Technical Supplement to: Efficient kernel-based semiparametric IV estimation with an application to resolving a puzzle on the estimates of the return to schooling

Feng Yao
Department of Economics
West Virginia University
Morgantown, WV 26505 USA
email: feng.yao@mail.wvu.edu
Voice: + 1 304 293 7867

Junsen Zhang
Department of Economics
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
email: jszhang@cuhk.edu.hk
Voice: + 852 2609 8186

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, $Z_i^c\in R^{l_c}, Z_i^d\in R^{l_d}, K_h(z^c)=\frac{1}{h^{l_c}}K(\frac{z^c}{h})$, and $K(.)$ is a kernel function defined on R^{l_c}. If we have

L_1. $K(.)$ is bounded with compact support and for Euclidean norm $||.||$,

$$|u^jK(u)-v^jK(v)|\le c_K||u-v||,\ for\ 0\le|j|\le J.$$

L_2. $g(u)$ is a measurable function of u_i and $E|g(u)|^s<\infty$ for $s>2$.

L_3. $\sup_{z\in G}\int|g(u)|^s f_{z,u}(z,u)du<\infty$, $f_{z|u}(z)<\infty$, and $f_{z,u}(z,u)$ is continuous around z^c.

L_4. $|w(Z_i^c-z^c;z)|<\infty$, $\forall z^d\in G^d$, a compact subset of R^{l_d}, $|w(Z_i^c-z^c;z^c,z^d)-w(Z_i^c-z_k^c;z_k^c,z^d)|\le c||z^c-z_k^c||$.

L_5. $nh^{l_c}\to\infty$.

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_c},

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

Proof. Let's define

$$S_{n,j}^B(x)=\frac{1}{n}\sum_{i=1}^{n}K_h(Z_i^c-z^c)\left(\frac{Z_i^c-z^c}{h}\right)^j I(Z_i^d=z^d)g(U_i)w(Z_i^c-z^c;z)I(|g(U_i)|\le B_n),$$

where $B_1\le B_2\le\cdots$ 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_c 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, $l_nr_n^{l_c}=c$, c a constant. Suppose we let $l_n=\left(\frac{n}{ln(n)h^{l_c+2}}\right)^{\frac{l_c}{2}}$, then $r_n=c/l_n^{\frac{1}{l_c}}$. Since

$$\begin{aligned}
&\sup_{z\in G}|S_{n,j}^B(z)-E(S_{n,j}^B(z))|\\
=\ &max_{1\le k\le l_n}\sup_{z\in I_k\cap G}|S_{n,j}^B(z)-E(S_{n,j}^B(z))\\
=\ &max_{1\le k\le l_n}\sup_{z^c\in I_k\cap G}|S_{n,j}^B(z^c,z_k^d)-S_{n,j}^B(z_k)\\
&+S_{n,j}^B(z_k)-ES_{n,j}^B(z_k)+ES_{n,j}^B(z_k)-ES_{n,j}^B(z^c,z_k^d)|\\
\le\ &max_{1\le k\le l_n}\sup_{z^c\in I_k\cap G}|S_{n,j}^B(z^c,z_k^d)-S_{n,j}^B(z_k)|\\
&+max_{1\le k\le l_n}|S_{n,j}^B(z_k)-ES_{n,j}^B(z_k)|\\
&+max_{1\le k\le l_n}\sup_{z^c\in I_k\cap G}|ES_{n,j}^B(z_k)-ES_{n,j}^B(z^c,z_k^d)|\\
=\ &I_1+I_2+I_3
\end{aligned}$$

The lemma is proved if we can show

(1) $I_0=\sup_{z\in G}|S_{n,j}(z)-E(S_{n,j}(z))-[S_{n,j}^B(z)-E(S_{n,j}^B(z))]|=O_{a.s.}(B_n^{1-s})$ for $B_n^{1-s}=O(\left(\frac{ln(n)}{nh^{l_c}}\right)^{\frac{1}{2}})$, $\sum_{i=1}^{\infty}B_i^{-s}<\infty$.

(2) $I_1=O_{a.s.}(\left(\frac{ln(n)}{nh^{l_c}}\right)^{\frac{1}{2}})$. (3) $I_2=O_p(\left(\frac{ln(n)}{nh^{l_c}}\right)^{\frac{1}{2}})$. (4) $I_3=O_{a.s.}(\left(\frac{ln(n)}{nh^{l_c}}\right)^{\frac{1}{2}})$.

(1) $I_0\le\sup_{z\in G}|S_{n,j}(z)-S_{n,j}^B(z)|+\sup_{z\in G}|E(S_{n,j}(z)-S_{n,j}^B(z))|=I_{01}+I_{02}$. We note

$$I_{01} = \sup_{z\in G}\left|\frac{1}{n}\sum_{i=1}^{n} K_h(Z_i^c - z^c)\left(\frac{Z_i^c - z^c}{h}\right)^j I(Z_i^d = z^d)g(U_i)w(Z_i)I(|g(U_i)| > B_n)\right|.$$

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

$$P(\cup_{m=i}^{\infty}\{\omega : |g(U_m)| > B_m\}) < \epsilon, \quad \text{or } P(\cap_{m=i}^{\infty}\{\omega : |g(U_m)| \le B_m\}) > 1 - \epsilon.$$

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

When $i = \{1, 2, \cdots, i'\}$, $P(|g(U_i)| \le B_n) > 1 - \epsilon$. To see this, $\forall\epsilon > 0$, and sufficiently large n, $P(|g(U_i)| > B_n) < \frac{E|g(U_i)|^s}{B_n^s} < \frac{c}{B_n^s} < \epsilon$, since $E|g(U_i)|^s < \infty$ 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..

$I_{02} = \sup_{z\in G}|EK_h(Z_i^c - z^c)\left(\frac{Z_i^c - z^c}{h}\right)^j I(Z_i^d = z^d)g(U_i)w(Z_i^c - z^c; z)I(|g(U_i)| > B_n)$. Let $\frac{Z_i^c - z^c}{h} = (\frac{Z_{i1}^c - z_1^c}{h}, \cdots, \frac{Z_{il_c}^c - z_{l_c}^c}{h}) = \Psi_i = (\Psi_{i1}, \cdots, \Psi_{il_c})$, so $Z_i^c = (z_1^c + h\Psi_{i1}, \cdots, z_{l_c}^c + h\Psi_{il_c}) = z^c + h\Psi_i$, $|\frac{\partial Z_i^c}{\partial \Psi_i}| = h^{l_c}$. By change of variable,

$$\begin{aligned}
I_{02} = \quad & \sup_{z\in G}|\textstyle\sum_{Z_i^d = z^d}\int K(\Psi_i)\Psi^j \int w(h\Psi_i; z)g(U_i)I(|g(U_i)| > B_n) \\
& \times f_{z,u}(z^c + h\Psi_i, z^d, U_i)dU_i d\Psi_i| \\
\le \quad & c\textstyle\int |K(\Psi_i)\Psi_i^j| d\Psi_i \sup_{z\in G}\int |g(U_i)| f_{z,U_i}(z, U_i)I(|g(U_i)| > B_n)dU_i \\
\le \quad & c\sup_{z\in G}[\textstyle\int |g(U_i)|^s f_{z,U_i}(z, U_i)dU_i]^{\frac{1}{s}}[\int I(|g(U_i)| > B_n)f_{z,U_i}(z, U_i)dU_i]^{1-\frac{1}{s}} \\
\le \quad & c[\textstyle\int I(|g(U_i)| > B_n)f_{z|U_i}(z)dU_i]^{1-\frac{1}{s}} \\
\le \quad & c[E(I(|g(U_i)| > B_n)]^{1-\frac{1}{s}} = c[P(|g(U_i)| > B_n)]^{1-\frac{1}{s}} \\
\le \quad & c[\tfrac{E|g(U_i)|^s}{B_n^s}]^{1-\frac{1}{s}} \le cB_n^{1-s}.
\end{aligned}$$

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.

$$\begin{aligned}
(2) \quad & |S_{n,j}^B(z^c, z_k^d) - S_{n,j}^B(z_k)| \\
= \quad & |\tfrac{1}{nh^{l_c}}\sum_{i=1}^{n}[K_h(Z_i^c - z^c)\left(\tfrac{Z_i^c - z^c}{h}\right)^j w(Z_i^c - z^c; z^c, z_k^d) \\
& -K_h(Z_i^c - Z_k^c)\left(\tfrac{Z_i^c - Z_k^c}{h}\right)^j w(Z_i^c - z_k^c; z_k)]I(Z_i^d = Z_k^d)g(U_i)I(|g(U_i)| \le B_n)| \\
\le \quad & \tfrac{1}{nh^{l_c}}\sum_{i=1}^{n}[|[K(Z_i^c - z^c)\left(\tfrac{Z_i^c - z^c}{h}\right)^j - K(Z_i^c - z_k^c)\left(\tfrac{Z_i^c - z_k^c}{h}\right)^j]w(Z_i^c - z^c; z^c, z_k^d)| \\
& +|K(Z_i^c - Z_k^c)\left(\tfrac{Z_i^c - Z_k^c}{h}\right)^j [w(Z_i^c - z^c; z^c, z_k^d) - w(Z_i^c - z_k^c; z_k)]|] \\
& \times I(Z_i^d = Z_k^d)g(U_i)I(|g(U_i)| \le B_n)| \\
\le \quad & \tfrac{1}{nh^{l_c}}\sum_{i=1}^{n}[c\tfrac{||Z_k^c - z^c||}{h} + c||Z_k^c - z^c||]|g(U_i)| \text{ by } L_1 \text{ and } L_4,
\end{aligned}$$

since $z \in I_k$ for some k, $||Z_k^c - z^c|| \le cr_n$ and with L_4,
$I_1 \le c\frac{r_n}{h^{l_c+1}}\frac{1}{n}\sum_{i=1}^{n}|g(U_i)|$, by L_2 and Kolmogorov's Theorem,

$$\frac{1}{n}\sum_{i=1}^{n}|g(U_i)| \overset{a.s.}{\to} E|g(U_i)| < \infty.$$

So $I_1 \le c\frac{r_n}{h^{l_c+1}} = \frac{c}{h^{l_c+1}}(\frac{n}{ln(n)h^{l_c+2}})^{-\frac{1}{2}} = c(\frac{ln(n)}{nh^{l_c}})^{\frac{1}{2}}$ *a.s.*.

We could show (4) $I_3 = O_{a.s.}\left(\frac{ln(n)}{nh^{l_c}}\right)^{\frac{1}{2}}$ similarly.
(3) It is sufficient to show $\exists$ a constant $\Delta > 0$ and $N > 0$ such that $\forall \epsilon > 0$ and $n > N$, $P(\left(\frac{ln(n)}{nh^{l_c}}\right)^{\frac{1}{2}} I_2 \geq \Delta) < \epsilon$.
Let $\epsilon_n = \left(\frac{ln(n)}{nh^{l_c}}\right)^{\frac{1}{2}}\Delta$, then $P(I_2 \geq \epsilon_n) \leq \sum_{k=1}^{l_n} P(|S_{n,j}^B(z_k) - ES_{n,j}^B(z_k)| \geq \epsilon_n)$.

$$\begin{aligned} & |S_{n,j}^B(z_k) - ES_{n,j}^B(z_k)| = |\tfrac{1}{n}\textstyle\sum_{i=1}^n W_{in}| \\ = \; & |\tfrac{1}{n}\textstyle\sum_{i=1}^n[\frac{1}{h^{l_c}}K(Z_i^c - z_k^c)\left(\frac{Z_i^c - z_k^c}{h}\right)^j I(Z_i^d = z_k^d)g(U_i)w(Z_i^c - z_k^c; z_k)I(|g(U_i)| \leq B_n) \\ & -\tfrac{1}{h^{l_c}}EK(Z_i^c - z_k^c)\left(\frac{Z_i^c - z_k^c}{h}\right)^j I(Z_i^d = z_k^d)g(U_i)w(Z_i^c - z_k^c; z_k)I(|g(U_i)| \leq B_n)]|. \end{aligned}$$

Since $EW_{in} = 0$, $|W_{in}| \leq 2c\frac{B_n}{h^{l_c}}$ by L_1 and L_4, and $\{W_{in}\}_{i=1}^n$ is an independent sequence, by Bernstein's inequality,

$$P(|S_{n,j}^B(z_k) - ES_{n,j}^B(z_k)| \geq \epsilon_n) < 2exp\left(\frac{-nh^{l_c}\epsilon_n^2}{2h^{l_c}\bar{\sigma}^2 + \frac{2}{3}B_n\epsilon_n}\right),$$

where $\bar{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n V(W_{in}) = EW_{in}^2 = I_{21} - I_{22}^2$

$$\begin{aligned} = \; & \tfrac{1}{h^{2l_c}}EK^2(Z_i^c - z^c)(\left(\frac{Z_i^c - z_k^c}{h}\right)^j)^2 I(Z_i^d = z_k^d)g^2(U_i)w(Z_i^c - z_k^c; z_k)^2 I(|g(U_i)| \leq B_n)]| \\ & -\left[\tfrac{1}{h^{l_c}}EK(Z_i^c - z_k^c)\left(\frac{Z_i^c - z_k^c}{h}\right)^j I(Z_i^d = z_k^d)g(U_i)w(Z_i^c - z_k^c; z_k)I(|g(U_i)| \leq B_n)]\right]^2. \end{aligned}$$

$$\begin{aligned} I_{22} = \; & \textstyle\sum_{Z_i^d = z_k^d}\int K(\Psi)\Psi_i^j g(U_i)w(h\Psi_i; z_k)I(|g(U_i)| \leq B_n)f_{z,u}(z_k^c + h\Psi_i, Z_i^d, U_i)d\psi_i dU_i \\ \leq \; & c\textstyle\int |K(\Psi)\Psi_i^j||g(U_i)|f_{z,u}(Z_k^c + h\Psi_i, Z_k^d, U_i)d\psi_i dU_i \\ \to \; & c\textstyle\int |K(\Psi)\Psi_i^j|d\Psi_i \int |g(U_i)|f_{z,u}(z_k, U_i)dU_i < \infty, \text{ with } L_1, L_3 \text{ and } L_4. \end{aligned}$$

