Identifying the Average Treatment Effect in a Two Threshold Model - Supplemental Appendix

Arthur Lewbel and Thomas Tao Yang[2]

Boston College

This online appendix provides proofs for Theorem 2.4, 2.6, 2.8 and associated lemmas, which are asymptotics for the following estimators in the paper,

$$\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\widehat{\mathrm{E}}\left(\left.\frac{D_iY_i}{f(v_i|x_i)}\right|x_i\right)}{\widehat{\mathrm{E}}\left(\left.\frac{D_i}{f(v_i|x_i)}\right|x_i\right)}-\frac{\widehat{\mathrm{E}}\left(\left.\frac{(1-D_i)Y_i}{f(v_i|x_i)}\right|x_i\right)}{\widehat{\mathrm{E}}\left(\left.\frac{1-D_i}{f(v_i|x_i)}\right|x_i\right)}\right], \tag{1}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\widehat{\mathrm{E}}\left(\left.\frac{D_iY_i}{\widehat{f}(v_i|x_i)}\right|x_i\right)}{\widehat{\mathrm{E}}\left(\left.\frac{D_i}{\widehat{f}(v_i|x_i)}\right|x_i\right)}-\frac{\widehat{\mathrm{E}}\left(\left.\frac{(1-D_i)Y_i}{\widehat{f}(v_i|x_i)}\right|x_i\right)}{\widehat{\mathrm{E}}\left(\left.\frac{1-D_i}{\widehat{f}(v_i|x_i)}\right|x_i\right)}\right], \tag{2}$$

$$\frac{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{D_{it}Y_{it}}{\widehat{f}_{v_t}(v_{it})}}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{D_{it}}{\widehat{f}_{v_t}(v_{it})}}-\frac{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{(1-D_{it})Y_{it}}{\widehat{f}_{v_t}(v_{it})}}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{(1-D_{it})}{\widehat{f}_{v_t}(v_{it})}}. \tag{3}$$

Lemma 4.1 *Suppose random variables* $\underset{1\times 1}{s_i}$, $\underset{k\times 1}{x_i}$ *are i.i.d. across* $i=1,2,...,n$, *and* s_i *is bounded.* $r(x_i)$ *is a bounded real function of* x_i. *The following Lipschitz conditions hold:*

$$|E(s_i|x_i+e_x)-E(s_i|x_i)|\le M_1\|e_x\|,$$

$$|r(x_i+e_x)-r(x_i)|\le M_2\|e_x\|,$$

$$|f_x(x_i+e_x)-f_x(x_i)|\le M_3\|e_x\|$$

for some positive number M_1, M_2, *and* M_3. $r_i(x_i)$, $E(s_i|x_i)$, $f_x(x_i)$ *are p-th order differentiable and the p-th order derivatives are bounded.*

[2] Corresponding Author: Arthur Lewbel, Department of Economics, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA, 02467, USA. (617)-552-3678, lewbel@bc.edu, https://www2.bc.edu/~lewbel/. Thomas Tao Yang, Department of Economics, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA, 02467, USA. yangta@bc.edu.

Let

$$
\begin{aligned}
U_n &= \frac{1}{n}\sum_{i=1}^{n}\left\{r(x_i)\widehat{E}\left(s_i|x_i\right)\widehat{f}(x_i) - E\left[r(x_i)s_j\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)\right]\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{r(x_i)\left[\frac{1}{(n-1)h^k}\sum_{j=1,j\neq i}^{n} s_jK\left(\frac{x_j-x_i}{h}\right)\right] - E\left[r(x_i)s_j\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)\right]\right\},
\end{aligned}
$$

where $K(.)$ *is the kernel function defined in Assumption 4.3.* $n^{1-\varepsilon}h^{2k+2} \to \infty$, *and* $nh^{2p} \to 0$, *as* $n \to \infty$, *for a very small* $\varepsilon > 0$. *Define term*

$$
\widetilde{U}_n = \frac{1}{n}\sum_{i=1}^{n}\left\{\left[r(x_i)s_i + r(x_i)E(s_i|x_i)\right]f_x(x_i) - 2E\left[r(x_i)E(s_i|x_i)f_x(x_i)\right]\right\},
$$

then $\sqrt{n}(U_n - \widetilde{U}_n) = o_p(1)$. *Furthermore, since* $K\left(\frac{x_j-x_i}{h}\right)$ *is a p-th order kernel, and by assumption that* $nh^{2p} \to 0$, $E\Big[r(x_i)s_j\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)\Big]$ *in* U_n *could be replaced by* $E[r(x_i)E(s_i|x_i)f_x(x_i)]$ *and the conclusion still holds.*

The Lemma is just a modification of Lemma 3.2 in Powell et al. (1988). It is pretty straightforward, given the property of U-statistics and uniformly convergence property of the nonparametric regression. The Lipschitz condition is imposed to control the residual term; the existence and boundedness conditions are imposed to control the biased term. Estimator (1) is consist of several terms like U_n and a residual term of the order $o_p\left(\frac{1}{\sqrt{n}}\right)$.

Proof of Lemma 4.1. Rewrite the first term in U_n, indexed by U_{1n} as

$$
\begin{aligned}
U_{1n} &= \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} r(x_i)s_j\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right) \\
&= \frac{2}{n(n-1)}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{1}{2}\left(r(x_i)s_j + r(x_j)s_i\right)\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)
\end{aligned}
\tag{4}
$$

Index $z_i = (s_i, x_i)$, and define $p(z_i, z_j) = \frac{1}{2}\left[r(x_i)s_j + r(x_j)s_i\right]\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)$. Since the kernel function we adopt is symmetric, $p(z_i, z_j) = p(z_j, z_i)$. It is easy to see that $\mathrm{E}(\|p(z_i, z_j)\|^2) = o(n)$. According to Lemma 3.2 in Powell et al. (1989),

$$
U_{1n} - \widetilde{U}_{1n} = o_p\left(\frac{1}{\sqrt{n}}\right), \tag{5}
$$

where

$$\widetilde{U}_{1n} = \mathrm{E}\left[p(z_j, z_i)\right] + \frac{2}{n}\sum_{i=1}^{n}\left[\mathrm{E}\left(p(z_j, z_i)|z_i\right) - \mathrm{E}\left(p(z_j, z_i)\right)\right]. \qquad (6)$$

$E\left[p(z_j, z_i)|z_i\right]$ is

$$\begin{aligned}
\mathrm{E}\left[p(z_j, z_i)|z_i\right] &= \int_{\mathbb{R}^k} \frac{1}{2}\left[r(x_i)\mathrm{E}(s_j|x_j) + r(x_j)s_i\right] \frac{1}{h^k} K\left(\frac{x_j - x_i}{h}\right) f_x(x_j)dx_j \\
&= \int_{\mathbb{R}^k} \frac{1}{2}\left[r(x_i)\mathrm{E}(s_i|x_i + hu) + r(x_i + hu)s_i\right] K(u) f_x(x_i + hu)du \\
&= \frac{1}{2}\left[r(x_i)\mathrm{E}(s_i|x_i) + r(x_i)s_i\right] f_x(x_i) + R_i,
\end{aligned}$$

where second equality is obtained via variable transformation $u = \frac{x_j - x_i}{h}$, and R_i is the residual term, which could be decomposed into two components R_{1i} and R_{2i},

$$\begin{aligned}
R_i &= R_{1i} - R_{2i} \\
&= \int_{\mathbb{R}^k} \frac{1}{2}\left[r(x_i)\mathrm{E}(s_i|x_i + hu) + r(x_i + hu)s_i - r(x_i)\mathrm{E}(s_i|x_i) - r(x_i)s_i\right] K(u) f_x(x_i)du \\
&\quad - \int_{\mathbb{R}^k} \frac{1}{2}\left[r(x_i)\mathrm{E}(s_i|x_i) + r(x_i)s_i\right] K(u)\left[f_x(x_i + hu) - f_x(x_i)\right] du.
\end{aligned}$$

$\mathrm{E}[p(z_j, z_i)|z_i] - \mathrm{E}[p(z_j, z_i)]$ in Equation (6) thus could be written as

$$\begin{aligned}
& \mathrm{E}\left[p(z_j, z_i)|z_i\right] - \mathrm{E}\left[p(z_j, z_i)\right] \qquad (7) \\
=\ & \frac{1}{2}\left[r(x_i)\mathrm{E}(s_i|x_i) + r(x_i)s_i\right] f_x(x_i) - \mathrm{E}\left[\frac{1}{2}\left[r(x_i)\mathrm{E}(s_i|x_i) + r(x_i)s_i\right] f_x(x_i)\right] \\
& + R_{1i} - \mathrm{E}(R_{1i}) - \left[R_{2i} - \mathrm{E}(R_{2i})\right] \\
=\ & \frac{1}{2}\left[r(x_i)\mathrm{E}(s_i|x_i) + r(x_i)s_i\right] f_x(x_i) - \mathrm{E}\left[r(x_i)\mathrm{E}(s_i|x_i)f_x(x_i)\right] + R_{1i} - \mathrm{E}(R_{1i}) - \left[R_{2i} - \mathrm{E}(R_{2i})\right].
\end{aligned}$$

Then under Lipschitz conditions and bound conditions given in the Lemma,

$$\mathrm{E}(R_{1i}^2) = O(h^2),$$

$$\mathrm{E}\left(R_{2i}^2\right) = O(h^2).$$

Thus by Lindberg Levy CLT

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}[R_{1i}-\mathrm{E}(R_{1i})]-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}[R_{2i}-\mathrm{E}(R_{2i})]\xrightarrow{p}0. \tag{8}$$

Equation (5), (6), (7), and (8) imply that $U_{1n}-\mathrm{E}[p(z_j,z_i)]$ is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\{[r(x_i)\mathrm{E}(s_i|x_i)+r(x_i)s_i]f_x(x_i)-2\mathrm{E}\left[r(x_i)\mathrm{E}(s_i|x_i)f_x(x_i)\right]\}+o_p\left(\frac{1}{\sqrt{n}}\right),$$

which is the first conclusion of the Lemma.

