(Z_t) - g_1,k(Z_t) + g_k(Z_t) - g_1,k(Z_t)] \epsilon_t$$
$$= \frac{1}{n} \sum_t \left[ \frac{1}{\sigma(Z_t)} - \frac{1}{\sigma^2(Z_t)} \right] W_{t,k} \epsilon_t + o_p(n^{-1/2})$$
$$= \frac{1}{n} \sum_t \frac{\sigma(Z_t) - \sigma^2(Z_t)}{\sigma^3(Z_t)} W_{t,k} \epsilon_t + o_p(n^{-1/2})$$
$$= \frac{1}{n} \sum_t \frac{\sigma(Z_t)}{\sigma^2(Z_t)} W_{t,k} \epsilon_t + \hat{D}_{6k}$$
$$= D_{1k} + \cdots + \hat{D}_{6k},$$

where the last three equalities uses the result (2) above and $\text{sup}_{Z_t \in G_1} |\frac{1}{\sigma(Z_t)} - \frac{1}{\sigma^2(Z_t)}| = O_p(nz_1)$.

We show below that $D_{1k} = o_p(n^{-1/2})$ for $i = 1, \cdots, 6$, which implies $C_{1k} = o_p(n^{-1/2})$.

$$D_{1k} = \frac{1}{n} \sum_t \left[ \frac{1}{\sigma(Z_t)} - \frac{1}{\sigma(Z_t^*)} \right] W_{t,k} \epsilon_t$$

$$= \frac{1}{n} \sum_t \frac{\sigma(Z_t)}{\sigma(Z_t^*)} W_{t,k} \epsilon_t$$

$$= \frac{1}{n} \sum_t \left[ \frac{1}{\sigma(Z_t^*)} - \frac{1}{\sigma(Z_t)} \right] K_{11}(Z_t - Z_{1t}) (\epsilon_t^2 - \sigma^2(Z_t))$$.

We follow the argument in $C_{32k}$ of Theorem 1 to perform H-decomposition to obtain

$$D_{1k} = \frac{1}{n} \sum_t E(\phi_{nti}|Z_t, \epsilon_t) - \frac{1}{n} \sum_t \phi_{nti} + o_p(n^{-1/2}) = O_p(n^{-1/2})$$

since $E\phi_{nti}^2 \leq CE\psi_{nti}^2 \leq C \frac{\sigma(Z_t)}{h_{1t}^2} E\left[ \frac{W_{t,k}^2}{\sigma^2(Z_t)} K_{11}(Z_t - Z_{1t}) \right] (E(\epsilon_t^2|Z_t) + 4\sigma^2(Z_t)) = O(h_{1t}^{-1/2})$.

$$D_{2k} = \frac{1}{n} \sum_t \frac{\sigma(Z_t)}{\sigma(Z_t^*)} W_{t,k} \epsilon_t$$

$$= \frac{1}{n} \sum_t \frac{\sigma(Z_t)}{\sigma(Z_t^*)} W_{t,k} \epsilon_t$$

$$= \frac{1}{n} \sum_t \left[ \frac{1}{\sigma(Z_t^*)} - \frac{1}{\sigma(Z_t)} \right] K_{11}(Z_t - Z_{1t}) (\sigma_t^2(Z_t^*) + \sigma_t^2(Z_t))$$.

Similarly, $D_{2k} = \frac{1}{n} \sum_t E(\phi_{nti}|Z_t, \epsilon_t) + o_p(n^{-1/2})$. $E\phi_{nti} = O(h_{1t}^{-1/2})$ and

$$E(\phi_{nti}|Z_t, \epsilon_t) = \frac{1}{n} \sum_t \frac{\sigma(Z_t)}{\sigma(Z_t^*)} W_{t,k} \epsilon_t$$

$$= \frac{\sigma(Z_t)}{\sigma(Z_t^*)} K_{11}(Z_t - Z_{1t}) + o(h_{1t}^{-1/2})$$.

So $E(\phi_{nti}|Z_t, \epsilon_t) = 0$ and $E[E(\phi_{nti}|Z_t, \epsilon_t)] = O(h_{1t}^{-1/2})$, so $\frac{1}{n} \sum_t E(\phi_{nti}|Z_t, \epsilon_t) = o_p(n^{-1/2})$.