Similarly $h^{l_c}I_{21} = O(1)$. So $2h^{l_c}\bar{\sigma}^2 < \infty$. If $B_n\epsilon_n < \infty$, then $C_n = 2h^{l_c}\bar{\sigma}^2 + \frac{2}{3}B_n\epsilon_n < \infty$, then

$$\begin{aligned} P(I_2 \geq \epsilon_n) \leq \; & l_n 2exp\left(\frac{-nh^{l_c}\epsilon_n^2}{2h^{l_c}\bar{\sigma}^2 + \frac{2}{3}B_n\epsilon_n}\right) \\ = \; & \left(\frac{n}{ln(n)h^{l_c+2}}\right)^{\frac{l_c}{2}} 2exp\left(\frac{-nh^{l_c}\left(\left(\frac{ln(n)}{nh^{l_c}}\right)^{\frac{1}{2}}\Delta\right)^2}{C_n}\right) = \frac{2n^{\frac{l_c}{2} - \frac{\Delta^2}{C_n}}}{(ln(n))^{\frac{l_c}{2}}h^{l_c + \frac{l_c^2}{2}}} \to 0. \end{aligned}$$

Above is true since $C_n < \infty$, if we let $\Delta^2 \geq C_n(1 + l_c)$, then $\frac{2n^{\frac{l_c}{2} - \frac{\Delta^2}{C_n}}}{(ln(n))^{\frac{l_c}{2}}h^{l_c + \frac{l_c^2}{2}}} \leq \frac{2}{(ln(n))^{\frac{l_c}{2}}(nh^{l_c})^{1 + \frac{l_c}{2}}} \to 0$ by L_5.
If we let $B_n = n^{\frac{1}{s}+\delta}$ for $s > 2$ and $\delta > 0$, then $B_n\epsilon_n < \infty$ for sufficiently large s. To see this, $B_n\epsilon_n = n^{\frac{1}{s}+\delta}\Delta(\frac{ln(n)}{nh^{l_c}})^{\frac{1}{2}}$. By L_5, we could let $(nh^{l_c})^{-\frac{1}{2}} = n^{-\frac{1}{2}+\delta_1}$ for $\frac{1}{2} > \delta_1 > 0$, then $B_n\epsilon_n = n^{\frac{1}{s}-\frac{1}{2}+\delta+\delta_1}\Delta(ln(n))^{\frac{1}{2}}$. If we let $s > [\frac{1}{2} - \delta - \delta_1]^{-1}$, then $B_n\epsilon_n \to 0$.
It is easy to see that for $B_n = n^{\frac{1}{s}+\delta}$, we easily have $\sum_{i=1}^\infty B_i^{-s} < \infty$. Furthermore $B_n^{1-s} < n^{\frac{1}{2}-\delta}$, so $B_n^{1-s} = O\left(\frac{ln(n)}{nh^{l_c}}\right)^{\frac{1}{2}}$.

Theorem 1: *Proof.* Note $Y_t - \hat{E}(Y|Z_{1t}) = (X_t - \hat{E}(X|Z_{1t}))\beta + m(Z_{1t}) - \hat{E}(m(z_1)|Z_{1t}) + \epsilon_t - \hat{E}(\epsilon|Z_{1t})$, so we could write

$$\begin{aligned} & \hat{\beta} - \beta = [(\tfrac{1}{n}\hat{W}'\check{X})^{-1} - (EW_t'W_t)^{-1} + (EW_t'W_t)^{-1}] \\ & \times \underbrace{\frac{1}{n}\hat{W}'(\vec{m} - \hat{E}(m(z_1)|\vec{Z}_1) + \vec{\epsilon} - \hat{E}(\epsilon|\vec{Z}_1))}_{C}, \end{aligned}$$

where $\vec{m} = \{m(Z_{1t})\}_{t=1}^n$, $\hat{E}(m(z_1)|\vec{Z}_1) = \{\hat{E}(m(z_1)|Z_{1t})\}_{t=1}^n$, $\vec{\epsilon} = \{\epsilon_t\}_{t=1}^n$, $\hat{E}(\epsilon|\vec{Z}_1) = \{\hat{E}(\epsilon|Z_{1t})\}_{t=1}^n$.
Let's denote $\hat{E}(X_k|\mathbf{Z_t}) = \hat{g}_k(\mathbf{Z_t})$ and $\hat{E}(X_k|Z_{1t}) = \hat{g}_{1,k}(Z_{1t})$, then $\hat{W}_{t,k} = \hat{g}_k(\mathbf{Z_t}) - g_k(\mathbf{Z_t}) + g_{1,k}(Z_{1t}) - \hat{g}_{1,k}(Z_{1t}) + g_k(\mathbf{Z_t}) - g_{1,k}(Z_{1t})$. The $(i,j)th$ element of $\frac{1}{n}\hat{W}'\check{X}$ is

$$\tfrac{1}{n}\textstyle\sum_{t=1}^n \hat{W}_{t,i}\check{X}_{t,j} = \tfrac{1}{n}\sum_{t=1}^n \hat{W}_{t,i}(X_{t,j} - \hat{g}_j(\mathbf{Z_t})) + \tfrac{1}{n}\sum_{t=1}^n \hat{W}_{t,i}\hat{W}_{t,j}.$$

$$
\begin{aligned}
& \frac{1}{n}\sum_{t=1}^n \hat{W}_{t,i}\hat{W}_{t,j} \\
= \ & \frac{1}{n}\sum_t[\hat{g}_i(\mathbf{Z_t})-g_i(\mathbf{Z_t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})]+\frac{1}{n}\sum_t[\hat{g}_i(\mathbf{Z_t})-g_i(\mathbf{Z_t})][g_{1,j}(Z_{1t})-\hat{g}_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t[\hat{g}_i(\mathbf{Z_t})-g_i(\mathbf{Z_t})][g_j(\mathbf{Z_t})-g_{1,j}(Z_{1t})]+\frac{1}{n}\sum_t[g_{1,i}(Z_{1t})-\hat{g}_{1,i}(Z_{1t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})] \\
& +\frac{1}{n}\sum_t[g_{1,i}(Z_{1t})-\hat{g}_{1,i}(Z_{1t})][g_{1,j}(Z_{1t})-\hat{g}_{1,j}(Z_{1t})]+\frac{1}{n}\sum_t[g_{1,i}(Z_{1t})-\hat{g}_{1,i}(Z_{1t})][g_j(\mathbf{Z_t})-g_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})]+\frac{1}{n}\sum_t[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})][g_{1,j}(Z_{1t})-\hat{g}_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})][g_j(\mathbf{Z_t})-g_{1,j}(Z_{1t})] \\
= \ & A_1+A_2+\cdots+A_9
\end{aligned}
$$

$$
\begin{aligned}
& \frac{1}{n}\sum_{t=1}^n \hat{W}_{t,i}(X_{t,j}-\hat{g}_j(\mathbf{Z_t}))=\frac{1}{n}\sum_{t=1}^n \hat{W}_{t,i}(g_j(\mathbf{Z_t})-\hat{g}_j(\mathbf{Z_t})+e_{jt}) \\
= \ & (-1)\frac{1}{n}\sum_t[\hat{g}_i(\mathbf{Z_t})-g_i(\mathbf{Z_t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})]+\frac{1}{n}\sum_t[\hat{g}_i(\mathbf{Z_t})-g_i(\mathbf{Z_t})]e_{jt} \\
& -\frac{1}{n}\sum_t[g_{1,i}(Z_{1t})-\hat{g}_{1,i}(Z_{1t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})]+\frac{1}{n}\sum_t[g_{1,i}(Z_{1t})-\hat{g}_{1,i}(Z_{1t})]e_{jt} \\
& -\frac{1}{n}\sum_t[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})]+\frac{1}{n}\sum_t[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})]e_{jt} \\
= \ & -A_1+A_{10}-A_4+A_{11}-A_7+A_{12}.
\end{aligned}
$$

Similarly, for $k=1,2,\cdots,K$, the kth element of C is

$$
\begin{aligned}
C_k = \ & \frac{1}{n}\sum_{t=1}^n \hat{W}_{t,k}(m(Z_{1t})-\hat{E}(m(z_1)|Z_{1t})+\epsilon_t-\hat{E}(\epsilon|Z_{1t})) \\
= \ & \frac{1}{n}\sum_t[\hat{g}_k(\mathbf{Z_t})-g_k(\mathbf{Z_t})][m(Z_{1t})-\hat{E}(m(z_1)|Z_{1t})] \\
& +\frac{1}{n}\sum_t[g_{1,k}(Z_{1t})-\hat{g}_{1,k}(Z_{1t})][m(Z_{1t})-\hat{E}(m(z_1)|Z_{1t})] \\
& +\frac{1}{n}\sum_t[g_k(\mathbf{Z_t})-g_{1,k}(Z_{1t})][m(Z_{1t})-\hat{E}(m(z_1)|Z_{1t})] \\
& +\frac{1}{n}\sum_t[\hat{g}_k(\mathbf{Z_t})-g_k(\mathbf{Z_t})][\epsilon_t-\hat{E}(\epsilon|Z_{1t})] \\
& +\frac{1}{n}\sum_t[g_{1,k}(Z_{1t})-\hat{g}_{1,k}(Z_{1t})][\epsilon_t-\hat{E}(\epsilon|Z_{1t})] \\
& +\frac{1}{n}\sum_t[g_k(\mathbf{Z_t})-g_{1,k}(Z_{1t})][\epsilon_t-\hat{E}(\epsilon|Z_{1t})] \\
= \ & C_{1k}+C_{2k}+\cdots+C_{6k}.
\end{aligned}
$$

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

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

so together we have $\frac{1}{n}\hat{W}'\check{X}-EW_t'W_t=o_p(1)$. By $A1(3)$ and Slutsky' Theorem, $(\frac{1}{n}\hat{W}'\check{X})^{-1}-(EW_t'W_t)^{-1}=o_p(1)$.
(2) Denote $lnz_1=(\frac{ln(n)}{nh_1^{l_{1c}}})^{1/2}+h_1^{s_1}$, $lnz=(\frac{ln(n)}{nh^{l_c}})^{1/2}+h^s$. We show that

$C_{1k}=O_p(lnz*lnz_1)$, $C_{2k}=O_p(lnz_1^2)$, $C_{3k}=o_p(n^{-1/2})$, $C_{4k}=o_p(n^{-1/2})+O_p(lnz*(\frac{ln(n)}{nh_1^{l_{1c}}})^{1/2})$,

$C_{5k}=o_p(n^{-1/2})+O_p(lnz_1*(\frac{ln(n)}{nh^{l_c}})^{1/2})$.

For $C_6=[C_{61},C_{62},\cdots,C_{6K}]'$, $\sqrt{n}C_6\xrightarrow{d}N(0,\Phi_0)$, where Φ_0 is defined in Theorem 1.
With A5, $nh^{2l_c}\to\infty$, $nh_1^{2l_{1c}}\to\infty$, $nh^{4s}\to 0$ and $nh_1^{4s_1}\to 0$. It implies that $C_{1k}+\cdots+C_{5k}=o_p(n^{-1/2})+O_p(lnz^2+lnz_1^2)=o_p(n^{-1/2})$.
Combining results in (1) and (2) and using A1(3), we conclude

$$\sqrt{n}(\hat{\beta}-\beta)\xrightarrow{d}N(0,(EW_t'W_t)^{-1}\Phi_0(EW_t'W_t)^{-1}).$$

(1) (a) Define $\hat{f}_1(z_{10})=\frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^n K_{1I}(Z_{1i}-z_{10})$. We first show $\sup_{z_{10}\in G_1}|\hat{f}_1(z_{10})-f_1(z_{10})|=O_p(lnz_1)$. We apply Lemma 1 with $S_{n,0}(z_{10})=\hat{f}_1(z_{10})$, so

$$\sup_{z_{10}\in G_1}|\hat{f}_1(z_{10})-E\hat{f}_1(z_{10})|=O_p((\frac{nh_1^{l_{1c}}}{ln(n)})^{-\frac{1}{2}}).$$

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),

$E\hat{f}(z_{10}) = \int \frac{1}{h_1^{l_{1c}}} K_1(\frac{Z_{1i}^c - z_{10}^c}{h_1}) f_1(Z_{1i}^c, z_{10}^d) dZ_{1i}^c$, with $\Psi_i = \frac{Z_{1i}^c - z_{10}^c}{h_1}$,
$= \int K_1(\Psi_i) f_1(z_{10}^c + h_1\Psi_i, z_{10}^d) d\Psi_i$, with A2(1)
$= \int K_1(\Psi_i)[f_1(z_{10}^c, z_{10}^d) + \sum_{|j|=1}^{s_1-1} \frac{h_1^{|j|}}{j!} \frac{\partial^j f_1(z_{10}^c, z_{10}^d)}{\partial(z_{10}^c)^j} \Psi_i^j + \sum_{|j|=s_1} \frac{h_1^{s_1}}{j!} \frac{\partial^j f_1(z_{10*}^c, z_{10}^d)}{\partial(z_{10}^c)^j} \Psi_i^j] d\Psi_i$
$= f_1(z_{10}) + O(h_1^{s_1})$ uniformly $\forall z_{10} \in G_1$ by A3, A2(1) and Dominated Convergence Theorem, where z_{10*}^c is between z_{10}^c and Z_{1i}^c. So $\sup_{z_{10} \in G_1} |E\hat{f}_1(z_{10}) - f_1(z_{10})| = O(h_1^{s_1})$.

(b) Define $\hat{f}(\mathbf{z_0}) = \frac{1}{nh^{l_c}} \sum_{i=1}^n K_I(\mathbf{Z_i} - \mathbf{z_0})$. Similarly, we obtain $\sup_{\mathbf{z_0} \in G} |\hat{f}(\mathbf{z_0}) - f_z(\mathbf{z_0})| = O_p(lnz)$ with A2(4), (5) and A3.