Second conclusion follows by the assumption that $nh^{2p}\to 0$ and those bound conditions and by the fact that

$$\mathrm{E}\left[r(x_i)s_j\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)\right]-\mathrm{E}\left[r(x_i)\mathrm{E}(s_i|x_i)f_x(x_i)\right]=o_p(h^p),$$

the equality holds by the assumption that kernel function is of order p. ■

Proof of Theorem 2.4. The last conclusion of this theorem follows immediately after the first conclusion via Lindeberg central limit theorem. For the first conclusion, it is enough to show that $\frac{1}{n}\sum_{i=1}^{n}\left[\widehat{\psi}_1(x_i)-\mathrm{E}(Y_1)\right]$ is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\left[\psi_1(x_i)+\frac{h_{1i}}{\mathrm{E}(g_{1i}|x_i)}-\frac{\mathrm{E}(h_{1i}|x_i)g_{1i}}{[\mathrm{E}(g_{1i}|x_i)]^2}-\mathrm{E}(Y_1)\right]+o_p\left(\frac{1}{\sqrt{n}}\right). \tag{9}$$

Note that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\left[\widehat{\psi}_1(x_i)-\psi_1(x_i)\right] &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\widehat{\mathrm{E}}(h_{1i}|x_i)}{\widehat{\mathrm{E}}(g_{1i}|x_i)}-\frac{\mathrm{E}(h_{1i}|x_i)}{\mathrm{E}(g_{1i}|x_i)}\right]\\ &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\widehat{\mathrm{E}}(\widetilde{h}_{1i}|x_i)}{\widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i)}-\frac{\mathrm{E}(\widetilde{h}_{1i}|x_i)}{\widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i)}+\frac{\mathrm{E}(\widetilde{h}_{1i}|x_i)}{\widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i)}-\frac{\mathrm{E}(\widetilde{h}_{1i}|x_i)}{\mathrm{E}(\widetilde{g}_{1i}|x_i)}\right]\\ &= \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\widehat{\mathrm{E}}(\widetilde{h}_{1i}|x_i)-\mathrm{E}(\widetilde{h}_{1i}|x_i)}{\mathrm{E}(\widetilde{g}_{1i}|x_i)}-\frac{\mathrm{E}(\widetilde{h}_{1i}|x_i)\left[\widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i)-\mathrm{E}(\widetilde{g}_{1i}|x_i)\right]}{\left[\mathrm{E}(\widetilde{g}_{1i}|x_i)\right]^2}+R_{ni}\right],\end{aligned}$$

where the second equality follows by the fact that $\frac{\widehat{\mathrm{E}}(h_{1i}|x_i)}{\widehat{\mathrm{E}}(g_{1i}|x_i)} = \frac{\widehat{\mathrm{E}}(\widetilde{h}_{1i}|x_i)}{\widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i)}$, $\frac{\mathrm{E}(h_{1i}|x_i)}{\mathrm{E}(g_{1i}|x_i)} = \frac{\mathrm{E}(\widetilde{h}_{1i}|x_i)}{\mathrm{E}(\widetilde{g}_{1i}|x_i)}$, and

$$R_{ni} = \frac{\left[\widehat{\mathrm{E}}(\widetilde{h}_{1i}|x_i) - \mathrm{E}(\widetilde{h}_{1i}|x_i)\right]\left[\mathrm{E}(\widetilde{g}_{1i}|x_i) - \widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i)\right]}{\mathrm{E}(\widetilde{g}_{1i}|x_i)\widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i)} - \frac{\mathrm{E}(\widetilde{h}_{1i}|x_i)\left[\mathrm{E}(\widetilde{g}_{1i}|x_i) - \widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i)\right]^2}{\left[\mathrm{E}(\widetilde{g}_{1i}|x_i)\right]^2 \widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i)}.$$

Under Assumption 4.2, $n^{1-\varepsilon}h^{2k+2} \to \infty$ and from Li and Racine (2007) Chapter 2.3

$$\sup\left|\widehat{\mathrm{E}}(\widetilde{h}_{1i}|x_i) - \mathrm{E}(\widetilde{h}_{1i}|x_i)\right| = O_p\left[(n^{1-\varepsilon}h^{k})^{-\frac{1}{2}}\right],$$

$$\sup\left|\widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i) - \mathrm{E}(\widetilde{g}_{1i}|x_i)\right| = O_p\left[(n^{1-\varepsilon}h^{k})^{-\frac{1}{2}}\right],$$

for any arbitrary small $\varepsilon > 0$. So $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sup|R_{ni}|$ is $O_p(\frac{1}{n^{\frac{1}{2}-\varepsilon}h^k})$. Since $n^{1-\varepsilon}h^{2k+2} \to \infty$, $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}|R_{ni}|$ is $o_p(1)$. Thus, $\frac{1}{n}\sum_{i=1}^{n}\left[\widehat{\psi}_1(x_i) - \psi_1(x_i)\right]$ is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\widehat{\mathrm{E}}(\widetilde{h}_{1i}|x_i) - \mathrm{E}(\widetilde{h}_{1i}|x_i)}{\mathrm{E}(\widetilde{g}_{1i}|x_i)} - \frac{\mathrm{E}(\widetilde{h}_{1i}|x_i)\left[\widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i) - \mathrm{E}(\widetilde{g}_{1i}|x_i)\right]}{\left[\mathrm{E}(\widetilde{g}_{1i}|x_i)\right]^2}\right] + o_p\left(\frac{1}{\sqrt{n}}\right). \tag{10}$$

Note that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{\mathrm{E}}(\widetilde{h}_{1i}|x_i)}{\mathrm{E}(\widetilde{g}_{1i}|x_i)} &= \frac{1}{n}\sum_{i=1}^{n}\frac{\frac{1}{(n-1)h^k}\sum_{j=1,j\neq i}^{n}h_{1j}K\left(\frac{x_j-x_i}{h}\right)}{\mathrm{E}(\widetilde{g}_{1i}|x_i)} \\ &= \frac{1}{n}\sum_{i=1}^{n}r_1(x_i)\left[\frac{1}{(n-1)h^k}\sum_{j=1,j\neq i}^{n}s_{1j}K\left(\frac{x_j-x_i}{h}\right)\right].\end{aligned}$$

Under Assumption 4.4, applying Lemma 4.1 by letting $r(x_i) = r_1(x_i), s_j = s_{1j}$, one can get

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{\mathrm{E}}(\widetilde{h}_{1i}|x_i) - \mathrm{E}(\widetilde{h}_{1i}|x_i)}{\mathrm{E}(\widetilde{g}_{1i}|x_i)} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{h_{1i}}{\mathrm{E}(g_{1i}|x_i)} - \mathrm{E}(Y_1)\right] + o_p\left(\frac{1}{\sqrt{n}}\right). \tag{11}$$

Similarly, we can prove that

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\mathrm{E}(\widetilde{h}_{1i}|x_i)\left[\widehat{\mathrm{E}}(\widetilde{g}_{1i}|x_i) - \mathrm{E}(\widetilde{g}_{1i}|x_i)\right]}{\left[\mathrm{E}(\widetilde{g}_{1i}|x_i)\right]^2} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\mathrm{E}(h_{1i}|x_i)g_{1i}}{\left[\mathrm{E}(g_{1i}|x_i)\right]^2} - \mathrm{E}(Y_1)\right] + o_p\left(\frac{1}{\sqrt{n}}\right). \tag{12}$$

Combining Equation (10) (11) (12) gives that

$$\frac{1}{n}\sum_{i=1}^{n}\left[\widehat{\psi}_1(x_i)-\psi_1(x_i)\right]=\frac{1}{n}\sum_{i=1}^{n}\left[\frac{h_{1i}}{\mathrm{E}(g_{1i}|x_i)}-\frac{\mathrm{E}(h_{1i}|x_i)g_{1i}}{\left[\mathrm{E}(g_{1i}|x_i)\right]^2}\right]+o_p\left(\frac{1}{\sqrt{n}}\right).$$