So in all, $D_{2k} = o_p(n^{-1/2})$.

$$D_{3k} = \frac{1}{n} \sum_t \frac{\sigma(Z_t)}{\sigma(Z_t^*)} W_{t,k} \epsilon_t$$

$$= -2 \frac{1}{n} \sum_t \sum_{j=1}^{n} \frac{W_{t,k} \epsilon_t \epsilon_j}{h_{1j}^{1/2} f_1(Z_t) f_1(Z_t^*)} K_{11}(Z_t - Z_{1t}) K_{11}(Z_{1j} - Z_{1t})$$.

When $t \neq i$ and $j$, let $\phi_{nij} = \psi_{nij} + \psi_{nij} + \psi_{nij} + \psi_{nij} + \psi_{nij} + \psi_{nij}$, $D_{3k} = (-2) \frac{1}{n} \frac{6}{n^2} - \left( \begin{array}{c} n \\ 3 \end{array} \right) - 1 \left( \begin{array}{c} n \\ 3 \end{array} \right) - 1 \left( \begin{array}{c} n \\ 3 \end{array} \right) - 1 \left( \begin{array}{c} n \\ 3 \end{array} \right) - 1 \left( \begin{array}{c} n \\ 3 \end{array} \right) - 1 \left( \begin{array}{c} n \\ 3 \end{array} \right) - 1 \left( \begin{array}{c} n \\ 3 \end{array} \right)$$.
where $H_n^{(j)} = O_p((n^{-3/2}h_n^{2/3}))$. Here $\theta_n = \sigma_n^2 = \sigma_n^2 = 0$.

$\sigma_3^2 \leq CE\psi_n^2 = O(h_n^{2/3})$, so $H_n^{(3)} = O_p(n^{-3/2}h_n^{1/3}) = o_p(n^{-1/2})$. So $D_{3k} = o_p(n^{-1/2})$. We obtain that $D_{3k} = o_p(n^{-1/2})$ when $t = i = j$, $t = i$, $i = j$, and $t = j$ easily.

$$D_{4k} = \frac{1}{n} \sum_t \frac{S_t(Z_{i1})}{\sigma_t(Z_{i1})} W_{t,k} \epsilon_t = -2 * \frac{h_t^3}{n^3} \sum_t \sum_{t,k=1}^n \frac{W_{t,k} \epsilon_t}{h_t^3 f_1(Z_{i1})} K_{1f}(Z_{11} - Z_{11}) \epsilon_t DFM_{n}(Z_{i1}).$$

We show in a similar fashion that $D_{4k} = (-2h_t^3)O_p(n^{-1}h_t^{1/3}) = o_p(n^{-1/2})$ when $t \neq i$ and $C_{4k} = o_p(n^{-1/2}h_t^1) = o_p(n^{-1/2})$ when $t = i$. So $C_{4k} = o_p(n^{-1/2})$.

$$D_{5k} = \frac{1}{n} \sum_t \frac{S_t(Z_{i1})}{\sigma_t(Z_{i1})} W_{t,k} \epsilon_t$$

$$= \frac{1}{n} \sum_t \sum_{t,k=1}^n \frac{W_{t,k} \epsilon_t}{h_t^3 f_1(Z_{i1})} \left[ K_{1f}(Z_{11} - Z_{11})(Z_{ij} - Z_{11})m(2)(Z_{11}, Z_{11}')(Z_{ij} - Z_{11})' \right]$$

$$- E(K_{1f}(Z_{11} - Z_{11})(Z_{ij} - Z_{11})m(2)(Z_{ij}, Z_{ij}')(Z_{ij} - Z_{11})' | Z_i)$$

$$= o_p(n^{-1/2})$$

$D_{6k}$ can be shown with similar arguments.

$$D_{6k} = \frac{1}{n} \sum_t \frac{S_t(Z_{i1})}{\sigma_t(Z_{i1})} W_{t,k} \epsilon_t$$

$$= 2 \sum_{t,k=1}^n (\beta_k - \bar{\beta}_k) \frac{1}{n} \sum_t \sum_{t,k=1}^n \frac{W_{t,k} \epsilon_t}{h_t^3 f_1(Z_{i1})} K_{1f}(Z_{11} - Z_{11}) \epsilon_t e_{1,k}$$

$$= o_p(n^{-1/2})$$ with similar arguments.

References


[37] Su, L., I. Murtazashvili, and A. Ullah, 2011, Local linear GMM estimation of functional coefficient IV models with an application to estimating the rate of return to schooling, manuscript, Singapore Management University.