(c) Define $S_{1n}(z_{10}) = \begin{bmatrix} s_{1n0}(z_{10}) & s_{1n1}(z_{10}) \\ s_{1n1}'(z_{10}) & s_{1n2}(z_{10}) \end{bmatrix}$, where $s_{1nj}(z_{10}) = \frac{1}{nh_1^{l_{1c}}} \sum_{i=1}^n K_{1I}(Z_{1i} - z_{10})(\frac{Z_{1i}^c - z_{10}^c}{h_1})^j$ for $j = 0, 1$, and $s_{1n2}(z_{10}) = \frac{1}{nh_1^{l_{1c}}} \sum_{i=1}^n K_{1I}(Z_{1i} - z_{10})(\frac{Z_{1i}^c - z_{10}^c}{h_1})'(\frac{Z_{1i}^c - z_{10}^c}{h_1})$. Let

$$W_1(\frac{Z_{1i}^c - z_{10}^c}{h_1}, z_{10}) = (1 \ 0_{l_{1c}}') S_{1n}^{-1}(z_{10})(1 \ \ \frac{Z_{1i}^c - z_{10}^c}{h_1})' K_{1I}(Z_{1i} - z_{10}).$$

Then $\hat{E}(A|z_{10}) = \frac{1}{nh^{l_{1c}}} \sum_{i=1}^n W_1(\frac{Z_{1i}^c - z_{10}^c}{h_1}, z_{10}) A_i$.

With (1)(a), we have $s_{1n0}(z_{10}) = f_1(z_{10}) + O_p(lnz_1)$. With similar arguments , we have $s_{1n1}(z_{10}) = O_p((\frac{ln(n)}{nh_1^{l_{1c}}})^{1/2} + h_1^{s_1-1}) = O_p(h_1^{s_1-1})$ and $s_{1n2}(z_{10}) = O_p((\frac{ln(n)}{nh_1^{l_{1c}}})^{1/2} + h_1^{s_1-2}) = O_p(h_1^{s_1-2})$ uniformly over the compact set G_1, where the second equalities follow if we choose $h_1 = O(n^{-1/(2s_1+l_{1c})})$.

So for any y_i^*,

$$\begin{aligned} & \frac{1}{nh^{l_{1c}}} \sum_{i=1}^n W_1(\frac{Z_{1i}^c - z_{10}^c}{h_1}, z_{10}) y_i^* - \frac{1}{nh^{l_c} f_1(z_{10})} \sum_{i=1}^n K_{1I}(Z_{1i} - z_{10}) y_i^* \\ = & \frac{1}{nh^{l_c}} \sum_{i=1}^n [W_1(\frac{Z_{1i}^c - z_{10}^c}{h_1}, z_{10}) - \frac{1}{f_1(z_{10})} K_{1I}(Z_{1i} - z_{10})] y_i^* \\ = & [(1 \ 0_{l_{1c}}') S_{1n}^{-1}(z_{10}) - (\frac{1}{f_1(z_{10})} \ 0_{l_{1c}}')] \begin{bmatrix} \frac{1}{nh^{l_c}} \sum_{i=1}^n K_{1I}(Z_{1i} - z_{10}) y_i^* \\ \frac{1}{nh^{l_c}} \sum_{i=1}^n K_{1I}(Z_{1i} - z_{10})(\frac{Z_{1i}^c - z_{10}^c}{h_1}) y_i^* \end{bmatrix}. \end{aligned}$$

Then $\frac{1}{nh^{l_c}} \sum_{i=1}^n W_1(\frac{Z_{1i}^c - z_{10}^c}{h_1}, z_{10}) y_i^* = \frac{1}{nh^{l_{1c}} f_1(z_{10})} \sum_{i=1}^n K_{1I}(Z_{1i} - z_{10}) y_i^* (1 + O_p(h_1))$ since
$[(1 \ 0_{l_{1c}}') S_{1n}^{-1}(z_{10}) - (\frac{1}{f_1(z_{10})} \ 0_{l_{1c}}')] = [O_p(lnz_1) \ O_p(h_1)]$ and we expect
$\frac{1}{nh^{l_{1c}}} \sum_{i=1}^n K_{1I}(Z_{1i} - z_{10})(\frac{Z_{1i}^c - z_{10}^c}{h_1}) y_i^* = O_p(\frac{1}{nh^{l_{1c}}} \sum_{i=1}^n K_{1I}(Z_{1i} - z_{10}) y_i^*)$.

Define $S_n(\mathbf{z_0}) = \begin{bmatrix} s_{n0}(\mathbf{z_0}) & s_{n1}(\mathbf{z_0}) \\ s_{n1}'(\mathbf{z_0}) & s_{n2}(\mathbf{z_0}) \end{bmatrix}$, where $s_{nj}(\mathbf{z_0}) = \frac{1}{nh^{l_c}} \sum_{i=1}^n K_I(\mathbf{Z_i} - \mathbf{z_0})(\frac{\mathbf{Z_i}^c - \mathbf{z_0}^c}{h})^j$ for $j = 0, 1$, and $s_{n2}(\mathbf{z_0}) = \frac{1}{nh^{l_c}} \sum_{i=1}^n K_I(\mathbf{Z_i} - \mathbf{z_0})(\frac{\mathbf{Z_i}^c - \mathbf{z_0}^c}{h})'(\frac{\mathbf{Z_i}^c - \mathbf{z_0}^c}{h})$. Let

$$W(\frac{\mathbf{Z_i}^c - \mathbf{z_0}^c}{h}, \mathbf{z_0}) = (1 \ 0_{l_c}') S_n^{-1}(\mathbf{z_0})(1 \ \ \frac{\mathbf{Z_i}^c - \mathbf{z_0}^c}{h})' K_I(\mathbf{Z_i} - \mathbf{z_0}).$$

Then $\hat{E}(A|\mathbf{z_0}) = \frac{1}{nh^{l_c}} \sum_{i=1}^n W(\frac{\mathbf{Z_i}^c - \mathbf{z_0}^c}{h}, \mathbf{z_0}) A_i$.

With (1)(b), we have $s_{n0}(\mathbf{z_0}) = f(\mathbf{z_0}) + O_p(lnz)$. With similar arguments , we have $s_{n1}(\mathbf{z_0}) = O_p((\frac{ln(n)}{nh^{l_c}})^{1/2} + h^{s-1}) = O_p(h^{s-1})$ and $s_{n2}(\mathbf{z_0}) = O_p((\frac{ln(n)}{nh^{l_c}})^{1/2} + h^{s-2}) = O_p(h^{s-2})$ uniformly over the compact set G, where the second equality follows if we choose $h = O(n^{-1/(2s+l_c)})$.

So for any y_i^*, $\frac{1}{nh^{l_c}} \sum_{i=1}^n W(\frac{\mathbf{Z_i}^c - \mathbf{z_0}^c}{h}, \mathbf{z_0}) y_i^* = \frac{1}{nh^{l_c} f(\mathbf{z_0})} \sum_{i=1}^n K_I(\mathbf{Z_i} - \mathbf{z_0}) y_i^* (1 + O_p(h))$ since

$[(1\ 0'_{l_c})S_n^{-1}(\mathbf{z_0}) - (\frac{1}{f(\mathbf{z_0})}\ 0'_{l_c})] = [O_p(lnz)\ O_p(h)]$ and we expect
$\frac{1}{nh^{l_c}}\sum_{i=1}^{n} K_I(\mathbf{Z_i}-\mathbf{z_0})(\frac{\mathbf{Z_i}^c-\mathbf{z_0}^c}{h})y_i^* = O_p(\frac{1}{nh^{l_c}}\sum_{i=1}^{n} K_I(\mathbf{Z_i}-\mathbf{z_0})y_i^*)$.

(d) We show $\sup_{z_{10}\in G_1}|\hat{g}_{1,j}(z_{10}) - g_{1,j}(z_{10})| = O_p(lnz_1)$.
Since $X_{i,j} = e_{1,ji} + g_{1,j}(Z_{1i})$, let $g_{1,j}^{(1)}(z_{10}) = \frac{\partial}{\partial(z_1^c)'}g_{1,j}(z_{10})$, $g_{1,j}^{(2)}(z_{10}) = \frac{\partial^2}{\partial(z_1^c)'\partial(z_1^c)}g_{1,j}(z_{10})$,

$$\begin{aligned}
&\hat{g}_{1,j}(z_{10}) - g_{1,j}(z_{10}) = \hat{E}(X_{0,j}|z_{10}) - g_{1,j}(z_{10})\\
=\ &\frac{1}{nh^{l_{1c}}}\sum_{i=1}^{n} W_1(\frac{Z_{1i}^c - z_{10}^c}{h_1}, z_{10})[X_{i,j} - g_{1,j}(z_{10}) - (Z_{1i}^c - z_{10}^c)g_{1,j}^{(1)}(z_{10})]\\
=\ &\frac{1}{nh^{l_{1c}}f_1(z_{10})}\sum_{i=1}^{n} K_{1I}(Z_{1i}-z_{10})[e_{1,ji} + \frac{1}{2}(Z_{1i}^c - z_{10}^c)g_{1,j}^{(2)}(z_{10}^{*c}, z_{10}^d)(Z_{1i}^c - z_{10}^c)](1+O_p(h_1)),
\end{aligned}$$

where $z_{10}^{*c} = \lambda z_{10}^c + (1-\lambda)Z_{1i}^c$ for some $\lambda \in (0,1)$. We apply Lemma 1 and use assumptions A2(1)-(3), A3, A4(1) to obtain the claimed result in the same fashion as in (1)(a).

(e) We show $\sup_{\mathbf{z_0}\in G}|\hat{g}_j(\mathbf{z_0}) - g_j(\mathbf{z_0})| = O_p(lnz)$.
Since $X_{i,j} = e_{ji} + g_j(\mathbf{Z_i})$, let $g_j^{(1)}(\mathbf{z_0}) = \frac{\partial}{\partial(z^c)'}g_j(\mathbf{z_0})$, $g_j^{(2)}(\mathbf{z_0}) = \frac{\partial^2}{\partial(z^c)'\partial(z^c)}g_j(\mathbf{z_0})$,

$$\begin{aligned}
&\hat{g}_j(\mathbf{z_0}) - g_j(\mathbf{z_0}) = \hat{E}(X_{0,j}|\mathbf{z_0}) - g_j(\mathbf{z_0})\\
=\ &\frac{1}{nh^{l_c}}\sum_{i=1}^{n} W(\frac{\mathbf{Z_i}^c - \mathbf{z_0}^c}{h}, \mathbf{z_0})[X_{i,j} - g_j(\mathbf{z_0}) - (\mathbf{Z_i}^c - \mathbf{z_0}^c)g_j^{(1)}(\mathbf{z_0})]\\
=\ &\frac{1}{nh^{l_c}f(\mathbf{z_0})}\sum_{i=1}^{n} K_I(\mathbf{Z_i}-\mathbf{z_0})[e_{ji} + \frac{1}{2}(\mathbf{Z_i}^c - \mathbf{z_0}^c)g_j^{(2)}(\mathbf{z_0}^{*c}, \mathbf{z_0}^d)(\mathbf{Z_i}^c - \mathbf{z_0}^c)](1+O_p(h)).
\end{aligned}$$

We apply Lemma 1 and use assumptions A2(4)-(6), A3, A4(1) to obtain the claimed result.

$A_1 = \frac{1}{n}\sum_t[\hat{g}_i(\mathbf{Z_t}) - g_i(\mathbf{Z_t})][\hat{g}_j(\mathbf{Z_t}) - g_j(\mathbf{Z_t})] = O_p(lnz^2)$ with result in (e) and A5.
Similarly, we use results in (d) and (e) to show terms A_2, A_4 and A_5 are $o_p(1)$.
$A_3 \le \sup_{\mathbf{z_0}\in G}|\hat{g}_i(\mathbf{z_0}) - g_i(\mathbf{z_0})|\frac{1}{n}\sum_t|g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})| = o_p(1)\frac{1}{n}\sum_t|g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})|$, since $\mathbf{Z_t}$ is iid, by Khinchin's theorem, $\frac{1}{n}\sum_t|g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})| \xrightarrow{p} E|g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})|$, provided $E|g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})| < \infty$. Since $E|g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})| = E|E(X_{t,j}|\mathbf{Z_t}) - E(X_{t,j}|Z_{1t})| < 2E(|X_t|) < \infty$ by assumption A4, $A_3 = o_p(1)$. Similar arguments show that A_6, A_7 and A_8 are $o_p(1)$.
By Khinchin's theorem, $A_9 \xrightarrow{p} E[g_i(\mathbf{Z_t}) - g_{1,i}(Z_{1t})][g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})]$, which is the $(i,j)th$ element of $EW_t'W_t$, provided $E[g_i(\mathbf{Z_t}) - g_{1,i}(Z_{1t})][g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})] < \infty$.
$E[g_i(\mathbf{Z_t}) - g_{1,i}(Z_{1t})][g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})] \le E|g_i(\mathbf{Z_t})g_i(\mathbf{Z_t})| + E|g_i(\mathbf{Z_t})g_{1,j}(Z_{1t})| + E|g_{1,i}(Z_{1t})g_j(\mathbf{Z_t})| + E|g_{1,i}(Z_{1t})g_{1,j}(Z_{1t})| \le \infty$ by Cauchy-Schwartz inequality and A4(1).

Since $E|e_{jt}| = E|X_{t,j} - g_j(\mathbf{Z_t})| < \infty$, A_{10} and A_{11} are $o_p(1)$ with similar argument.
Since $E(g_i(\mathbf{Z_t}) - g_{1,i}(Z_{1t}))e_{jt} = 0$, $E[(g_i(\mathbf{Z_t}) - g_{1,i}(Z_{1t}))e_{jt}]^2 < \infty$, $A_{12} = O_p(n^{-1/2}) = o_p(1)$.