Adding both sides of the above equation with $\frac{1}{n}\sum_{i=1}^{n}\left[\psi_1(x_i)-\mathrm{E}(Y_1)\right]$ gives that $\frac{1}{n}\sum_{i=1}^{n}\left[\widehat{\psi}_1(x_i)-\mathrm{E}(Y_1)\right]$ is equal to Equation (9). The first conclusion is proved. ■

Lemma 4.2 *Assume we observe* $\underset{(k+1)\times 1}{W_i}=(\underset{k\times 1}{X_i}\ \underset{1\times 1}{V_i}),\ \underset{1\times 1}{s_i},\ \underset{(k+2)\times 1}{Z_i}=(\underset{(k+1)\times 1}{W_i}\ \underset{1\times 1}{s_i})$, *which are i.i.d. across i.* $r(x_i)$ *is a real function of* x_i. s_i, $r(x_i)$ *and density function* f_x, f_w *are bounded.* $E(s_i|w_i)$, $r(x_i)$, f_x *and* f_w *are p-th order differentiable, and p-th order derivatives are bounded.* $E(s_i|w_i)$, $r(x_i)$, f_x, f_w *satisfy the Lipschitz condition*

$$|E(s_i|w_i+e_w)-E(s_i|w_i)|\le M_1\left\|e_w\right\|,$$

$$|r(x_i+e_x)-r(x_i)|\le M_2\left\|e_x\right\|,$$

$$|f_x(x_i+e_x)-f_x(x_i)|\le M_3\left\|e_x\right\|,$$

$$|f_w(w_i+e_w)-f_w(w_i)|\le M_4\left\|e_w\right\|$$

for some positive M_1, M_2, M_3, M_4. *Under above assumptions, the following term*

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}r(x_i)\widehat{E}\left(s_i\widehat{f}_w(w_i)\middle|\,x_i\right)\widehat{f}_x(x_i) \qquad (13)\\
=\;&\frac{1}{n}\sum_{i=1}^{n}r(x_i)\left\{\frac{1}{n-1}\sum_{j=1,j\neq i}^{n}\frac{1}{h^k}s_jK\left(\frac{x_j-x_i}{h}\right)\left[\frac{1}{n-1}\sum_{l=1,l\neq j}^{n}\frac{1}{h^{k+1}}K\left(\frac{w_l-w_j}{h}\right)\right]\right\}\\
=\;&\frac{1}{n(n-1)^2}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\sum_{l=1,l\neq j}^{n}r(x_i)s_j\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)\frac{1}{h^{k+1}}K\left(\frac{w_l-w_j}{h}\right)
\end{aligned}$$

is equal to

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\{r(x_i)s_if_w(w_i)f_x(x_i)+r(x_i)f_x(x_i)E\left[s_if_w(w_i)|\,x_i\right]\\
&+r(x_i)E\left(s_i|w_i\right)f_w(w_i)f_x(x_i)-2E[r(x_i)s_if_w(w_i)f_x(x_i)]\}+o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

After decomposing Estimator (2), the influence function of our estimator is consist of several terms like Equation (13). The proof of this lemma looks tedious but it only repeatedly uses the U-statistics results from Powell et al. (1988) and uniformly convergence property of nonparametric estimates.

Proof of Lemma 4.2. Rewrite Equation (13),

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} r(x_i)\widehat{\mathrm{E}}\left(s_i\widehat{f}_w(w_i)\Big| x_i\right)\widehat{f}_x(x_i) \\
=\; & \frac{1}{n(n-1)h^k}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} r(x_i)s_j\widehat{f}_w(w_j)K\left(\frac{x_j-x_i}{h}\right) \\
=\; & \frac{1}{n(n-1)^2h^{2k+1}}\sum_{i=1}^{n} r(x_i)\sum_{j=1,j\neq i}^{n}\sum_{l=1,l\neq j}^{n} s_jK\left(\frac{x_j-x_i}{h}\right)K\left(\frac{w_l-w_j}{h}\right).
\end{aligned}$$

Drop one term in the last summation to make it look more like a U-statistics,

$$\frac{1}{n}\sum_{i=1}^{n} r(x_i)\frac{1}{(n-1)(n-2)h^{2k+1}}\sum_{j=1,j\neq i}^{n}\sum_{l=1,l\neq i,j}^{n} s_jK\left(\frac{x_j-x_i}{h}\right)K\left(\frac{w_l-w_j}{h}\right). \tag{14}$$

Since all terms inside summation are bounded, one term dropped is of the order $o_p(\frac{1}{n^3h^{2k+1}})$, which is of course $o_p(\frac{1}{\sqrt{n}})$ by assumption $nh^{2k+1}\to\infty$.

The structure of proof is as follows.

First we will show that

$$\frac{1}{(n-1)(n-2)h^{2k+1}}\sum_{j=1,j\neq i}^{n}\sum_{l=1,l\neq i,j}^{n} s_jK\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h}) \tag{15}$$

is equal to

$$\begin{aligned}
&\mathrm{E}\left[\frac{s_j}{h^{2k+1}}K\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h})\Big| x_i\right] + \frac{1}{n-1}\sum_{j=1,j\neq i}^{n}\left\{\frac{1}{h^k}s_jK\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\right. \\
&+\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right) - \mathrm{E}\left[\frac{1}{h^k}s_jK\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\Big| x_i\right] \\
&\left.-\mathrm{E}\left[\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\Big| x_i\right] + R_{nij} - \mathrm{E}(R_{nij}|x_i)\right\} + o_p\left(\frac{1}{\sqrt{n}}\right),
\end{aligned} \tag{16}$$

where R_{nij} is some residual term that will be defined later.

Then substitute Equation (16) into Equation (14), the new expression becomes a standard

U-statistics. We then could show the new expression is equal to

$$
\begin{aligned}
&\mathrm{E}\left[r(x_i)s_i f_w(w_i)f_x(x_i)\right] + \frac{1}{n}\sum_{i=1}^{n}[r(x_i)s_i f_w(w_i)f_x(x_i) + r(x_i)\mathrm{E}\left(s_i|w_i\right) f_w(w_i)f_x(x_i) \\
&-2\mathrm{E}\left[r(x_i)s_i f_w(w_i)f_x(x_i)\right]] + o_p\left(\frac{1}{\sqrt{n}}\right),
\end{aligned}
$$

by which the conclusion of the lemma follows immediately.

Step 1: As discussed, we consider first the following term,

$$
\begin{aligned}
&\frac{1}{(n-1)(n-2)h^{2k+1}}\sum_{j=1,j\neq i}^{n}\sum_{l=1,l\neq i,j}^{n} s_j K\left(\frac{x_j-x_i}{h}\right)K\left(\frac{w_l-w_j}{h}\right) \qquad (17)\\
= \;&\frac{2}{(n-1)(n-2)}\sum_{j=1,j\neq i}^{n-1}\sum_{l=j+1,l\neq i}^{n}\frac{1}{2}\left[s_j K\left(\frac{x_j-x_i}{h}\right)+s_l K\left(\frac{x_l-x_i}{h}\right)\right]\frac{1}{h^{2k+1}}K\left(\frac{w_l-w_j}{h}\right).
\end{aligned}
$$

Let

$$
P_1(z_j,z_l;x_i) = \frac{1}{2}\left[s_j K\left(\frac{x_j-x_i}{h}\right)+s_l K\left(\frac{x_l-x_i}{h}\right)\right]\frac{1}{h^{2k+1}}K\left(\frac{w_l-w_j}{h}\right),
$$

then Equation (17) becomes

$$
\frac{2}{(n-1)(n-2)}\sum_{j=1,j\neq i}^{n-1}\sum_{l=j+1,l\neq i}^{n} P_1(z_j,z_l;x_i), \qquad (18)
$$

following Powell et al. (1989), we first verify $\mathrm{E}\big[P_1(z_j,z_l;x_i)^2|x_i\big] = o_p(n)$.

$$
\begin{aligned}
&\mathrm{E}\left[P_1(z_j,z_l;x_i)^2|x_i\right]\\
= \;&\int\int_{\Omega_{w_j,w_l}}\mathrm{E}\left\{\left[\frac{1}{2}\left[s_j K\left(\frac{x_j-x_i}{h}\right)+s_l K\left(\frac{x_l-x_i}{h}\right)\right]\frac{1}{h^{2k+1}}K\left(\frac{w_l-w_j}{h}\right)\right]^2\middle| w_j,w_l\right\}\\
&f_w(w_j)f_w(w_l)\mathrm{d}w_j\mathrm{d}w_l.\\
= \;&\int\int_{\Omega_{u_j,u_l}}\frac{1}{h^{2k+1}}\mathrm{E}\left\{\left[\frac{1}{2}\left[s_j K\left(u_i\right)+s_l K\left(u_j+hu_l\right)\right]K(u_l)\right]^2\middle|(x_i+hu_j,v_j),(x_i+hu_j+hu_l,v_j+hu_l)\right\}\\
&f_w(x_i+hu_j,v_j)f_w\left(x_i+hu_j+hu_l,v_j+hu_l\right)\mathrm{d}u_j\mathrm{d}v_j\mathrm{d}u_l\\
= \;&O_p\left(\frac{1}{h^{2k+1}}\right) = o_p(n),
\end{aligned}
$$

where the second equality holds by variable changes $u_l = \frac{w_l-w_j}{h}$, $u_j = \frac{x_j-x_i}{h}$, the third equality holds because of those bound conditions, and the last equality holds by assumption that $nh^{2k+1} \to$