(2) (a) $\sup_{Z_{1t}\in G_1}|\hat{E}(m(z_1)|Z_{1t}) - m(Z_{1t})| = O_p(h_1^2(\frac{ln(n)}{nh_1^{l_{1c}}})^{1/2} + h_1^{s_1})$.

$$\begin{aligned}
&\hat{E}(m(z_1)|Z_{1t}) - m(Z_{1t})\\
=\ &\frac{1}{nh^{l_{1c}}}\sum_{i=1}^{n} W_1(\frac{Z_{1i}^c - Z_{1t}^c}{h_1}, Z_{1t})[m(Z_{1i}) - m(Z_{1t}) - (Z_{1i}^c - Z_{1t}^c)m^{(1)}(Z_{1t})]\\
=\ &[\frac{1}{nh^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n} K_{1I}(Z_{1i}-Z_{1t})\frac{1}{2}(Z_{1i}^c - Z_{1t}^c)m^{(2)}(Z_{1t}^{*c}, Z_{1t}^d)(Z_{1i}^c - Z_{1t}^c)'\\
&-E_t(\frac{1}{nh^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n} K_{1I}(Z_{1i}-Z_{1t})\frac{1}{2}(Z_{1i}^c - Z_{1t}^c)m^{(2)}(Z_{1t}^{*c}, Z_{1t}^d)(Z_{1i}^c - Z_{1t}^c)')\\
&+E_t(\frac{1}{nh^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n} K_{1I}(Z_{1i}-Z_{1t})\frac{1}{2}(Z_{1i}^c - Z_{1t}^c)m^{(2)}(Z_{1t}^{*c}, Z_{1t}^d)(Z_{1i}^c - Z_{1t}^c)')](1+O_p(h_1))\\
=\ &[VFM_n(Z_{1t}) + h_1^{s_1}DFM_n(Z_{1t}](1+O_p(h_1)).
\end{aligned}$$

$m^{(1)}(Z_{1t}) = \frac{\partial}{\partial(z_1^c)'}m(Z_{1t})$, $m^{(2)}(Z_{1t}) = \frac{\partial^2}{\partial(z_1^c)'\partial(z_1^c)}m(Z_{1t})$, and $Z_{1t}^{*c} = \lambda Z_{1t}^c + (1-\lambda)Z_{1i}^c$ for $\lambda \in (0,1)$.
$E_t(\cdot)$ refers to the conditional expectation given $\mathbf{Z_t}$. The claim is shown with (2)(b) and (c).

$(b)\quad E_t(\frac{1}{nh^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n} K_{1I}(Z_{1i}-Z_{1t})\frac{1}{2}(Z_{1i}^c-Z_{1t}^c)m^{(2)}(Z_{1t}^{*c},Z_{1t}^d)(Z_{1i}^c-Z_{1t}^c)')$

$= \frac{h_1^2}{f_1(Z_{1t})}\int K_1(\psi_1)I(Z_{1i}^d=Z_{1t}^d)\frac{1}{2}\psi_1 m^{(2)}(Z_{1t}^c+\lambda h_1\psi_1,Z_{1t}^d)\psi_1' f_1(Z_{1t}^c+h_1\psi_1,Z_{1t}^d)d\psi_1$

$= \frac{1}{f_1(Z_{1t})}\int K_1(\psi_1)I(Z_{1i}^d=Z_{1t}^d)[\sum_{|j|=2}^{s_1-1}\frac{h_1^{|j|}}{j!}\frac{\partial^{(j)}m(Z_{1t})}{\partial(Z_{1t}^c)^j}(\psi_1)^j+\sum_{|j|=s_1}\frac{h_1^{|j|}}{j!}\frac{\partial^{(j)}m(Z_{1t}^{*c},Z_{1t}^d)}{\partial(Z_{1t}^c)^j}(\psi_1)^j]$

$*[f_1(Z_{1t})+\sum_{|j|=1}^{s_1-3}\frac{h_1^{|j|}}{j!}\frac{\partial^{(j)}f_1(Z_{1t})}{\partial(Z_{1t}^c)^j}(\psi_1)^j+\sum_{|j|=s_1-2}\frac{h_1^{|j|}}{j!}\frac{\partial^{(j)}f_1(Z_{1t}^{*c},Z_{1t}^d)}{\partial(Z_{1t}^c)^j}(\psi_1)^j]d\psi_1$

$= \frac{h_1^{s_1}}{f_1(Z_{1t})}\int K_1(\psi_1)I(Z_{1i}^d=Z_{1t}^d)[(\sum_{|j|=2}\frac{1}{j!}\frac{\partial^{(j)}m(Z_{1t})}{\partial(Z_{1t}^c)^j}(\psi_1)^j)(\sum_{|j|=s_1-2}\frac{1}{j!}\frac{\partial^{(j)}f_1(Z_{1t})}{\partial(Z_{1t}^c)^j}(\psi_1)^j)$

$+f_1(Z_{1t})\sum_{|j|=s_1}\frac{1}{j!}\frac{\partial^{(j)}m(Z_{1t})}{\partial(Z_{1t}^c)^j}(\psi_1)^j]d\psi_1(1+o_p(1))$

$= h_1^{s_1}DFM_n(Z_{1t})=O_p(h_1^{s_1})$ uniformly over G_1.

(c) $\sup_{Z_{1t}\in G_1}|VFM_n(Z_{1t})|=O_p(h_1^2(\frac{ln(n)}{nh_1^{l_{1c}}})^{1/2})$ follows from Lemma 1 and assumption A2(7).

(d) With Lemma 1, A1(2), A2(1)-(3), A3 and A4(2), we have $\sup_{Z_{1t}\in G_1}|\hat{E}(\epsilon|Z_{1t})|=O_p((\frac{nh_1^{l_{1c}}}{ln(n)})^{-\frac{1}{2}})$.

$C_{1k}\le \sup_{\mathbf{Z_t}\in G}|\hat{g}_k(\mathbf{Z_t})-g_k(\mathbf{Z_t})|*|m(Z_{1t})-\hat{E}(m(z_1)|Z_{1t})|=O_p(lnz*lnz_1)$.

$C_{2k}\le \sup_{Z_{1t}\in G_1}|\hat{g}_{1,k}(Z_{1t})-g_{1,k}(Z_{1t})|*|m(Z_{1t})-\hat{E}(m(z_1)|Z_{1t})|=O_p(lnz_1^2)$.

$C_{3k} = \frac{1}{n}\sum_t[g_k(\mathbf{Z_t})-g_{1,k}(Z_{1t})][m(Z_{1t})-\hat{E}(m(z_1)|Z_{1t})]$

$= \frac{1}{n}\sum_t W_{t,k}(-h_1^{s_1}DFM_n(Z_{1t})-VFM_n(Z_{1t}))$

$= -(C_{31k}+C_{32k})=o_p(n^{-1/2})$.

Note $DFM_n(Z_{1t})$ depends only on Z_{1t} and is bounded.
$EW_{t,k}DFM_n(Z_{1t})=E[E(g_k(\mathbf{Z_t})-g_{1,k}(Z_{1t})|Z_{1t})DFM_n(Z_{1t})]=0$, and $E(W_{t,k}DFM_n(Z_{1t}))^2<\infty$ by A4, so $C_{31k}=O_p(n^{-1/2}h_1^{s_1})=o_p(n^{-1/2})$.

$C_{32k} = \frac{1}{n^2}\sum\sum_{t\neq i}\frac{W_{t,k}}{h_1^{l_{1c}}f_1(Z_{1t})}[K_{1I}(Z_{1i}-Z_{1t})\frac{1}{2}(Z_{1i}^c-Z_{1t}^c)m^{(2)}(Z_{1t}^{*c},Z_{1t}^d)(Z_{1i}^c-Z_{1t}^c)'$

$-E_t(K_{1I}(Z_{1i}-Z_{1t})\frac{1}{2}(Z_{1i}^c-Z_{1t}^c)m^{(2)}(Z_{1t}^{*c},Z_{1t}^d)(Z_{1i}^c-Z_{1t}^c)')]$

$= \frac{1}{n^2}\sum\sum_{t\neq i}\psi_{nti}=\frac{1}{2n^2}\sum\sum_{t\neq i}\underbrace{(\psi_{nti}+\psi_{nit})}_{\phi_{nti}}=\frac{1}{n^2}\sum\sum_{t<i}\phi_{nti}$

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

where the second to last equality uses the fact that ϕ_{nti} is symmetric in $(\mathbf{Z_t},\mathbf{Z_i})$ and the last equality uses the H-decomposition for U-statistics with sample size dependent kernel (Theorem 1 in Yao and Martins-Filho (2013)).

Since $E(W_{t,k}|Z_{1t})=0$, $E\phi_{tni}0$. $E(\phi_{nti}|\mathbf{Z_t})=0$ as well.

$E\phi_{tni}^2 \le CE\psi_{tni}^2$

$\le CE[\frac{W_{t,k}^2}{h_1^{2l_{1c}}f_1^2(Z_{1t})}[K_{1I}^2(Z_{1i}-Z_{1t})(\frac{1}{2}(Z_{1i}^c-Z_{1t}^c)m^{(2)}(Z_{1t}^{*c},Z_{1t}^d)(Z_{1i}^c-Z_{1t}^c)')^2$

$+(E_t(K_{1I}(Z_{1i}-Z_{1t})\frac{1}{2}(Z_{1i}^c-Z_{1t}^c)m^{(2)}(Z_{1t}^{*c},Z_{1t}^d)(Z_{1i}^c-Z_{1t}^c)')^2]$

$=O_p(h_1^{-l_{1c}}h_1^4)$,

so $C_{32k}=O_p(n^{-1}h_1^{-l_{1c}/2}h_1^2)=o_p(n^{-1/2})$.

$C_{4k} = \frac{1}{n}\sum_t[\hat{g}_k(\mathbf{Z_t})-g_k(\mathbf{Z_t})][\epsilon_t-\hat{E}(\epsilon|Z_{1t})]=\frac{1}{n}\sum_t[\hat{g}_k(\mathbf{Z_t})-g_k(\mathbf{Z_t})]\epsilon_t+O_p(lnz*(\frac{ln(n)}{nh_1^{l_{1c}}})^{1/2})$

$= C_{41k}+O_p(lnz*(\frac{ln(n)}{nh_1^{l_{1c}}})^{1/2})$.

$C_{41k} = \frac{1}{n^2}\sum_{t=1}^{n}\sum_{i=1}^{n}\underbrace{\frac{\epsilon_t K_I(\mathbf{Z_t}-\mathbf{Z_i})}{h^{l_c}f(\mathbf{Z_t})}e_{ki}}_{\psi_{1nti}}+\frac{1}{n^2}\sum\sum_{t\neq i}\underbrace{\frac{\epsilon_t K_I(\mathbf{Z_t}-\mathbf{Z_i})}{h^{l_c}f(\mathbf{Z_t})}\frac{1}{2}(\mathbf{Z_i}^c-\mathbf{Z_t}^c)g_k^{(2)}(\mathbf{Z_t}^{*c},\mathbf{Z_t}^d)(\mathbf{Z_i}^c-\mathbf{Z_t}^c)'}_{\psi_{2nti}}$

$= C_{411k}+C_{412k}$.

Following C_{32k}, for $\phi_{1nti}=\psi_{1nti}+\psi_{1nit}$, when $t\neq i$,

$$\begin{aligned} C_{411k} = & \frac{1}{n}\sum_t E(\phi_{1nti}|\mathbf{Z_t},\epsilon_t) - \frac{1}{2}E\phi_{1nti} + O_p(n^{-1}(E\phi_{1nti}^2)^{1/2}) \\ = & O_p(n^{-1}(E\phi_{1nti}^2)^{1/2}) = O_p(n^{-1}h^{-l_c/2}) = o_p(n^{-1/2}), \end{aligned}$$

since $E(\epsilon_t|\mathbf{Z_t}) = 0$ and $E(e_{ki}|\mathbf{Z_i}) = 0$ and $E\phi_{1nti}^2 \leq CE\psi_{1nti}^2 \leq CE(\frac{\epsilon_t^2 K_I^2(\mathbf{Z_i}-\mathbf{Z_t})}{h^{2l_c} f^2(\mathbf{Z_t})} e_{ki}^2) = O(h^{-l_c})$ by assumption A4.
When $t = i$, $C_{411k} = \frac{K_I(0)}{nh^{l_c}}\frac{1}{n}\sum_t \frac{\epsilon_t e_{kt}}{f(\mathbf{Z_t})} = O_p(n^{-1}h^{-l_c}) = o_p(n^{-1/2})$.
Similarly, for $\phi_{2nti} = \psi_{2nti} + \psi_{2nit}$, when $t \neq i$,

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

Since $E(\epsilon_t|\mathbf{Z_t}) = 0$, $E\phi_{2nti} = 0$. $E\phi_{2nti}^2 \leq CE\psi_{2nti}^2 = O(h^{-l_c+4})$ by assumption A4, so $O_p(n^{-1}(E\phi_{2nti}^2)^{1/2}) = o_p(n^{-1/2})$. Furthermore, $E(\phi_{2nti}|\mathbf{Z_t},\epsilon_t) = E(\psi_{2nti}|\mathbf{Z_t},\epsilon_t) = \frac{\epsilon_t}{f(\mathbf{Z_t})}\int K(\psi)\frac{1}{2}h^2\psi g_k^{(2)}(\mathbf{Z_t}^c+\lambda h\psi, \mathbf{Z_t}^d)\psi' f(\mathbf{Z_t}^c + h\psi, \mathbf{Z_t}^d)d\psi = \frac{\epsilon_t}{f(\mathbf{Z_t})}O(h^s)$, $E(E(\psi_{2nti}|\mathbf{Z_t},\epsilon_t)) = 0$, $E(E(\psi_{2nti}|\mathbf{Z_t},\epsilon_t))^2 = O(h^{2s})$, so $\frac{1}{n}\sum_t E(\phi_{2nti}|\mathbf{Z_t},\epsilon_t) = O_p(n^{-1/2}h^s) = o_p(n^{-1/2})$. So $C_{412k} = o_p(n^{-1/2})$ and $C_{41k} = o_p(n^{-1/2})$.

$$\begin{aligned} C_{5k} = & \frac{1}{n}\sum_t [g_{1,k}(Z_{1t}) - \hat{g}_{1,k}(Z_{1t})][\epsilon_t - \hat{E}(\epsilon|Z_{1t})] \\ = & \frac{1}{n}\sum_t [g_{1,k}(Z_{1t}) - \hat{g}_{1,k}(Z_{1t})]\epsilon_t + O_p(lnz_1(\frac{ln(n)}{nh^{l_c}})^{1/2}) \\ = & C_{51k} + O_p(lnz_1(\frac{ln(n)}{nh^{l_c}})^{1/2}). \\ C_{51k} = & -\frac{1}{n^2}\sum_{t=1}^{n}\sum_{i=1}^{n}\frac{\epsilon_t K_{1I}(Z_{1i}-Z_{1t})}{h_1^{l_{1c}} f_1(Z_{1t})}[e_{1,ki} + \frac{1}{2}(Z_{1i}^c - Z_{1t}^c)g_{1,k}^{(2)}(Z_{1t}^{*c}, Z_{1t}^d)(Z_{1i}^c - Z_{1t}^c)']. \end{aligned}$$

We follow similar argument as in C_{41k} to obtain $C_{51k} = o_p(n^{-1/2})$.