∞. According to Lemma 3.2 in Powell et al. (1989), Equation (18) is equal to[3]

$$\mathrm{E}\left[P_1(z_j, z_l; x_i)|x_i\right] + \frac{2}{n}\sum_{j=1}^{n}\left\{\mathrm{E}\left[P_1(z_j, z_l; x_i)|z_j, x_i\right] - \mathrm{E}\left[P_1(z_j, z_l; x_i)|x_i\right]\right\} + o_p\left(\frac{1}{\sqrt{n}}\right). \tag{19}$$

Term inside the summation in Equation (19) could be analyzed as follows,

$$\begin{aligned}
& \mathrm{E}\left[P_1(z_j, z_l; x_i)|z_j, x_i\right] - \mathrm{E}\left[P_1(z_j, z_l; x_i)|x_i\right] \\
= & \frac{1}{2}\mathrm{E}\left[\frac{s_j}{h^{2k+1}}K\left(\frac{x_j - x_i}{h}\right)K\left(\frac{w_j - w_l}{h}\right)\middle| z_j, x_i\right] + \frac{1}{2}\mathrm{E}\left[\frac{s_l}{h^{2k+1}}K\left(\frac{x_l - x_i}{h}\right)K\left(\frac{w_j - w_l}{h}\right)\middle| z_j, x_i\right] \\
& -\frac{1}{2}\mathrm{E}\left[\frac{s_j}{h^{2k+1}}K\left(\frac{x_j - x_i}{h}\right)K\left(\frac{w_j - w_l}{h}\right)\middle| x_i\right] - \frac{1}{2}\mathrm{E}\left[\frac{s_l}{h^{2k+1}}K\left(\frac{x_l - x_i}{h}\right)K\left(\frac{w_j - w_l}{h}\right)\middle| x_i\right].
\end{aligned}$$

Since

$$\begin{aligned}
& \mathrm{E}\left[\frac{s_j}{h^{2k+1}}K\left(\frac{x_j - x_i}{h}\right)K\left(\frac{w_j - w_l}{h}\right)\middle| z_j, x_i\right] \\
= & \int_{\Omega_{w_l}} \frac{s_j}{h^{2k+1}}K\left(\frac{x_j - x_i}{h}\right)K\left(\frac{w_j - w_l}{h}\right)f_w(w_l)\mathrm{d}w_l \\
= & \frac{s_j}{h^k}K\left(\frac{x_j - x_i}{h}\right)f_w(w_j) + \frac{s_j}{h^k}K\left(\frac{x_j - x_i}{h}\right)\int_{\Omega_{u_l}} K\left(u_l\right)\left[f_w(w_j + hu_l) - f_w(w_j)\right]\mathrm{d}u_l,
\end{aligned}$$

and similarly

$$\begin{aligned}
& \mathrm{E}\left[\frac{s_l}{h^{2k+1}}K\left(\frac{x_l - x_i}{h}\right)K\left(\frac{w_j - w_l}{h}\right)\middle| z_j, x_i\right] \\
= & \frac{\mathrm{E}\left[s_j|w_j\right]}{h^k}K\left(\frac{x_j - x_i}{h}\right)f_w(w_j) + \frac{1}{h^k}\int_{\Omega_{u_l}}\left[\mathrm{E}\left[s_j|w_j + hu_l\right]K\left(\frac{x_j + hu_l - x_i}{h}\right)f_w(w_j + hu_l)\right. \\
& \left. -\mathrm{E}\left[s_j|w_j\right]K\left(\frac{x_j - x_i}{h}\right)f_w(w_j)\right]K\left(u_l\right)\mathrm{d}u_l,
\end{aligned}$$

the following holds

$$\begin{aligned}
& \mathrm{E}\left[P_1(z_j, z_l; x_i)|z_j, x_i\right] - \mathrm{E}\left[P_1(z_j, z_l; x_i)|x_i\right] \\
= & \frac{1}{2}\frac{s_j}{h^k}K\left(\frac{x_j - x_i}{h}\right)f_w(w_j) + \frac{1}{2}\frac{\mathrm{E}\left[s_j|w_j\right]}{h^k}K\left(\frac{x_j - x_i}{h}\right)f_w(w_j) - \frac{1}{2}\mathrm{E}\left[\frac{s_j}{h^k}K\left(\frac{x_j - x_i}{h}\right)f_w(w_j)\middle| x_i\right] \\
& -\frac{1}{2}\mathrm{E}\left[\frac{\mathrm{E}\left[s_j|w_j\right]}{h^k}K\left(\frac{x_j - x_i}{h}\right)f_w(w_j)\middle| x_i\right] + R_{nij} - \mathrm{E}\left(R_{nij}|x_i\right),
\end{aligned} \tag{20}$$

[3]The argument in Powell et al. (1989) is without conditional expectation. With conditional expectation, everything follows very similarly to the case without conditional expectation.

where

$$R_{nij} = \frac{s_j}{h^k} K\left(\frac{x_j - x_i}{h}\right) \int K(u_l)\left[f_w(w_j + hu_l) - f_w(w_j)\right] \mathrm{d}u_l \tag{21}$$

$$+\frac{1}{h^k}\int_{\Omega_{u_l}} \left[\mathrm{E}\left[s_j|w_j + hu_l\right] K\left(\frac{x_j + hu_l - x_i}{h}\right) f_w(w_j + hu_l)\right.$$

$$\left. -\mathrm{E}\left[s_j|w_j\right] K\left(\frac{x_j - x_i}{h}\right) f_w(w_j) K(u_l)\, \mathrm{d}u_l\right]. \tag{22}$$

Equation (20) is the same as the term inside the summation of Equation (16), together with the fact that $p(z_j, z_l; x_i)$ are symmetric for z_j, z_l conditional on x_i and thus

$$\mathrm{E}\left[P_1(z_j, z_l; x_i)|x_i\right] = \mathrm{E}\left[s_j K\left(\frac{x_j - x_i}{h}\right) K(\frac{w_l - w_j}{h}) \middle| x_i\right],$$

Equation (19) is equal to (16). By the assumption that $r(x_i)$ is bounded, the $o_p(\frac{1}{\sqrt{n}})$ term is still $o_p(\frac{1}{\sqrt{n}})$ after summing over i.

To accomplish the second part of proof, we will first discuss the residual term of the order $o_p(\frac{1}{\sqrt{n}})$ and then the influence term.

Step 2 (residual term): Here we claim that

$$\frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \left[r(x_i) R_{nij} - r(x_i) \mathrm{E}(R_{nij}|x_i)\right] = o_p(\frac{1}{\sqrt{n}}). \tag{23}$$

Let $R_{nij} = R_{1nij} + R_{2nij}$, where

$$R_{1nij} = \frac{s_j}{h^k} K(\frac{x_j - x_i}{h}) \int_{\Omega_{u_l}} K(u_l)\left[f_w(w_j + hu_l) - f_w(w_j)\right] \mathrm{d}u_l,$$

$$R_{2nij} = \frac{1}{h^k}\int_{\Omega_{u_l}} \left[\mathrm{E}\left[s_j|w_j + hu_l\right] K(\frac{x_j + hu_l - x_i}{h}) f_w(w_j + hu_l) - \mathrm{E}\left[s_j|w_j\right] K(\frac{x_j - x_i}{h}) f_w(w_j)\right] K(u_l)\, \mathrm{d}u_l.$$

So the object of this small session is

$$\frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \left[r(x_i) R_{1nij} - r(x_i) \mathrm{E}(R_{1nij}|x_i) + r(x_i) R_{2nij} - r(x_i) \mathrm{E}(R_{2nij}|x_i)\right].$$

For $\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} r\left(x_i\right) R_{1nij}$, it is a standard U-statistics after being written as

$$\begin{aligned}
& \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} \frac{r\left(x_i\right) s_j}{h^k} K\left(\frac{x_j - x_i}{h}\right) \int_{\Omega_{u_l}} K\left(u_l\right)\left[f_w(w_j + hu_l) - f_w(w_j)\right] \mathrm{d}u_l \qquad (24) \\
= & \frac{1}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left[P_{21}(z_i, z_j) + P_{22}(z_i, z_j)\right] = \frac{2}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{n} P_2(z_i, z_j),
\end{aligned}$$

where

$$P_{21}(z_i, z_j) = r\left(x_j\right) s_i \int_{\Omega_{u_l}} K\left(u_l\right)\left[f_w(w_i + hu_l) - f_w(w_i)\right] \mathrm{d}u_l \frac{1}{h^k} K\left(\frac{x_j - x_i}{h}\right),$$

$$P_{22}(z_i, z_j) = r\left(x_i\right) s_j \int_{\Omega_{u_l}} K\left(u_l\right)\left[f_w(w_j + hu_l) - f_w(w_j)\right] \mathrm{d}u_l \frac{1}{h^k} K\left(\frac{x_j - x_i}{h}\right),$$

$$P_2(z_i, z_j) = \frac{1}{2}\left[P_{21}(z_i, z_j) + P_{22}(z_i, z_j)\right].$$

For the similar reason as in Step 1, easy to show that $\mathrm{E}\left[P_2(z_i, z_j)^2\right] = o_p(n)$. Then we could apply Lemma 3.2 in Powell et al. (1989) again to get that

$$\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} r\left(x_i\right) R_{1nij} = \mathrm{E}\left[P_2(z_i, z_j)\right] + \frac{2}{n}\sum_{i=1}^{n}\left[\mathrm{E}\left[P_2(z_i, z_j)|z_i\right] - \mathrm{E}\left[P_2(z_i, z_j)\right]\right] + o_p\left(\frac{1}{\sqrt{n}}\right). \qquad (25)$$