$C_{6k} = \frac{1}{n}\sum_t [g_k(\mathbf{Z_t}) - g_{1,k}(Z_{1t})]\epsilon_t - \frac{1}{n}\sum_t [g_k(\mathbf{Z_t}) - g_{1,k}(Z_{1t})]\hat{E}(\epsilon|Z_{1t})] = C_{61k} - C_{62k}$.
Since $\hat{E}(\epsilon|Z_{1t}) = \frac{1}{nh_1^{l_{1c}} f_1(Z_{1t})}\sum_{i=1}^{n} K_{1I}(Z_{1i} - Z_{1t})\epsilon_i(1 + O_p(h_1))$, with similar arguments as in C_{411k},

$C_{62k} = \frac{1}{n^2}\sum\sum_{t\neq i}\frac{W_{t,k}}{h_1^{l_{1c}} f_1(Z_{1t})} K_{1I}(Z_{1i} - Z_{1t})\epsilon_i(1 + O_p(h_1)) = o_p(n^{-1/2})$.

$C_{61k} = \frac{1}{n}\sum_t W_{t,k}\epsilon_t$. Since $E[W_{t,k}\epsilon_t] = 0$, $E[W_{t,k}\epsilon_t] = E\sigma^2(\mathbf{Z_t})(g_k(\mathbf{Z_t}) - g_{1,k}(Z_{1t}))^2$, by Central Limit Theorem (Lindeberg-Lévy), with assumption A4, we have

$$\sqrt{n}C_{61k} \xrightarrow{d} N(0, E\sigma^2(\mathbf{Z_t})W_{t,k}^2).$$

Finally with the Cramer-Rao device, we obtain

$$\sqrt{n}C_6 \xrightarrow{d} N(0, \Phi_0).$$

Theorem 3: *Proof.*
Let's define the infeasible estimator
$\hat{\beta}^I = (\hat{W}'\Omega^{-1}(\vec{Z}_1)\check{X})^{-1}\hat{W}'\Omega^{-1}(\vec{Z}_1)(Y - \hat{E}(Y|\vec{Z}_1))$, where the true $\sigma^2(Z_{1t})$ is known in $\Omega^{-1}(\vec{Z}_1)$.
In the following we show
(1) $\sqrt{n}(\hat{\beta}^I - \beta) \xrightarrow{d} N(0, (E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t)^{-1})$.
(2) $\sup_{Z_{1t}\in G_1}|\hat{\sigma}^2(Z_{1t}) - \sigma^2(Z_{1t})| = O_p((\frac{nh_1^{l_{1c}}}{ln(n)})^{-\frac{1}{2}}) + O_p(h_1^{s_1}) + O_p(n^{-\frac{1}{2}})$.
Result (2) might be of use by itself. Here repeated use of (2) enables us to obtain
(3) $\sqrt{n}(\hat{\beta}^I - \hat{\beta}^H) = o_p(1)$.
The conclusion of Theorem 3 follows from (1) and (3).

$$\begin{aligned} (1)\ \hat{\beta}^I - \beta & = [(\frac{1}{n}\hat{W}'\Omega^{-1}(\vec{Z}_1)\check{X})^{-1} - (E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t)^{-1} + (E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t)^{-1}] \\ & \times \underbrace{\frac{1}{n}\hat{W}'\Omega^{-1}(\vec{Z}_1)(Y - \hat{E}(Y|\vec{Z}_1))}_{C} \end{aligned}$$

(a) The $(i,j)th$ element of $\frac{1}{n}\hat{W}'\Omega^{-1}(\vec{Z}_1)\check{X}$ is

$\frac{1}{n}\sum_{t=1}^n \frac{1}{\sigma^2(Z_{1t})}\hat{W}_{t,i}\hat{X}_{t,j}$
$\frac{1}{n}\sum_{t=1}^n \frac{1}{\sigma^2(Z_{1t})}\hat{W}_{t,i}(X_{t,j}-\hat{g}_j(\mathbf{Z_t})) + \frac{1}{n}\sum_{t=1}^n \frac{1}{\sigma^2(Z_{1t})}\hat{W}_{t,i}\hat{W}_{t,j}.$

$$\begin{aligned}
& \frac{1}{n}\sum_{t=1}^n \frac{1}{\sigma^2(Z_{1t})}\hat{W}_{t,i}\hat{W}_{t,j} \\
= & \frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[\hat{g}_i(\mathbf{Z_t})-g_i(\mathbf{Z_t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[\hat{g}_i(\mathbf{Z_t})-g_i(\mathbf{Z_t})][g_{1,j}(Z_{1t})-\hat{g}_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[\hat{g}_i(\mathbf{Z_t})-g_i(\mathbf{Z_t})][g_j(\mathbf{Z_t})-g_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_{1,i}(Z_{1t})-\hat{g}_{1,i}(Z_{1t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_{1,i}(Z_{1t})-\hat{g}_{1,i}(Z_{1t})][g_{1,j}(Z_{1t})-\hat{g}_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_{1,i}(Z_{1t})-\hat{g}_{1,i}(Z_{1t})][g_j(\mathbf{Z_t})-g_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})][g_{1,j}(Z_{1t})-\hat{g}_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})][g_j(\mathbf{Z_t})-g_{1,j}(Z_{1t})] \\
= & A_1 + A_2 + \cdots + A_9
\end{aligned}$$

$$\begin{aligned}
& \frac{1}{n}\sum_{t=1}^n \frac{1}{\sigma^2(Z_{1t})}\hat{W}_{t,i}(X_{t,j}-\hat{g}_j(\mathbf{Z_t})) = \frac{1}{n}\sum_{t=1}^n \frac{1}{\sigma^2(Z_{1t})}\hat{W}_{t,i}(g_j(\mathbf{Z_t})-\hat{g}_j(\mathbf{Z_t})+e_{jt}) \\
= & (-1)\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[\hat{g}_i(\mathbf{Z_t})-g_i(\mathbf{Z_t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})] + \frac{1}{n}\frac{1}{\sigma^2(Z_{1t})}\sum_t[\hat{g}_i(\mathbf{Z_t})-g_i(\mathbf{Z_t})]e_{jt} \\
& -\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_{1,i}(Z_{1t})-\hat{g}_{1,i}(Z_{1t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})] + \frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_{1,i}(Z_{1t})-\hat{g}_{1,i}(Z_{1t})]e_{jt} \\
& -\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})][\hat{g}_j(\mathbf{Z_t})-g_j(\mathbf{Z_t})] + \frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})]e_{jt} \\
= & -A_1 + A_{10} - A_4 + A_{11} - A_7 + A_{12}.
\end{aligned}$$

Since Z_{1t} is iid, $\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})} \xrightarrow{p} E\frac{1}{\sigma^2(Z_{1t})} < \infty$ by A6(1), we follow the proof of Theorem 1 to obtain $A_i = o_p(1)$, $i = 1, \cdots, 8, 10, 11, 12$,

$$A_9 - E\frac{1}{\sigma^2(Z_{1t})}[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})][g_j(\mathbf{Z_t})-g_{1,j}(Z_{1t})] = o_p(1),$$

provided $E\frac{1}{\sigma^2(Z_{1t})}[g_i(\mathbf{Z_t})-g_{1,i}(Z_{1t})][g_j(\mathbf{Z_t})-g_{1,j}(Z_{1t})] < \infty$, which is true given A6(1) and A4(1). So together we have $\frac{1}{n}\hat{W}'\Omega^{-1}(\vec{Z}_1)\check{X} - E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t = o_p(1)$. By $A6(2)$ and Slutsky' Theorem, $(\frac{1}{n}\hat{W}'\Omega^{-1}(\vec{Z}_1)\check{X})^{-1} - (E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t)^{-1} = o_p(1)$.

(b) Similarly, for $k = 1, 2, \cdots, K$, the kth element of C is

$$\begin{aligned}
C_k = & \frac{1}{n}\sum_{t=1}^n \frac{1}{\sigma^2(Z_{1t})}\hat{W}_{t,k}(m(Z_{1t})-\hat{E}(m(z_1)|Z_{1t})+\epsilon_t-\hat{E}(\epsilon|Z_{1t})) \\
= & \frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[\hat{g}_k(\mathbf{Z_t})-g_k(\mathbf{Z_t})][m(Z_{1t})-\hat{E}(m(z_1)|Z_{1t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_{1,k}(Z_{1t})-\hat{g}_{1,k}(Z_{1t})][m(Z_{1t})-\hat{E}(m(z_1)|Z_{1t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_k(\mathbf{Z_t})-g_{1,k}(Z_{1t})][m(Z_{1t})-\hat{E}(m(z_1)|Z_{1t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[\hat{g}_k(\mathbf{Z_t})-g_k(\mathbf{Z_t})][\epsilon_t-\hat{E}(\epsilon|Z_{1t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_{1,k}(Z_{1t})-\hat{g}_{1,k}(Z_{1t})][\epsilon_t-\hat{E}(\epsilon|Z_{1t})] \\
& +\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_k(\mathbf{Z_t})-g_{1,k}(Z_{1t})][\epsilon_t-\hat{E}(\epsilon|Z_{1t})] \\
= & C_{1k} + C_{2k} + \cdots + C_{6k}
\end{aligned}$$

Since $\frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})} \xrightarrow{p} E\frac{1}{\sigma^2(Z_{1t})} < \infty$, we follow proof of Theorem 1 to obtain $C_{ik} = o_p(n^{-\frac{1}{2}})$ for $i = 1, 2, 3, 4, 5$ with the additional assumption A6(1).

$$\begin{aligned}
C_{6k} = & \frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_k(\mathbf{Z_t})-g_{1,k}(Z_{1t})]\epsilon_t - \frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}[g_k(\mathbf{Z_t})-g_{1,k}(Z_{1t})]\hat{E}(\epsilon|Z_{1t}) \\
= & C_{61k} + o_p(n^{-1/2})
\end{aligned}$$

following similar arguments as in C_{6k} in Theorem 1. Again.

$C_{61k} = \frac{1}{n}\sum_t \frac{1}{\sigma^2(Z_{1t})}W_{t,k}\epsilon_t$. Since $E[\frac{1}{\sigma^2(Z_{1t})}W_{t,k}\epsilon_t] = 0$, $E[\frac{1}{\sigma^2(Z_{1t})}W_{t,k}\epsilon_t]^2 = E\frac{1}{\sigma^2(\mathbf{Z_t})}(g_k(\mathbf{Z_t}) - g_{1,k}(Z_{1t}))^2$, by Central Limit Theorem (Lindeberg-Lévy), with assumption A6, we have

$$\sqrt{n}C_{61k} \xrightarrow{d} N(0, E\frac{W_{t,k}^2}{\sigma^2(Z_{1t})}).$$

Finally with the Cramer-Rao device, for $C_6 = [C_{61}, C_{62}, \cdots, C_{6K}]'$, we obtain

$$\sqrt{n}C_6 \xrightarrow{d} N(0, E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t).$$

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

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

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

$$\begin{aligned}\tilde{\epsilon}_t^2 = \; & (m(Z_{1t}) - \hat{E}(m(z_1)|Z_{1t}))^2 + (\epsilon_t - \hat{E}(\epsilon|Z_{1t}))^2 + ((X_t - \hat{E}(X|Z_{1t}))(\beta - \tilde{\beta}))^2 \\ & +2(m(Z_{1t}) - \hat{E}(m(z_1)|Z_{1t}))(\epsilon_t - \hat{E}(\epsilon|Z_{1t})) \\ & +2(m(Z_{1t}) - \hat{E}(m(z_1)|Z_{1t}))(X_t - \hat{E}(X|Z_{1t}))(\beta - \tilde{\beta}) \\ & +2(\epsilon_t - \hat{E}(\epsilon|Z_{1t}))(X_t - \hat{E}(X|Z_{1t}))(\beta - \tilde{\beta}) \\ & = I_1 + \cdots + I_6.\end{aligned}$$

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

$$\begin{aligned} & \hat{E}(I_1|Z_{1t}) \\ = \; & \tfrac{1}{nh_1^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^n K_1(\tfrac{Z_{1i}^c - Z_{1t}^c}{h_1})I(Z_{1i}^d = Z_{1t}^d)(m(Z_{1i}) - \hat{E}(m(z_1)|Z_{1i}))^2(1 + O_p(h_1)) \\ \leq \; & \sup_{Z_{1i}\in G_1}|m(Z_{1i}) - \hat{E}(m(z_1)|Z_{1i})|^2[\tfrac{1}{f_1(Z_{1t})}(1 + O_p(h_1))] \underbrace{\frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^n |K_1(\frac{Z_{1i}^c - Z_{1t}^c}{h_1})|I(Z_{1i}^d = Z_{1t}^d)}_{I_{11}}.\end{aligned}$$

We notice that $|K_1(\cdot)|$ satisfies the Lipschitz condition given assumption A3. So with assumption A6(3), we apply Lemma 1 and obtain $\sup_{Z_{1t}\in G_1}|I_{11} - EI_{11}| = O_p((\frac{nh_1^{l_{1c}}}{lnn})^{-\frac{1}{2}})$.
$EI_{11} \to f_1(Z_{1t})\int |K_1(\psi)|d\psi < \infty$ uniformly in $Z_{1t} \in G_1$.
So $I_{11} = O_p(1)$ uniformly. With result in (2)(a) in Theorem 1, we conclude
$\sup_{Z_{1t}\in G_1}|\hat{E}(I_1|Z_{1t})| = (O_p(h_1^2(\frac{nh_1^{l_{1c}}}{lnn})^{-\frac{1}{2}}) + O(h_1^{s_1}))^2 = o_p(n^{-\frac{1}{2}})$ with assumption A5.