So

$$\begin{aligned}
& \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\left[r\left(x_i\right) R_{1nij} - r\left(x_i\right)\mathrm{E}\left(R_{1nij}|x_i\right)\right] \qquad (26) \\
= & \frac{1}{n}\sum_{i=1}^{n}\left[2\mathrm{E}\left[P_2(z_i, z_j)|z_i\right] - r\left(x_i\right)\mathrm{E}\left(R_{1nij}|x_i\right) - \mathrm{E}\left[P_2(z_i, z_j)\right]\right] + o_p\left(\frac{1}{\sqrt{n}}\right), \\
= & \frac{1}{n}\sum_{i=1}^{n}\left[\mathrm{E}\left[P_{21}(z_i, z_j)|z_i\right] + \mathrm{E}\left[P_{22}(z_i, z_j)|z_i\right] - r\left(x_i\right)\mathrm{E}\left(R_{1nij}|x_i\right) - \mathrm{E}\left[P_2(z_i, z_j)\right]\right] + o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Since z_i and z_j are identically distributed,

$$E\left[P_{21}(z_i, z_j)\right] = E\left[P_{22}(z_i, z_j)\right] = \mathrm{E}\left[P_2(z_i, z_j)\right]. \qquad (27)$$

By changing variables $u_j = \frac{x_j - x_i}{h}$ in the integral,

$$
\begin{aligned}
\mathrm{E}\left[P_{21}(z_i, z_j)|z_i\right] &= s_i \int_{\Omega_{u_l}} K\left(u_l\right)\left[f_w(w_i + hu_l) - f_w(w_i)\right] \mathrm{d}u_l \int_{\Omega_{x_j}} \frac{r(x_j)}{h^k} K\left(\frac{x_j - x_i}{h}\right) f_x(x_j) \mathrm{d}x_j \\
&= s_i \int_{\Omega_{u_l}} K\left(u_l\right)\left[f_w(w_i + hu_l) - f_w(w_i)\right] \mathrm{d}u_l \int_{\Omega_{u_j}} r\left(x_i + hu_j\right) K\left(u_j\right) f_x(x_i + hu_j) \mathrm{d}u_j
\end{aligned}
$$

$\frac{1}{h^k}$ has been cancelled out above, then by the assumptions that the first derivative of f_w and each term above are bounded, we know that

$$
\mathrm{E}\left[\mathrm{E}\left[P_{21}(z_i, z_j)|z_i\right]^2\right] = O(h^2). \tag{28}
$$

Note that

$$
r\left(x_i\right) \mathrm{E}\left(R_{1nij}|x_i\right) = r\left(x_i\right) \mathrm{E}\left[\frac{s_j}{h^k} K(\frac{x_j - x_i}{h}) \int_{\Omega_{u_l}} K\left(u_l\right)\left[f_w(w_j + hu_l) - f_w(w_j)\right] \mathrm{d}u_l \middle| x_i\right]
$$

which is exactly $\mathrm{E}[P_{22}(z_i, z_j)|z_i]$. Therefore, Equation (26) is equal to the following

$$
\begin{aligned}
& \frac{1}{n} \sum_{i=1}^{n} \left[\mathrm{E}\left[P_{21}(z_i, z_j)|z_i\right] - \mathrm{E}\left[P_2(z_i, z_j)\right]\right] + o_p\left(\frac{1}{\sqrt{n}}\right), \\
= \; & \frac{1}{n} \sum_{i=1}^{n} \left[\mathrm{E}\left[P_{21}(z_i, z_j)|z_i\right] - \mathrm{E}\left[P_{21}(z_i, z_j)\right]\right] + o_p\left(\frac{1}{\sqrt{n}}\right),
\end{aligned} \tag{29}
$$

where the last equality follows by Equation (27). By Equation (28) and Lindeberg–Levy central limit theorem,

$$
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\mathrm{E}\left[P_{21}(z_i, z_j)|z_i\right] - \mathrm{E}\left[P_{21}(z_i, z_j)\right]\right] \xrightarrow{p} 0,
$$

which implies that Equation (26) is $o_p(\frac{1}{\sqrt{n}})$.

Similarly

$$
\frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \left[r\left(x_i\right) R_{2nij} - r\left(x_i\right) \mathrm{E}\left(R_{2nij}|x_i\right)\right] = o_p(\frac{1}{\sqrt{n}}).
$$

Then Equation (23) is proved.

Step 3 (influence term): Taking results from Step 1 and 2, we know that the influence term for Equation (13) is

$$\frac{1}{n}\sum_{i=1}^{n} r(x_i)\mathrm{E}\left[\frac{s_j}{h^{2k+1}}K\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h})\bigg| x_i\right]$$
$$+\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} r(x_i)\left\{\frac{1}{h^k}s_jK\left(\frac{x_j-x_i}{h}\right)f_w(w_j)+\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w(w_j)\right.$$
$$\left.-\mathrm{E}\left[\frac{1}{h^k}s_jK\left(\frac{x_j-x_i}{h}\right)f_w(w_j)\bigg| x_i\right]-\mathrm{E}\left[\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w(w_j)\bigg| x_i\right]\right\}.$$

Let

$$\Delta_1=\frac{1}{n}\sum_{i=1}^{n} r(x_i)\mathrm{E}\left[\frac{s_j}{h^{2k+1}}K\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h})\bigg| x_i\right],$$

$$\Delta_2=\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} r(x_i)\left\{\frac{1}{h^k}s_jK\left(\frac{x_j-x_i}{h}\right)f_w(w_j)-\mathrm{E}\left[\frac{1}{h^k}s_jK\left(\frac{x_j-x_i}{h}\right)f_w(w_j)\bigg| x_i\right]\right\},$$

$$\begin{aligned}\Delta_3 &= \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} r(x_i)\left\{\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w(w_j)\right.\\ &\quad \left.-\mathrm{E}\left[\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w(w_j)\bigg| x_i\right]\right\}.\end{aligned}$$

Easy the see that the influence term is $\Delta_1+\Delta_2+\Delta_3$, we will discuss the property of them one by one.

For Δ_1, the proof is very similar to the proof of the second conclusion in Lemma 4.1. Note that

$$\begin{aligned}\Delta_1-\mathrm{E}(\Delta_1) &= \Delta_1-\mathrm{E}\left[\frac{r(x_i)s_j}{h^{2k+1}}K\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h})\right] \\ &= \frac{1}{n}\sum_{i=1}^{n}\left\{r(x_i)f_x(x_i)\mathrm{E}\left[s_if_w(w_i)|\, x_i\right]-\mathrm{E}\left[r(x_i)s_if_x(x_i)f_w(w_i)\right]+R_{ni}-\mathrm{E}(R_{ni})\right\},\end{aligned} \tag{30}$$

where

$$R_{ni}=\mathrm{E}\left[\frac{r(x_i)s_j}{h^{2k+1}}K\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h})\bigg| x_i\right]-\mathrm{E}\left[r(x_i)s_if_x(x_i)f_w(w_i)|\, x_i\right].$$

Easy to see that $\mathrm{E}\big(|R_{ni}|^2\big) = O(h^2)$, since x_i is i.i.d. across i,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[R_{ni} - \mathrm{E}(R_{ni})\right] \xrightarrow{p} 0.$$

So the influence term for Δ_1 is

$$\mathrm{E}\left[\frac{r(x_i)s_j}{h^{2k+1}} K\left(\frac{x_j - x_i}{h}\right) K(\frac{w_l - w_j}{h})\right] + \frac{1}{n}\sum_{i=1}^{n} \left\{r(x_i)f_x(x_i)\mathrm{E}\left[s_i f_w(w_i)|\, x_i\right] - \mathrm{E}\left[r(x_i)s_i f_x(x_i) f_w(w_i)\right]\right\}. \tag{31}$$

Since the kernel function used here is of order p, and by assumption that f_w,$\mathrm{E}(s_i|x_i)$ are p-th order differentiable and p-th derivatives are bounded, we know that

$$\mathrm{E}\left[\frac{r(x_i)s_j}{h^{2k+1}} K\left(\frac{x_j - x_i}{h}\right) K(\frac{w_l - w_j}{h})\right] = \mathrm{E}\left[r(x_i)s_i f_x(x_i) f_w(w_i)\right] + O(h^p).$$

Therefore, by assumption $n^{\frac{1}{2}} h^p \to 0$, Equation (31) is equal to

$$\begin{aligned}
& \mathrm{E}\left[r(x_i)s_i f_x(x_i) f_w(w_i)\right] + \frac{1}{n}\sum_{i=1}^{n} \left\{r(x_i)\mathrm{E}\left[f_x(x_i)s_i f_w(w_i)|\, x_i\right] - \mathrm{E}\left[r(x_i)s_i f_x(x_i) f_w(w_i)\right]\right\} + o_p\left(\frac{1}{\sqrt{n}}\right) \\
= & \frac{1}{n}\sum_{i=1}^{n} r(x_i)\mathrm{E}\left[s_i f_x(x_i) f_w(w_i)|\, x_i\right] + o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Sum up the results for Δ_1 so far, Δ_1 is equal to

$$\frac{1}{n}\sum_{i=1}^{n} r(x_i)\mathrm{E}\left[s_i f_x(x_i) f_w(w_i)|\, x_i\right] + o_p\left(\frac{1}{\sqrt{n}}\right). \tag{32}$$

For Δ_2,

$$\begin{aligned}
\Delta_2 &= \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} r(x_i)\left\{\frac{1}{h^k} s_j K\left(\frac{x_j - x_i}{h}\right) f_w\left(w_j\right) - \mathrm{E}\left[\frac{1}{h^k} s_j K\left(\frac{x_j - x_i}{h}\right) f_w\left(w_j\right)\middle|\, x_i\right]\right\} \\
&= \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P_3(z_i, z_j) - \frac{1}{n}\sum_{i=1}^{n} r(x_i)\mathrm{E}\left[\frac{1}{h^k} s_j K\left(\frac{x_j - x_i}{h}\right) f_w\left(w_j\right)\middle|\, x_i\right], \qquad (33)
\end{aligned}$$

where

$$P_3(z_i, z_j) = \frac{1}{2h^k}\left[r(x_i)s_j f_w(w_j) + r(x_j)s_i f_w(w_i)\right] K\left(\frac{x_j - x_i}{h}\right).$$

Let

$$P_{31}(z_i, z_j) = \frac{1}{h^k} r(x_j) s_i f_w(w_i) K\left(\frac{x_j - x_i}{h}\right),$$

$$P_{32}(z_i, z_j) = \frac{1}{h^k} r(x_i) s_j f_w(w_j) K\left(\frac{x_i - x_j}{h}\right),$$

then

$$P_3(z_i, z_j) = \frac{1}{2}\left[P_{31}(z_i, z_j) + P_{32}(z_i, z_j)\right].$$

Notice that the structure of Δ_2 is the same as Equation (26) and P_3, P_{31}, P_{32} here are corresponding to P_2, P_{21}, P_{22} respectively. By the results in Step 2 that Equation (26) is equal to Equation (29), similarly, Δ_2 is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\left\{\mathrm{E}\left[P_{31}(z_i, z_j)\right| z_i\right] - \mathrm{E}\left[P_{31}(z_i, z_j)\right]\} + o_p\left(\frac{1}{\sqrt{n}}\right).$$

The above expression could be rewritten as

$$\begin{aligned} & \frac{1}{n}\sum_{i=1}^{n}\left\{s_i f_w(w_i)\mathrm{E}\left[\frac{r(x_j)}{h^k}K\left(\frac{x_j - x_i}{h}\right)\middle| z_i\right] - \mathrm{E}\left[P_{31}(z_i, z_j)\right]\right\}, \\ = \; & \frac{1}{n}\sum_{i=1}^{n}\left\{r(x_i)s_i f_w(w_i) f_x(x_i) - \mathrm{E}\left[r(x_i)s_i f_w(w_i) f_x(x_i)\right] + R_{ni} - \mathrm{E}\left(R_{ni}\right)\right\}, \end{aligned}$$

where

$$R_{ni} = s_i f_w(w_i)\mathrm{E}\left[\frac{r(x_j)}{h^k}K\left(\frac{x_j - x_i}{h}\right)\middle| z_i\right] - r(x_i)s_i f_w(w_i) f_x(x_i).$$

Easy to see that $\mathrm{E}(R_{ni}^2) = O(h^2)$, by i.i.d. assumption on z_i and Lindeberg Levy CLT,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[R_{ni} - \mathrm{E}\left(R_{ni}\right)\right] \xrightarrow{p} 0.$$

So Δ_2 is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\left\{r(x_i)s_i f_w(w_i) f_x(x_i) - \mathrm{E}\left[r(x_i)s_i f_w(w_i) f_x(x_i)\right]\right\} + o_p\left(\frac{1}{\sqrt{n}}\right) \qquad (34)$$

For Δ_3, it has the same structure as Δ_2. Similarly, one could get that Δ_3 is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\left\{r(x_i)\mathrm{E}\left(s_i|w_i\right)f_w(w_i)f_x(x_i)-\mathrm{E}\left[r(x_i)s_if_w(w_i)f_x(x_i)\right]\right\}+o_p\left(\frac{1}{\sqrt{n}}\right). \tag{35}$$

At this stage, from Equation (32), (34), and (35), we could get that $\Delta_1+\Delta_2+\Delta_3$ is equal to

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}\{r(x_i)s_if_w(w_i)f_x(x_i)+r(x_i)f_x(x_i)\mathrm{E}\left[s_if_w(w_i)|\,x_i\right] \\ &+r(x_i)\mathrm{E}\left(s_i|w_i\right)f_w(w_i)f_x(x_i)-2\mathrm{E}\left[r(x_i)s_if_w(w_i)f_x(x_i)\right]\}+o_p\left(\frac{1}{\sqrt{n}}\right).\end{aligned} \tag{36}$$

Sum up the results in Step 1, 2, and 3, we know that the Equation (13) is equal to Equation (36), which is the conclusion of this lemma. ■

Lemma 4.3 *Adopt the same notation and assumptions as in Lemma 4.2. Then*

$$\frac{1}{(n-1)}\sum_{j=1,j\neq i}^{n}s_j\widehat{f}_w(w_j)\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)-E\left[s_if_w(w_i)|x_i\right]f_x(x_i)=O_p(h^p)+O_p\left[\left(\frac{1}{n^{1-\varepsilon}h^k}\right)^{\frac{1}{2}}\right]. \tag{37}$$

This Lemma is a complement to Lemma 4.2. The result will be used to determine the rate of convergence of the residual terms after decomposing our estimator.

Proof of Lemma 4.3. The proof of this lemma use some results in lemma 4.2.

First note that

$$\begin{aligned}&\frac{1}{(n-1)}\sum_{j=1,j\neq i}^{n}s_j\widehat{f}_w(w_j)\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)-\mathrm{E}\left[s_if_w(w_i)|x_i\right]f_x(x_i)\\ =&\frac{1}{(n-1)}\sum_{j=1,j\neq i}^{n}s_j\widehat{f}_w(w_j)\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)-\widehat{\mathrm{E}}\left[s_if_w(w_i)|x_i\right]\widehat{f}_x(x_i)\\ &+\widehat{\mathrm{E}}\left[s_if_w(w_i)|x_i\right]\left(\widehat{f}_x(x_i)-f_x(x_i)\right)+\left(\widehat{\mathrm{E}}\left[s_if_w(w_i)|x_i\right]-\mathrm{E}\left[s_if_w(w_i)|x_i\right]\right)f_x(x_i).\end{aligned}$$

Under Assumption $n^{1-\varepsilon}h^{2k+2}\to\infty$, from Li and Racine (2007) Chapter 2.3,

$$\sup\left|\widehat{\mathrm{E}}\left[s_if_w(w_i)|x_i\right]-E\left[s_if_w(w_i)|x_i\right]\right|=O_p\left[\left(\frac{1}{n^{1-\varepsilon}h^k}\right)^{\frac{1}{2}}\right].$$

From Silverman (1978), $\sup\left|\widehat{f}_x(x_i) - f_x(x_i)\right|$ is also $O_p\left[\left(\frac{1}{n^{1-\varepsilon}h^k}\right)^{\frac{1}{2}}\right]$. So we only need to focus on

$$\frac{1}{(n-1)}\sum_{j=1,j\neq i}^{n} s_j\widehat{f}_w(w_j)\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right) - \widehat{\mathrm{E}}\left[s_i f_w(w_i)|x_i\right]\widehat{f}_x(x_i) \tag{38}$$

which is

$$\frac{1}{(n-1)^2h^{2k+1}}\sum_{j=1,j\neq i}^{n}\sum_{l=1,l\neq j}^{n} s_jK\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h}) - \frac{1}{(n-1)h^k}\sum_{j=1,j\neq i}^{n} s_jf_w(w_j)K\left(\frac{x_j-x_i}{h}\right).$$