$$\begin{aligned} & \hat{E}(I_3|Z_{1t}) \\ = \; & [(1 + O_p(h_1))\tfrac{1}{f_1(Z_{1t})}]\tfrac{1}{nh_1^{l_{1c}}}\sum_{i=1}^n K_1(\tfrac{Z_{1i}^c - Z_{1t}^c}{h_1})I(Z_{1i}^d = Z_{1t}^d)\sum_{k=1}^K\sum_{k'=1}^K (X_{i,k} - \hat{g}_{1,k}(Z_{1i})) \\ & \times(X_{i,k'} - \hat{g}_{1,k'}(Z_{1i}))(\beta_k - \tilde{\beta}_k)(\beta_{k'} - \tilde{\beta}_{k'}) \\ = \; & O_p(n^{-1})[(1 + O_p(h_1))\tfrac{1}{f_1(Z_{1t})}]\sum_{k=1}^K\sum_{k'=1}^K \tfrac{1}{nh_1^{l_{1c}}}\sum_{i=1}^n K_1(\tfrac{Z_{1i}^c - Z_{1t}^c}{h_1})I(Z_{1i}^d = Z_{1t}^d) \\ & \times[e_{1,ki}e_{1,k'i} + (g_{1,k}(Z_{1i}) - \hat{g}_{1,k}(Z_{1i}))e_{1,k'i} + (g_{1,k'}(Z_{1i}) - \hat{g}_{1,k'}(Z_{1i}))e_{1,ki} \\ & +(g_{1,k}(Z_{1i}) - \hat{g}_{1,k}(Z_{1i}))(g_{1,k'}(Z_{1i}) - \hat{g}_{1,k'}(Z_{1i}))] \\ = \; & O_p(n^{-1})[(1 + O_p(h_1))\tfrac{1}{f_1(Z_{1t})}]\sum_{k=1}^K\sum_{k'=1}^K I_{31} = O_p(n^{-1})[(1 + O_p(h_1))\tfrac{1}{f_1(Z_{1t})}]\sum_{k=1}^K\sum_{k'=1}^K [I_{311} + \cdots + I_{314}].\end{aligned}$$

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

$$I_{312} \leq \sup_{Z_{1i}\in G_1}|(g_{1,k}(Z_{1i} - \hat{g}_{1,k}(Z_{1i})| \underbrace{\frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^n |K_1(\frac{Z_{1i}^c - Z_{1t}^c}{h_1})|I(Z_{1i}^d = Z_{1t}^d)|e_{1,k'i}|}_{I_{3121}},$$

Since with assumption A2(3),A4(1) and A6(3), we apply Lemma 1 to have $\sup_{Z_{1t}\in G_1}|I_{3121}| = O_p(1)$. So $I_{312} = o_p(1)$ uniformly in $Z_{1t} \in G_1$. Similarly, $I_{313} = o_p(1)$ uniformly in $Z_{1t} \in G_1$.

$I_{311} = \frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n} K_1(\frac{Z_{1i}^c - Z_{1t}^c}{h_1}) I(Z_{1i}^d = Z_{1t}^d) e_{1,ki} e_{1,k'i}$.

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

So in all, we have $I_{31} = O_p(1)$ uniformly in $Z_{1t} \in G_1$ and $\hat{E}(I_3|Z_{1t}) = O_p(n^{-1})$ uniformly.

$$\begin{aligned}
\hat{E}(I_4|Z_{1t}) = \quad & [(1+O_p(h_1))\tfrac{1}{f_1(Z_{1t})}]\tfrac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n} K_1(\tfrac{Z_{1i}^c - Z_{1t}^c}{h_1}) I(Z_{1i}^d = Z_{1t}^d) 2(m(Z_{1i}) - \hat{E}(m(z_1)|Z_{1i})) \\
& \times(\epsilon_i - \hat{E}(\epsilon|Z_{1i})). \\
\leq \quad & [O_p(h_1^2(\tfrac{nh_1^{l_{1c}}}{lnn})^{-\frac{1}{2}}) + O_p(h_1^{s_1})]\underbrace{[\frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}|K_1(\frac{Z_{1i}^c - Z_{1t}^c}{h_1})| I(Z_{1i}^d = Z_{1t}^d)|\epsilon_i|]}_{I_{41}} + o_p(n^{-1/2})
\end{aligned}$$

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 obtain $\sup_{Z_{1t}\in G_1}|I_{41}| = O_p(1)$ and $\sup_{Z_{1t}\in G_1}|\hat{E}(I_4|Z_{1t})| = O_p(h_1^2(\frac{nh_1^{l_{1c}}}{lnn})^{-\frac{1}{2}}) + O_p(h_1^{s_1})$.

$$\begin{aligned}
\hat{E}(I_5|Z_{1t}) = \quad & [(1+O_p(h_1))\tfrac{1}{f_1(Z_{1t})}]\tfrac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n} K_1(\tfrac{Z_{1i}^c - Z_{1t}^c}{h_1}) I(Z_{1i}^d = Z_{1t}^d) 2(m(Z_{1i}) - \hat{E}(m(Z_{1i})|Z_{1i})) \\
& \times\sum_{k=1}^{K}(X_{i,k} - \hat{g}_{1,k}(Z_{1i}))(\beta_k - \tilde{\beta}_k) \\
\leq \quad & [O_p(h_1^2(\tfrac{nh_1^{l_{1c}}}{lnn})^{-\frac{1}{2}}) + O_p(h_1^{s_1})]O_p(n^{-\frac{1}{2}}) \\
& \times\sum_{k=1}^{K}\{\underbrace{\frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}|K_1(\frac{Z_{1i}^c - Z_{1t}^c}{h_1})| I(Z_{1i}^d = Z_{1t}^d)|X_{i,k} - g_{1,k}(Z_{1i})|}_{I_{51}} \\
& \underbrace{+\frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}|K_1(\frac{Z_{1i}^c - Z_{1t}^c}{h_1})| I(Z_{1i}^d = Z_{1t}^d)|g_{1,k}(Z_{1i}) - \hat{g}_{1,k}(Z_{1i})|}_{I_{52}}\}
\end{aligned}$$

We obtain easily that $\sup_{Z_{1t}\in G_1}|I_{52}| = o_p(1)$ from result on I_{11} and $\sup_{Z_{1t}\in G_1}|I_{51}| = O_p(1)$ as argued for term I_{3121}. So we conclude $\sup_{Z_{1t}\in G_1}|\hat{E}(I_5|Z_{1t})| = o_p(n^{-\frac{1}{2}})$.

$$\begin{aligned}
\hat{E}(I_6|Z_{1t}) = \quad & [(1+O_p(h_1))\tfrac{1}{f_1(Z_{1t})}]\tfrac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n} K_1(\tfrac{Z_{1i}^c - Z_{1t}^c}{h_1}) I(Z_{1i}^d = Z_{1t}^d)\sum_{k=1}^{K} 2(\epsilon_i - \hat{E}(\epsilon_i|Z_{1i})) \\
& \times(X_{i,k} - \hat{g}_{1,k}(Z_{1i}))(\beta_k - \tilde{\beta}_k) \\
\leq \quad & O_p(n^{-\frac{1}{2}})\sum_{k=1}^{K}[\tfrac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}|K_1(\tfrac{Z_{1i}^c - Z_{1t}^c}{h_1})| I(Z_{1i}^d = Z_{1t}^d)|\epsilon_i||X_{i,k} - g_{1,k}(Z_{1i})| \\
& + \sup_{Z_{1i}\in G_1}|g_{1,k}(Z_{1i}) - \hat{g}_{1,k}(Z_{1i})|\tfrac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}|K_1(\tfrac{Z_{1i}^c - Z_{1t}^c}{h_1})| I(Z_{1i}^d = Z_{1t}^d)|\epsilon_i| \\
& + \sup_{Z_{1i}\in G_1}|\hat{E}(\epsilon|Z_{1i})|\tfrac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}|K_1(\tfrac{Z_{1i}^c - Z_{1t}^c}{h_1})| I(Z_{1i}^d = Z_{1t}^d)|X_{i,k} - g_{1,k}(Z_{1i})| \\
& + \sup_{Z_{1i}\in G_1}|\hat{E}(\epsilon|Z_{1i})|\sup_{Z_{1i}\in G_1}|g_{1,k}(Z_{1i}) - \hat{g}_{1,k}(Z_{1i})|\tfrac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}|K_1(\tfrac{Z_{1i}^c - Z_{1t}^c}{h_1})| I(Z_{1i}^d = Z_{1t}^d)|] \\
= \quad & O_p(n^{-\frac{1}{2}})[I_{61} + \cdots + I_{64}]
\end{aligned}$$

It is easy to see that $\sup_{Z_{1t}\in G_1}|I_{64}| = o_p(1)$ as in term I_{11}. Similarly we have $\sup_{Z_{1t}\in G_1}|I_{62}| = o_p(1)$ and $\sup_{Z_{1t}\in G_1}|I_{63}| = o_p(1)$ with results on I_{41} and I_{3121}.

$$I_{61} \leq \frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}|K_1(\frac{Z_{1i}^c-Z_{1t}^c}{h_1})|I(Z_{1i}^d=Z_{1t}^d)|\epsilon_i||X_{i,k}| + \frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}|K_1(\frac{Z_{1i}^c-Z_{1t}^c}{h_1})|I(Z_{1i}^d=Z_{1t}^d)|\epsilon_i||g_{1,k}(Z_{1i})|$$
$$= I_{611}+I_{612}$$

With assumption A6(3), $E(|\epsilon_i|^{2+\delta_1}|X_{i,k}|^{2+\delta_1}|Z_{1i}) \leq [E(|\epsilon_i|^{4+2\delta_1}|Z_{1i})E(|X_{i,k}|^{4+2\delta_1}|Z_{1i})]^{\frac{1}{2}} < \infty$, so we use A6(3) and apply Lemma 1 to obtain $\sup_{Z_{1t}\in G_1}|I_{611}| = O_p(1)$.

$I_{612} \leq C\frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}|K_1(\frac{Z_{1i}^c-Z_{1t}^c}{h_1})|I(Z_{1i}^d = Z_{1t}^d)|\epsilon_i| = cI_{41}$, so we conclude $\sup_{Z_{1t}\in G_1}|I_{612}| = O_p(1)$. So $\sup_{Z_{1t}\in G_1}|I_{61}| = O_p(1)$ and in all $\sup_{Z_{1t}\in G_1}|\hat{E}(I_6|Z_{1t})| = O_p(n^{-\frac{1}{2}})$.

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

$$\hat{E}(I_2|Z_{1t}) = [(1+O_p(h_1))\frac{1}{f_1(Z_{1t})}]\frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}K_1(\frac{Z_{1i}^c-Z_{1t}^c}{h_1})I(Z_{1i}^d=Z_{1t}^d)\{\epsilon_i^2-2\epsilon_i\hat{E}(\epsilon|Z_{1i})+(\hat{E}(\epsilon|Z_{1i}))^2\}$$
$$= \frac{1}{f_1(Z_{1t})}\frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}K_1(\frac{Z_{1i}^c-Z_{1t}^c}{h_1})I(Z_{1i}^d=Z_{1t}^d)\epsilon_i^2 + O_p((\frac{nh_1^{l_{1c}}}{lnn})^{-\frac{1}{2}}).$$

(b) So we have from above

$$\hat{\sigma}^2(Z_{1t}) = \hat{E}(\tilde{\epsilon}^2|Z_{1t})$$
$$= [(1+O_p(h_1))\frac{1}{f_1(Z_{1t})}]\underbrace{\frac{1}{nh_1^{l_{1c}}}\sum_{i=1}^{n}K_1(\frac{Z_{1i}^c-Z_{1t}^c}{h_1})I(Z_{1i}^d=Z_{1t}^d)\epsilon_i^2}_{I} + O_p(n^{-\frac{1}{2}}) + O_p((\frac{nh_1^{l_{1c}}}{ln(n)})^{-\frac{1}{2}}) + O_p(h_1^{s_1}).$$

With A6(3) and A4(4), we apply Lemma 1 to obtain $\sup_{Z_{1t}\in G_1}|I-EI| = O_p((\frac{nh_1^{l_{1c}}}{ln(n)})^{-\frac{1}{2}})$. With a change of variable and using A6(1) and A2(1),

$$EI = \int K_1(\psi)\sigma^2(Z_{1t}^c+h_1\psi, Z_{1t}^d)f_1(Z_{1t}^c+h_1\psi, Z_{1t}^d)d\psi$$
$$= \int K_1(\psi)[\sigma^2(Z_{1t}) + \sum_{|j|=1}^{s_1}\frac{\partial^j}{\partial(Z_{1t}^c)^j}\sigma^2(Z_{1t})\frac{h_1^{|j|}\psi^j}{j!}$$
$$+ \sum_{|j|=s_1}(\frac{\partial^j}{\partial(Z_{1t}^c)^j}\sigma^2(Z_{1t}^c*, Z_{1t}^d) - \frac{\partial^j}{\partial(Z_{1t}^c)^j}\sigma^2(Z_{1t}))\frac{h_1^{|j|}\psi^j}{j!}][f_1(Z_{1t}) + \sum_{|l|=1}^{s_1}\frac{\partial^l}{\partial(Z_{1t}^c)^l}f_1(Z_{1t})\frac{h_1^{|l|}\psi^l}{l!}$$
$$+ \sum_{|l|=s_1}(\frac{\partial^l}{\partial(Z_{1t}^c)^l}f_1(Z_{1t}^c*, Z_{1t}^d) - \frac{\partial^l}{\partial(Z_{1t}^c)^l}f_1(Z_{1t}))\frac{h_1^{|l|}\psi^l}{l!}]d\psi$$
$$= \sigma^2(Z_{1t})f_1(Z_{1t}) + O(h_1^{s_1}), \text{ with the additional assumption A6(1).}$$

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

Note $\hat{E}(I_i|Z_{1t}) = o_p(n^{-1/2})$ for $i = 1, 3, 5$,

$\hat{E}(I_2|Z_{1t}) = \frac{1}{nh_1^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n}K_{1I}(Z_{1i}-Z_{1t})(\epsilon_i^2 - 2\epsilon_i\hat{E}(\epsilon|Z_{1i}))(1+o_p(1)) + o_p(n^{-1/2})$,

$\hat{E}(I_4|Z_{1t}) = \frac{1}{nh_1^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n}K_{1I}(Z_{1i}-Z_{1t})2(m(Z_{1i} - \hat{E}(m(z_1)|Z_{1i})\epsilon_i + o_p(n^{-1/2})$,

$\hat{E}(I_6|Z_{1t}) = \frac{1}{nh_1^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n}K_{1I}(Z_{1i}-Z_{1t})2\epsilon_i\sum_{k=1}^{K}e_{1,ki}(\beta_k-\hat{\beta}_k) + o_p(n^{-1/2})$,

$$\hat{\sigma}^2(Z_{1t}) - \sigma^2(Z_{1t})$$
$$= \{\frac{1}{nh_1^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n}K_{1I}(Z_{1i}-Z_{1t})[\epsilon_i^2-\sigma^2(Z_{1i})]$$
$$+\frac{1}{nh_1^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n}K_{1I}(Z_{1i}-Z_{1t})[\sigma^2(Z_{1i})-\sigma^2(Z_{1t})-(Z_{1i}^c-Z_{1t}^c)\frac{\partial\sigma^2(Z_{1t})}{\partial(z_1^c)'}]$$
$$-\frac{2}{nh_1^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n}K_{1I}(Z_{1i}-Z_{1t})\epsilon_i\frac{1}{nh_1^{l_{1c}}f_1(Z_{1i})}\sum_{j=1}^{n}K_{1I}(Z_{1j}-Z_{1i})\epsilon_j$$
$$-\frac{2}{nh_1^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n}K_{1I}(Z_{1i}-Z_{1t})\epsilon_i h_1^{s_1}DFM_n(Z_{1i}) - \frac{2}{nh_1^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n}K_{1I}(Z_{1i}-Z_{1t})\epsilon_i VFM_n(Z_{1i})$$
$$+\frac{1}{nh_1^{l_{1c}}f_1(Z_{1t})}\sum_{i=1}^{n}K_{1I}(Z_{1i}-Z_{1t})2\epsilon_i\sum_{k=1}^{K}e_{1,ki}(\beta_k-\hat{\beta}_k)\}(1+o_p(1)) + o_p(n^{-1/2})$$
$$= \{S_1(Z_{1t})+\cdots+S_6(Z_{1t})\}(1+o_p(1)) + o_p(n^{-1/2}).$$ We use this alternative expression in (3).}

$$
\begin{array}{rl}
(3) & \sqrt{n}(\hat\beta^I - \hat\beta^H) \\
= & \sqrt{n}\{[(\hat W'\Omega^{-1}(\vec Z_1)\check X)^{-1} - (\hat W'\hat\Omega^{-1}(\vec Z_1)\check X)^{-1}]\hat W'\Omega^{-1}(\vec Z_1)(Y - \hat E(Y|\vec Z_1)) \\
+ & (\hat W'\hat\Omega^{-1}(\vec Z_1)\check X)^{-1}\hat W'[\Omega^{-1}(\vec Z_1) - \hat\Omega^{-1}(\vec Z_1)](Y - \hat E(Y|\vec Z_1))\}
\end{array}
$$

So we show

$$
\begin{array}{ll}
(a) & (\frac{1}{n}\hat W'\Omega^{-1}(\vec Z_1)\check X)^{-1} - (\frac{1}{n}\hat W'\hat\Omega^{-1}(\vec Z_1)\check X)^{-1} = o_p(1). \\
(b) & \sqrt{n}\frac{1}{n}\hat W'[\Omega^{-1}(\vec Z_1) - \hat\Omega^{-1}(\vec Z_1)](Y - \hat E(Y|\vec Z_1)) = o_p(1).
\end{array}
$$

Since in (1) we have $\sqrt{n}\frac{1}{n}\hat W'\Omega^{-1}(\vec Z_1)(Y - \hat E(Y|\vec Z_1)) = O_p(1)$, and with (a) $(\frac{1}{n}\hat W'\hat\Omega^{-1}(\vec Z_1)\check X)^{-1} \xrightarrow{p} (E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t)^{-1}$, $E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t$ is positive definite, the claim of (3) follows from (a) and (b).

We first note $\sup_{Z_{1t}\in G_1}|\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}|$
$\le [\inf_{Z_{1t}\in G_1}\sigma^2(Z_{1t})\inf_{Z_{1t}\in G_1}\hat\sigma^2(Z_{1t})]^{-1}\sup_{Z_{1t}\in G_1}|\hat\sigma^2(Z_{1t}) - \sigma^2(Z_{1t})|$.
With result (2) and A6(1), for large n, $\inf_{Z_{1t}\in G_1}\hat\sigma^2(Z_{1t}) > 0$, so

$$
\sup_{Z_{1t}\in G_1}|\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}| = O_p((\frac{nh_1^{l_{1c}}}{ln(n)})^{-\frac{1}{2}}) + O_p(h_1^{s_1}) + O_p(n^{-\frac{1}{2}}).
$$

(a) Since $\frac{1}{n}\hat W'\Omega^{-1}(\vec Z_1)\check X \xrightarrow{p} E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t$, which is positive definite, so by Slutsky's Theorem, $(\frac{1}{n}\hat W'\Omega^{-1}(\vec Z_1)\check X)^{-1} \xrightarrow{p} (E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t)^{-1}$.
If (a') $\frac{1}{n}\hat W'(\Omega^{-1}(\vec Z_1) - \hat\Omega^{-1}(\vec Z_1))\check X = o_p(1)$, then $\frac{1}{n}\hat W'\hat\Omega^{-1}(\vec Z_1)\check X \xrightarrow{p} E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t$ as well, and $(\frac{1}{n}\hat W'\hat\Omega^{-1}(\vec Z_1)\check X)^{-1} \xrightarrow{p} (E\frac{1}{\sigma^2(Z_{1t})}W_t'W_t)^{-1}$ so we have the claim in (a). So we only need to show (a').
The $(i,j)th$ element in $\frac{1}{n}\hat W'(\Omega^{-1}(\vec Z_1) - \hat\Omega^{-1}(\vec Z_1))\check X$ is

$$
\begin{array}{rl}
& \frac{1}{n}\sum_{t=1}^n \hat W_{t,i}(\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})})\check X_{t,j} \\
= & \frac{1}{n}\sum_{t=1}^n \hat W_{t,i}(\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})})(e_{jt} + g_j(\mathbf{Z_t}) - \hat g_j(\mathbf{Z_t})) + \frac{1}{n}\sum_{t=1}^n \hat W_{t,i}(\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})})\hat W_{t,j}.
\end{array}
$$

$$
\begin{array}{rl}
& \frac{1}{n}\sum_{t=1}^n \hat W_{t,i}(\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})})\hat W_{t,j} \\
= & \frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][\hat g_i(\mathbf{Z_t}) - g_i(\mathbf{Z_t})][\hat g_j(\mathbf{Z_t}) - g_j(\mathbf{Z_t})] \\
& +\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][\hat g_i(\mathbf{Z_t}) - g_i(\mathbf{Z_t})][g_{1,j}(Z_{1t}) - \hat g_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][\hat g_i(\mathbf{Z_t}) - g_i(\mathbf{Z_t})][g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][g_{1,i}(Z_{1t}) - \hat g_{1,i}(Z_{1t})][\hat g_j(\mathbf{Z_t}) - g_j(\mathbf{Z_t})] \\
& +\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][g_{1,i}(Z_{1t}) - \hat g_{1,i}(Z_{1t})][g_{1,j}(Z_{1t}) - \hat g_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][g_{1,i}(Z_{1t}) - \hat g_{1,i}(Z_{1t})][g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][g_i(\mathbf{Z_t}) - g_{1,i}(Z_{1t})][\hat g_j(\mathbf{Z_t}) - g_j(\mathbf{Z_t})] \\
& +\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][g_i(\mathbf{Z_t}) - g_{1,i}(Z_{1t})][g_{1,j}(Z_{1t}) - \hat g_{1,j}(Z_{1t})] \\
& +\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][g_i(\mathbf{Z_t}) - g_{1,i}(Z_{1t})][g_j(\mathbf{Z_t}) - g_{1,j}(Z_{1t})] \\
= & A_1 + \cdots + A_9
\end{array}
$$

$$
\begin{array}{rl}
& \frac{1}{n}\sum_{t=1}^n \hat W_{t,i}(\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})})(e_{jt} + g_j(\mathbf{Z_t}) - \hat g_j(\mathbf{Z_t})) \\
= & \frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][\hat g_i(\mathbf{Z_t}) - g_i(\mathbf{Z_t})]e_{jt} \\
& -\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][\hat g_i(\mathbf{Z_t}) - g_i(\mathbf{Z_t})][\hat g_j(\mathbf{Z_t}) - g_j(\mathbf{Z_t})] \\
& +\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][g_{1,i}(Z_{1t}) - \hat g_{1,i}(Z_{1t})]e_{jt} \\
& -\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][g_{1,i}(Z_{1t}) - \hat g_{1,i}(Z_{1t})][\hat g_j(\mathbf{Z_t}) - g_j(\mathbf{Z_t})] \\
& +\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][g_i(\mathbf{Z_t}) - g_{1,i}(Z_{1t})]e_{jt} \\
& -\frac{1}{n}\sum_t[\frac{1}{\sigma^2(Z_{1t})} - \frac{1}{\hat\sigma^2(Z_{1t})}][g_i(\mathbf{Z_t}) - g_{1,i}(Z_{1t})][\hat g_j(\mathbf{Z_t}) - g_j(\mathbf{Z_t})] \\
= & A_{10} - A_1 + A_{11} - A_4 + A_{12} - A_7.
\end{array}
$$

Since $\sup_{Z_{1t}\in G_1}|\frac{1}{\sigma^2(Z_{1t})}-\frac{1}{\hat\sigma^2(Z_{1t})}|=o_p(1)$, 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 kth element in $\frac{1}{n}\hat W'[\Omega^{-1}(\vec Z_1)-\hat\Omega^{-1}(\vec Z_1)](Y-\hat E(Y|\vec Z_1))$ is

$$\begin{aligned}C_k &= \tfrac{1}{n}\textstyle\sum_t[\frac{1}{\sigma^2(Z_{1t})}-\frac{1}{\hat\sigma^2(Z_{1t})}][\hat g_k(\mathbf{Z_t})-g_k(\mathbf{Z_t})][m(Z_{1t})-\hat E(m(z_1)|Z_{1t})]\\ &+\tfrac{1}{n}\textstyle\sum_t[\frac{1}{\sigma^2(Z_{1t})}-\frac{1}{\hat\sigma^2(Z_{1t})}][g_{1,k}(Z_{1t})-\hat g_{1,k}(Z_{1t})][m(Z_{1t})-\hat E(m(z_1)|Z_{1t})]\\ &+\tfrac{1}{n}\textstyle\sum_t[\frac{1}{\sigma^2(Z_{1t})}-\frac{1}{\hat\sigma^2(Z_{1t})}][g_k(\mathbf{Z_t})-g_{1,k}(Z_{1t})][m(Z_{1t})-\hat E(m(z_1)|Z_{1t})]\\ &+\tfrac{1}{n}\textstyle\sum_t[\frac{1}{\sigma^2(Z_{1t})}-\frac{1}{\hat\sigma^2(Z_{1t})}]\hat W_{t,k}\epsilon_t\\ &+\tfrac{1}{n}\textstyle\sum_t[\frac{1}{\sigma^2(Z_{1t})}-\frac{1}{\hat\sigma^2(Z_{1t})}]\hat W_{t,k}\hat E(\epsilon|Z_{1t})\\ &= C_{1k}+\cdots+C_{5k}.\end{aligned}$$

With $\sup_{Z_{1t}\in G_1}|\frac{1}{\sigma^2(Z_{1t})}-\frac{1}{\hat\sigma^2(Z_{1t})}|=O_p(lnz_1)$, and Theorem 1 proof (1)(d), (e), (2)(d), we easily have $\sqrt{n}C_{ik}=o_p(1)$ for $i=1,2,3,5$.

$$\begin{aligned}C_{4k} &= \tfrac{1}{n}\textstyle\sum_t[\frac{1}{\sigma^2(Z_{1t})}-\frac{1}{\hat\sigma^2(Z_{1t})}][\hat g_k(\mathbf{Z_t})-g_k(\mathbf{Z_t})+g_{1,k}(Z_{1t})-\hat g_{1,k}(Z_{1t})+g_k(\mathbf{Z_t})-g_{1,k}(Z_{1t})]\epsilon_t\\ &= \tfrac{1}{n}\textstyle\sum_t[\frac{1}{\sigma^2(Z_{1t})}-\frac{1}{\hat\sigma^2(Z_{1t})}]W_{t,k}\epsilon_t+o_p(n^{-1/2})\\ &= \tfrac{1}{n}\textstyle\sum_t\frac{\hat\sigma^2(Z_{1t})-\sigma^2(Z_{1t})}{\sigma^4(Z_{1t})}W_{t,k}\epsilon_t+o_p(n^{-1/2})\\ &= \tfrac{1}{n}\textstyle\sum_t\frac{S_1(Z_{1t})+\cdots+S_6(Z_{1t})}{\sigma^4(Z_{1t})}W_{t,k}\epsilon_t+o_p(n^{-1/2})\\ &= D_{1k}+\cdots+D_{6k},\end{aligned}$$

where the last three equalities uses the result (2) above and $\sup_{Z_{1t}\in G_1}|\frac{1}{\sigma^2(Z_{1t})}-\frac{1}{\hat\sigma^2(Z_{1t})}|=O_p(lnz_1)$.