Drop $l = i$ in the first part of the above expression to make it look more like a U-statistics. This term is of the order $O_p\left(\frac{1}{n^2h^{2k+1}}\right)$ which is $o_p\left[\left(\frac{1}{n^{1-\varepsilon}h^k}\right)^{\frac{1}{2}}\right]$ by some simple calculation, so this term could be ignored. After dropping one term, the above term becomes

$$\begin{aligned} &\frac{1}{(n-1)(n-2)h^{2k+1}}\sum_{j=1,j\neq i}^{n}\sum_{l=1,l\neq i,j}^{n} s_jK\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h}) \\ &-\frac{1}{(n-1)h^k}\sum_{j=1,j\neq i}^{n} s_jf_w(w_j)K\left(\frac{x_j-x_i}{h}\right). \end{aligned} \tag{39}$$

As already discussed in the Step 1 of Lemma 4.2,

$$\frac{1}{(n-1)(n-2)h^{2k+1}}\sum_{j=1,j\neq i}^{n}\sum_{l=1,l\neq i,j}^{n} s_jK\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h})$$

is $\sqrt{n}$-equivalent to Equation (16). For the convenience of reading, Equation (16) is rewritten as follows:

$$\begin{aligned} &\mathrm{E}\left[\frac{s_j}{h^{2k+1}}K\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h})\middle| x_i\right] + \frac{1}{n-1}\sum_{j=1,j\neq i}^{n}\left\{\frac{1}{h^k}s_jK\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\right. \\ &+\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right) - \mathrm{E}\left[\frac{1}{h^k}s_jK\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\middle| x_i\right] \\ &\left.-\mathrm{E}\left[\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\middle| x_i\right] + R_{nij} - \mathrm{E}(R_{nij}|x_i) + o_p\left(\frac{1}{\sqrt{n}}\right)\right\}, \end{aligned} \tag{40}$$

Substitute Equation (40) into Equation (39), then Equation (39) becomes

$$\begin{aligned}
&\mathrm{E}\left[\frac{s_j}{h^{2k+1}}K\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h})\bigg|\, x_i\right]-\mathrm{E}\left[\frac{1}{h^k}s_jK\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\bigg|\, x_i\right] \qquad (41)\\
&+\frac{1}{n-1}\sum_{j=1,j\neq i}^{n}\left\{\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)-\mathrm{E}\left[\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\bigg|\, x_i\right]\right.\\
&\left.+R_{nij}-\mathrm{E}(R_{nij}|x_i)+o_p\left(\frac{1}{\sqrt{n}}\right)\right\}.
\end{aligned}$$

We will discuss each term in Equation (41).

By the assumption that f_w,$\mathrm{E}(s_i|x_i)$, f_x are p-th order differentiable and p-th derivatives are bounded, and the kernel function we use is of p-th order, we know that

$$\mathrm{E}\left[\frac{s_j}{h^{2k+1}}K\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h})\bigg|\, x_i\right]=\mathrm{E}\left[s_if_x(x_i)f_w(w_i)|\,x_i\right]+O_p(h^p),$$

$$\mathrm{E}\left[\frac{s_j}{h^k}K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\bigg|\, x_i\right]=\mathrm{E}\left[s_if_x(x_i)f_w(w_i)|\,x_i\right]+O_p(h^p),$$

so

$$\mathrm{E}\left[\frac{s_j}{h^{2k+1}}K\left(\frac{x_j-x_i}{h}\right)K(\frac{w_l-w_j}{h})\bigg|\, x_i\right]-\mathrm{E}\left[\frac{s_j}{h^k}K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\bigg|\, x_i\right]=O_p(h^p). \qquad (42)$$

Let

$$\Delta=\frac{1}{n-1}\sum_{j=1,j\neq i}^{n}\left\{\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)-\mathrm{E}\left[\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\bigg|\, x_i\right]\right\}.$$

By i.i.d. assumption, j-th observation is independent of l-th observation when $j\neq l$. So

$$\mathrm{E}(\Delta^2|x_i)=\frac{1}{n-1}\mathrm{var}\left[\left(\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\right)^2\bigg|\, x_i\right].$$

By

$$
\begin{aligned}
& \mathrm{E}\left[\left(\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\right)^2\middle| x_i\right] \\
= & \int_{\Omega_{w_j}}\left(\frac{1}{h^k}\mathrm{E}(s_j|w_j)K\left(\frac{x_j-x_i}{h}\right)f_w\left(w_j\right)\right)^2 f_w(w_j)\mathrm{d}w_j \\
= & \frac{1}{h^k}\int_{\Omega_{u_j,v_j}}\left[\mathrm{E}(s_j|x_i+hu_j,v_j)K\left(u_j\right)f_w\left(x_i+hu_j,v_j\right)\right]^2 f_w(x_i+hu_j,v_j)\mathrm{d}u_j\mathrm{d}v_j = O_p(\frac{1}{h^k}),
\end{aligned}
$$

we know

$$
\mathrm{E}(\Delta^2|x_i) = O_p\left[\left(nh^k\right)^{-1}\right].
$$

According to Markov inequality,

$$
\Delta = O_p\left[\frac{1}{(nh^k)^{\frac{1}{2}}}\right]. \tag{43}
$$

Similarly, let

$$
R = \frac{1}{n}\sum_{j=1,j\neq i}^{n}\left[R_{nij}-\mathrm{E}(R_{nij}|x_i)\right],
$$

and we can get that

$$
R = o_p\left[\frac{1}{(nh^k)^{\frac{1}{2}}}\right]. \tag{44}
$$

Equation (38) is equal to Equation (41). The order of Equation (41) could be obtained from Equation (42), (43) and (44), which is $O_p(h^p)+O_p\left[\frac{1}{(nh^k)^{\frac{1}{2}}}\right]$. ■

Corollary 4.4 *Adopt the same notation and assumptions as in Lemma 4.2. Then*

$$
\widehat{E}\left[\frac{s_i}{\widehat{f}(v_i|x_i)}|x_i\right]\widehat{f}_x(x_i) - E\left[\frac{s_i}{f(v_i|x_i)}|x_i\right]f_x(x_i) = O_p(h^p)+O_p\left[\frac{1}{(nh^k)^{\frac{1}{2}}}\right]+O_p(\frac{1}{n^{1-\varepsilon}h^{2(k+1)}}). \tag{45}
$$

This Corollary is a direct result of Lemma 4.3. It is useful, since our estimator includes a nonparametric estimate of conditional density in the denominator.

Proof of Corollary 4.4.

$$
\begin{aligned}
& \widehat{\mathrm{E}}\left[\frac{s_i}{\widehat{f}(v_i|x_i)}|x_i\right]\widehat{f}_x(x_i) - \mathrm{E}\left[\frac{s_i}{f(v_i|x_i)}|x_i\right] f_x(x_i) \\
= & \widehat{\mathrm{E}}\left[\frac{s_i}{\widehat{f}(v_i|x_i)}|x_i\right]\widehat{f}_x(x_i) - \widehat{\mathrm{E}}\left[\frac{s_i}{f(v_i|x_i)}|x_i\right]\widehat{f}_x(x_i) + \widehat{\mathrm{E}}\left[\frac{s_i}{f(v_i|x_i)}|x_i\right]\left[\widehat{f}_x(x_i) - f_x(x_i)\right] \\
& + \left\{\widehat{\mathrm{E}}\left[\frac{s_i}{f(v_i|x_i)}|x_i\right] - \mathrm{E}\left[\frac{s_i}{f(v_i|x_i)}|x_i\right]\right\} f_x(x_i).
\end{aligned}
$$

All terms except the first term are easily seen to be $O_p\left[(n^{1-\varepsilon}h^k)^{-\frac{1}{2}}\right]$. For the first term

$$
\begin{aligned}
& \widehat{\mathrm{E}}\left[\frac{s_i}{\widehat{f}(v_i|x_i)}|x_i\right]\widehat{f}_x(x_i) - \widehat{\mathrm{E}}\left[\frac{s_i}{f(v_i|x_i)}|x_i\right]\widehat{f}_x(x_i) \\
= & \frac{1}{n-1}\sum_{j=1,j\neq i}^{n} \frac{s_j \widehat{f}_x(x_j)}{\widehat{f}_w(w_j)}\frac{1}{h^k}K\left(\frac{x_j - x_i}{h}\right) - \frac{1}{n-1}\sum_{j=1,j\neq i}^{n} \frac{s_j f_x(x_j)}{f_w(w_j)}\frac{1}{h^k}K\left(\frac{x_j - x_i}{h}\right) \\
= & \frac{1}{n-1}\sum_{j=1,j\neq i}^{n} \frac{s_j\left[\widehat{f}_x(x_j) - f_x(x_j)\right]}{f_w(w_j)}\frac{1}{h^k}K\left(\frac{x_j - x_i}{h}\right) \qquad (46) \\
& + \frac{1}{n-1}\sum_{j=1,j\neq i}^{n} \frac{s_j f_x(x_j)\left[\widehat{f}_w(w_j) - f_w(w_j)\right]}{f_w^2(w_j)}\frac{1}{h^k}K\left(\frac{x_j - x_i}{h}\right) \qquad (47) \\
& + \frac{1}{n-1}\sum_{j=1,j\neq i}^{n} \frac{s_j\left[\widehat{f}_x(x_j) - f_x(x_j)\right]\left[\widehat{f}_w(w_j) - f_w(w_j)\right]}{f_w^2(w_j)}\frac{1}{h^k}K\left(\frac{x_j - x_i}{h}\right) \qquad (48) \\
& + \frac{1}{n-1}\sum_{j=1,j\neq i}^{n} \frac{s_j \widehat{f}_x(x_j)\left[\widehat{f}_w(w_j) - f_w(w_j)\right]^2}{f_w^2(w_j)\widehat{f}_w(w_j)}\frac{1}{h^k}K\left(\frac{x_j - x_i}{h}\right). \qquad (49)
\end{aligned}
$$