We show below that $D_{ik}=o_p(n^{-1/2})$ for $i=1,\cdots,6$, which implies $C_{4k}=o_p(n^{-1/2})$.

$$D_{1k}=\tfrac{1}{n}\textstyle\sum_t\frac{S_1(Z_{1t})}{\sigma^4(Z_{1t})}W_{t,k}\epsilon_t=\frac{1}{n^2}\sum_{t=1}^n\sum_{i=1}^n\underbrace{\frac{W_{t,k}\epsilon_t}{h_1^{l_{1c}}f_1(Z_{1t})\sigma^4(Z_{1t})}K_{1I}(Z_{1i}-Z_{1t})(\epsilon_i^2-\sigma^2(Z_{1i}))}_{\psi_{nti}}.$$

We follow the argument in C_{32k} of Theorem 1 to perform H-decomposition to obtain $D_{1k}=\frac{1}{n}\sum_t E(\phi_{nti}|\mathbf{Z_t},\epsilon_t)-\frac{1}{2}E\phi_{nti}+O_p(n^{-1}(E\phi_{nti}^2)^{1/2})=O_p(n^{-1}(E\phi_{nti}^2)^{1/2})=O_p(n^{-1}h_1^{-l_{1c}/2})$, since $E\phi_{tni}^2\le CE\psi_{tni}^2\le\frac{C}{h_1^{2l_c}}E[\frac{W_{t,k}^2\sigma^2(Z_{1t})}{\sigma^8(Z_{1t})f_1^2(Z_{1t})}K_{1I}^2(Z_{1i}-Z_{1t})(E(\epsilon_i^4|Z_{1i})+\sigma^4(Z_{1i}))]=O(h_1^{-l_{1c}})$.

$$\begin{aligned}D_{2k} &= \tfrac{1}{n}\textstyle\sum_t\frac{S_2(Z_{1t})}{\sigma^4(Z_{1t})}W_{t,k}\epsilon_t\\ &= \frac{1}{n^2}\sum_{t=1}^n\sum_{i=1}^n\underbrace{\frac{W_{t,k}\epsilon_t}{h_1^{l_{1c}}f_1(Z_{1t})\sigma^4(Z_{1t})}K_{1I}(Z_{1i}-Z_{1t})(\frac{1}{2}(Z_{1i}^c-Z_{1t}^c)\frac{\partial^2\sigma^2(Z_{1t}^{*c},Z_{1t}^d)}{\partial(z_1^c)'\partial z_1^c}(Z_{1i}^c-Z_{1t}^c)')}_{\psi_{nti}}.\end{aligned}$$

Similarly, $D_{2k}=\frac{1}{n}\sum_t E(\phi_{nti}|\mathbf{Z_t},\epsilon_t)+O_p(n^{-1}(E\phi_{nti}^2)^{1/2})$. $E\phi_{nti}^2=O(h_1^{-2l_{1c}+4})$ and

$$\begin{aligned}E(\phi_{nti}|\mathbf{Z_t},\epsilon_t) &= E(\psi_{nti}|\mathbf{Z_t},\epsilon_t)\\ &= \frac{h_1^2W_{t,k}\epsilon_t}{\sigma^4(Z_{1t})f_1(Z_{1t})}\int K_1(\psi_1)\tfrac{1}{2}\psi_1\frac{\partial^2\sigma^2(Z_{1t}^c+\lambda h_1\psi_1,Z_{1t}^d)}{\partial(z_1^c)'\partial z_1^c}\psi_1'f_1(Z_{1t}^c+h_1\psi_1,Z_{1t}^d)d\psi_1\\ &= \frac{W_{t,k}\epsilon_t}{\sigma^4(Z_{1t})f_1(Z_{1t})}O(h_1^{s_1}).\end{aligned}$$

So $E(E(\phi_{nti}|\mathbf{Z_t},\epsilon_t))=0$ and $E[E(\phi_{nti}|\mathbf{Z_t},\epsilon_t)]^2=O(h_1^{2s_1})$, so $\frac{1}{n}\sum_t E(\phi_{nti}|\mathbf{Z_t},\epsilon_t)=O_p(n^{-1/2}h_1^{s_1})$. So in all, $D_{2k}=o_p(n^{-1/2})$.

$$\begin{aligned}D_{3k} &= \tfrac{1}{n}\textstyle\sum_t\frac{S_3(Z_{1t})}{\sigma^4(Z_{1t})}W_{t,k}\epsilon_t\\ &= -2*\frac{1}{n^3}\sum_{t=1}^n\sum_{i=1}^n\sum_{j=1}^n\underbrace{\frac{W_{t,k}\epsilon_t\epsilon_i\epsilon_j}{h_1^{2l_{1c}}f_1(Z_{1t})f_1(Z_{1i})\sigma^4(Z_{1t})}K_{1I}(Z_{1i}-Z_{1t})K_{1I}(Z_{1j}-Z_{1i})}_{\psi_{ntij}}.\end{aligned}$$

When $t\neq i\neq j$, let $\phi_{ntij}=\psi_{ntij}+\psi_{ntji}+\psi_{nitj}+\psi_{nijt}+\psi_{njti}+\psi_{njit}$,

$$D_{3k}=(-2)\tfrac{1}{6}[\tfrac{6}{n^3}-\binom{n}{3}^{-1}+\binom{n}{3}^{-1}]\sum\sum_{t<i<j}\sum\phi_{ntij}.$$

$u_n = \left(\begin{array}{c} n \\ 3 \end{array} \right)^{-1} \sum\sum_{t<i<j}\sum \phi_{ntij}$ is a third degree U-statistics. We apply Theorem 1 of Yao and Martins-Filho (2013) to determine its order of magnitude. Using notations there,

$$u_n = \theta_n + \sum_{j=1}^{3} \left(\begin{array}{c} 3 \\ j \end{array} \right) H_n^{(j)}(w_{v_1}, \cdots, w_{v_j}), \text{ with } w = (Z, \epsilon) \ ,$$

where $H_n^{(j)} = O_p((n^{-j}\sigma_{jn}^2)^{1/2})$. Here $\theta_n = \sigma_{1n}^2 = \sigma_{2n}^2 = 0$.

$\sigma_{3n}^2 \leq CE\psi_{ntij}^2 = O(h_1^{-2l_{1c}})$, so $H_n^{(3)} = O_p(n^{-3/2}h_1^{-l_{1c}}) = o_p(n^{-1/2})$. 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.

$$D_{4k} = \frac{1}{n}\sum_t \frac{S_4(Z_{1t})}{\sigma^4(Z_{1t})} W_{t,k}\epsilon_t = -2 * \frac{h_1^{s_1}}{n^2} \sum_{t=1}^{n}\sum_{i=1}^{n} \underbrace{\frac{W_{t,k}\epsilon_t}{h_1^{l_{1c}} f_1(Z_{1t})\sigma^4(Z_{1t})} K_{1I}(Z_{1i} - Z_{1t})\epsilon_i DFM_n(Z_{1i})}_{\psi_{nti}} .$$

We show in a similar fashion that $C_{4k} = (-2h_1^{s_1})O_p(n^{-1}h_1^{-l_{1c}/2}) = o_p(n^{-1/2})$ when $t \neq i$ and $C_{4k} = o_p(n^{-1/2}h_1^{s_1}) = o_p(n^{-1/2})$ when $t = i$. So $C_{4k} = o_p(n^{-1/2})$.

$$\begin{aligned} D_{5k} = & \ \frac{1}{n}\sum_t \frac{S_5(Z_{1t})}{\sigma^4(Z_{1t})} W_{t,k}\epsilon_t \\ = & \ -\frac{1}{n^3}\sum_{t=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{W_{t,k}\epsilon_t\epsilon_i K_{1I}(Z_{1i}-Z_{1t})}{h_1^{2l_{1c}} f_1(Z_{1t}) f_1(Z_{1i})\sigma^4(Z_{1t})} [K_{1I}(Z_{1j} - Z_{1i})(Z_{1j}^c - Z_{1i}^c)m^{(2)}(Z_{1i}^{*c}, Z_{1i}^d)(Z_{1j}^c - Z_{1i}^c)' \\ & -E(K_{1I}(Z_{1j} - Z_{1i})(Z_{1j}^c - Z_{1i}^c)m^{(2)}(Z_{1i}^{*c}, Z_{1i}^d)(Z_{1j}^c - Z_{1i}^c)'|\mathbf{Z_i})] \\ = & \ o_p(n^{-1/2}) \text{ can be shown with similar arguments.} \\ D_{6k} = & \ \frac{1}{n}\sum_t \frac{S_6(Z_{1t})}{\sigma^4(Z_{1t})} W_{t,k}\epsilon_t = 2\sum_{k=1}^{K}(\beta_k - \hat{\beta}_k)\frac{1}{n^2}\sum_{t=1}^{n}\sum_{i=1}^{n} \frac{W_{t,k}\epsilon_t}{h_1^{l_{1c}} f_1(Z_{1t})\sigma^4(Z_{1t})} K_{1I}(Z_{1i} - Z_{1t})\epsilon_i e_{1,ki} \\ = & \ o_p(n^{-1/2}) \text{ with similar arguments.} \end{aligned}$$

References

[1] Ai, C. and X. Chen, 2003, Efficient estimation of models with conditional moment restrictions containing unknown functions. Econometrica, 71, 1795-1843.

[2] Angrist, J. D., and A. B. Krueger, 1991, Does compulsory school attendance affect schooling and earnings. Quarterly Journal of Economics, 106, 979-1014.

[3] Angrist, J. D., and A. B. Krueger, 1995, Split-sample instrumental variables estimates of the return to schooling. Journal of Business & Economic Statistics, 13, 225-235.

[4] Baltagi, B. H. and Q. Li, 2002, On instrumental variable estimation of semiparametric dynamic panel data models. Economics Letters, 76, 1-9.

[5] Bickel, P., 1982, On adaptive estimation. Annals of Statistics, 10, 647-671.

[6] Blundell, R., and J. L. Powell, 2003, Endogeneity in semiparametric and nonparametric regression models, in Advances in economics and econometrics: theory and applications. Vol. 2, ed. by M. Dewatripont, L. P. Hansen, S. J. Turnovsky. Cambridge, U.K., Cambridge university press, 312-357.

[7] Blundell, R., X. Chen, and D. Kristensen, 2007, Semi-nonparametric IV estimation of shape-invariant Engle curves. Econometrica, 75, 1613-1669.

[8] Camlong-Viot, C., J. M. Rodríguez-Póo, and P. Vieu, 2006, Nonparametric and semiparametric estimation of additive models with both discrete and continuous variables under dependence. The art of semiparametrics, 155-178, Physica-Verlag HD.

[9] Card, D., 1995, Using geographic variation in college proximity to estimate the return to schooling, in Aspects of labour market behavior: essays in honour of John Vanderkamp. Ed. L. N. Christophides, E. K. Grant and R. Swidinsky, 201-222. Toronto: University of Toronto Press.

[10] Card, D., 2001, Estimating the return to schooling: progress on some persistent econometric problems. Econometrica, 69, 1127-1160,

[11] Chen, X. and D. Pouzo, 2012, Estimation of nonparametric conditional moment models with possibly nonsmooth generalized residuals. Econometrica, 80, 277-321.

[12] Chamberlain, G., 1992, Efficiency bounds for semiparametric regression. Econometrica, 60, 567-596.

[13] Darolles, S., Y. Fan, J.-P. Florens, and E. Renault, 2011, Nonparametric Instrumental Regression. Econometrica, 79, 1541-1566.

[14] Davidson, R., and J. G. MacKinnon, 2004, Econometric theory and methods. Oxford university press, New York.

[15] Deschenes, O., 2007, Estimating the effects of family background on the return to schooling. Journal of Business & Economic Statistics, 25, 265-277.

[16] Delgado, M.A., and J. Mora, 1995, Nonparametric and semiparametric estimation with discrete regressors. Econometrica, 63, 1477-1484.

[17] Fan, J., 1992, Design adaptive nonparametric regression. Journal of the American Statistical Association, 87, 998-1004.

[18] Fan, J., W. Härdle and E. Mammen, 1998, Direct estimation of low dimensional components in additive models. The Annals of Statistics, 26, 943-971.

[19] Florens, J.-P.,J. Johanns, and S. Van Bellegem, 2011, Identification and estimation by penalization in nonparametric instrumental regression. Econometric Theory, 27, 472-496.

[20] Griliches, Z., 1977, Estimating the returns to schooling: some econometric problems. Econometrica, 45, 1-22.

[21] Härdle, W., H. Liang, and J. Gao, 2000, Partially linear models. Physica-Verlag.

[22] Imbens, G. W. and W. K. Newey, 2009, Identification and estimation of triangular simultaneous equations models without additivity. Econometrica, 77, 1481-1512.

[23] Kling, J. R., 2001, Interpreting instrumental variables estimates of the returns to schooling. Journal of Business & Economic Statistics, 19, 358-364

[24] Lewbel, A., 2012, Using heteroskedasticity to identify and estimate mismeasured and endogenous regressor models. Journal of Business & Economic Statistics, 30, 67-80.

[25] Li, Q. and J. Racine, 2007, Nonparametric econometrics: theory and practice. Princeton University Press, Princeton, NJ.

[26] Li, Q., and T. Stengos, 1996, Semiparametric estimation of partially linear panel data models. Journal of Econometrics, 71, 389-397.

[27] Manski, C. F., 1984, Adaptive estimation of nonlinear regression models. Econometric Reviews, 3, 145-194.

[28] Martins-Filho, C. and F. Yao, 2007, Nonparametric frontier estimation via local linear regression. Journal of Econometrics, 141, 283-319.

[29] Matzkin, R.L.,1994, Restrictions of economic theory in nonparametric methods. In Handbook of Econometrics: IV, edited by R. F. Engle, and D. L. McFadden, Elsevier Science, North Holland.

[30] Matzkin, R.L., 2008, Identification in nonparametric simultaneous equations. Econometrica, 76, 945-978.

[31] Newey, W. K., 1990, Efficient instrumental variable estimation of nonlinear models. Econometrica, 58, 809-837.

[32] Newey, W. K., 1993, Efficient estimation of models with conditional moment restrictions. In Handbook of Statistics, v 11, 419-454, G.S. Maddala, C.R. Rao, and H.D. Vinod, eds., Amsterdam: North-Holland.

[33] Otsu, T., 2011. Empirical likelihood estimation of conditional moment restriction models with unknown functions. Econometric Theory 27, 8-46.

[34] Racine, J. and Q. Li, 2004, Nonparametric estimation of regression functions with both categorical and continuous data. Journal of Econometrics, 119, 99-130.

[35] Robinson, P., 1988, Root-n-consistent semiparametric regression. Econometrica, 56, 931-954.

[36] Speckman, P., 1988, Kernel smoothing in partial linear models. Journal of the Royal Statistical Society, Series B, 50, 413-436.

[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.

[38] Ruppert, D., S. J. Sheather, M. P. Wand, 1995, An effective bandwidth selector for local least squares regression. Journal of the American Statistical Association, 90, 1257-1270.

[39] Theil, H., 1953, Repeated least squares applied to complete equation systems. The Hague: central planning bureau.

[40] White, H., 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica, 48, 817-838.

[41] White, H., 1982, Instrumental variables regression with independent observations. Econometrica, 50, 483-499.

[42] Yao, F. and C. Martins-Filho, 2013, An asymptotic characterization of finite degree U-statistics with sample size dependent kernels: applications to nonparametric estimators and test statistics, working paper, Economics department, West Virginia University.