According the results in Lemma 4.3, and by seeing that Equation (46) is exactly Equation (38), Equation (46 is of order $O_p(h^p) + O_p\left[(nh^k)^{-\frac{1}{2}}\right]$. For the same reason, Equation (47) is also of order $O_p(h^p) + O_p\left[(nh^k)^{-\frac{1}{2}}\right]$. From Silverman (1978),

$$
\sup\left|\widehat{f}_x(x_j) - f_x(x_j)\right| = O\left[(n^{1-\varepsilon}h^k)^{-\frac{1}{2}}\right],
$$

$$
\sup\left|\widehat{f}_w(w_j) - f_x(w_j)\right| = O\left[(n^{1-\varepsilon}h^{k+1})^{-\frac{1}{2}}\right],
$$

for any $\varepsilon > 0$, so together with the bounds condition on each term in Equation (48) and (49), we know that Equation (48) and (49) are of the order $O_p(\frac{1}{n^{1-\varepsilon}h^{2k}})$ and $O_p(\frac{1}{n^{1-\varepsilon}h^{2(k+1)}})$ respectively.

Combine results so far, the conclusion is proved. ■

Proof of Theorem 2.6. We will first derive the property of $\frac{1}{n}\sum_{i=1}^{n}\widehat{\psi}_1(x_i)$. It could be divided into several components as follows

$$\begin{aligned}\widehat{\psi}_1(x_i) &= \frac{\widehat{\mathrm{E}}\left(\widehat{h}_{1i}\middle|\, x_i\right)}{\widehat{\mathrm{E}}\left(\widehat{g}_{1i}\middle|\, x_i\right)} = \frac{\widehat{\mathrm{E}}\left(\widehat{\widetilde{h}}_{1i}\middle|\, x_i\right)}{\widehat{\mathrm{E}}\left(\widehat{\widetilde{g}}_{1i}\middle|\, x_i\right)} \\ &= \frac{\widehat{\mathrm{E}}\left(\widehat{\widetilde{h}}_{1i}\middle|\, x_i\right)}{\mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)} - \frac{\mathrm{E}\left(\widetilde{h}_{1i}\middle|\, x_i\right)\left[\widehat{\mathrm{E}}\left(\widehat{\widetilde{g}}_{1i}\middle|\, x_i\right) - \mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)\right]}{\left[\mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)\right]^2} + R_i, \end{aligned} \tag{50}$$

where

$$R_i = \frac{\left[\widehat{\mathrm{E}}\left(\widehat{\widetilde{h}}_{1i}\middle|\, x_i\right) - \mathrm{E}\left(\widetilde{h}_{1i}\middle|\, x_i\right)\right]\left[\widehat{\mathrm{E}}\left(\widehat{\widetilde{g}}_{1i}\middle|\, x_i\right) - \mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)\right]}{\left[\mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)\right]^2} + \frac{\widehat{\mathrm{E}}\left(\widehat{\widetilde{h}}_{1i}\middle|\, x_i\right)\left[\widehat{\mathrm{E}}\left(\widehat{\widetilde{g}}_{1i}\middle|\, x_i\right) - \mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)\right]^2}{\left[\mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)\right]^2}.$$

According to Corollary 4.4, and assumption that $\frac{1}{E(\widetilde{g}_{1i}|x_i)}$ is bounded, R_i is of order $O_p(h^{2p}) + O_p\left[\frac{1}{nh^k}\right] + O_p(\frac{1}{n^{2-\varepsilon}h^{4(k+1)}})$, which $o_p(\frac{1}{\sqrt{n}})$, by Assumption that $nh^{2p} \to 0$, and $n^{1-\varepsilon}h^{4k+4} \to \infty$. So $\frac{1}{n}\sum_{i=1}^{n}\widehat{\psi}_1(x_i)$ is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{\widehat{\mathrm{E}}\left(\widehat{\widetilde{h}}_{1i}\middle|\, x_i\right)}{\mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)} - \frac{\mathrm{E}\left(\widetilde{h}_{1i}\middle|\, x_i\right)\left[\widehat{\mathrm{E}}\left(\widehat{\widetilde{g}}_{1i}\middle|\, x_i\right) - \mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)\right]}{\left[\mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)\right]^2}\right\} + o_p\left(\frac{1}{\sqrt{n}}\right). \tag{51}$$

Notice that

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{\mathrm{E}}\left(\widehat{\widetilde{h}}_{1i}\middle|\, x_i\right)}{\mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)} \\ = &\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\frac{1}{\mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)}\frac{D_jY_j}{\widehat{f}(v_j|x_j)}\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right) \\ = &\frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\left\{\frac{1}{\mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)}\frac{D_jY_j\widehat{f}_x(x_j)}{f_{xv}(x_j,v_j)}\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)\right. \\ &\left. -\frac{1}{\mathrm{E}\left(\widetilde{g}_{1i}\middle|\, x_i\right)}\frac{D_jY_jf_x(x_j)\left[\widehat{f}_{xv}(x_j,v_j) - f_{xv}(x_j,v_j)\right]}{f_{xv}^2(x_j,v_j)}\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right) + R_{ij}\right\}, \end{aligned} \tag{52}$$

where

$$
\begin{aligned}
R_{ij} &= \frac{D_j Y_j \widehat{f}_x(x_j)\left[\widehat{f}_{xv}(x_j,v_j)-f_{xv}(x_j,v_j)\right]^2}{\mathrm{E}\left(\widetilde{g}_{1i}|\,x_i\right) f_{xv}^2(x_j,v_j)\widehat{f}_{xv}(x_j,v_j)}\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right) \\
&\quad -\frac{D_j Y_j\left[\widehat{f}_x(x_j)-f_x(x_j)\right]\left[\widehat{f}_{xv}(x_j,v_j)-f_{xv}(x_j,v_j)\right]}{\mathrm{E}\left(\widetilde{g}_{1i}|\,x_i\right) f_{xv}^2(x_j,v_j)}\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right).
\end{aligned}
$$

Consider the residual term first, again from Silverman (1978) and those bound conditions, we know that

$$
\sup|R_{ij}| = O_p\left(\frac{1}{n^{1-\varepsilon}h^{2k}}\right) = o_p\left(\frac{1}{\sqrt{n}}\right), \tag{53}
$$

the second equality follows by Assumption that $n^{1-\varepsilon}h^{4k+4}\to\infty$. Apply Lemma 4.2 (by letting $s_j = s_{5j}, r(x_i) = r_5(x_i)$) on the first term in Equation (52)[4],

$$
\frac{1}{n\,(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\frac{1}{\mathrm{E}\left(\widetilde{g}_{1i}|\,x_i\right)}\frac{D_j Y_j \widehat{f}_x(x_j)}{f_{xv}(x_j,v_j)}\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right),
$$

we know that is equal to

$$
\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{2\mathrm{E}(h_{1i}|x_i)}{\mathrm{E}(g_{1i}|x_i)}+\frac{h_{1i}}{\mathrm{E}(g_{1i}|x_i)}-2\mathrm{E}\left[\frac{\mathrm{E}(h_{1i}|x_i)}{\mathrm{E}(g_{1i}|x_i)}\right]\right\}+o_p\left(\frac{1}{\sqrt{n}}\right),
$$

which by further simplification is:

$$
\frac{1}{n}\sum_{i=1}^{n}\left[2\mathrm{E}\left(Y_1|x_i\right)+\frac{h_{1i}}{\mathrm{E}(g_{1i}|x_i)}-2\mathrm{E}\left(Y_1\right)\right]. \tag{54}
$$

The second component in Equation (52) can be rewritten as

$$
\begin{aligned}
&\frac{1}{n\,(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\left[\frac{D_j Y_j f_x(x_j)\widehat{f}_{xv}(x_j,v_j)}{\mathrm{E}\left(\widetilde{g}_{1i}|\,x_i\right) f_{xv}^2(x_j,v_j)}\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)\right. \\
&\left.-\frac{D_j Y_j f_x(x_j)}{\mathrm{E}\left(\widetilde{g}_{1i}|\,x_i\right) f_{xv}(x_j,v_j)}\frac{1}{h^k}K\left(\frac{x_j-x_i}{h}\right)\right].
\end{aligned} \tag{55}
$$

Apply Lemma 4.2 (by letting $s_j = s_{6j}, r(x_i) = r_6(x_i)$) on the first term in Equation (55), we could

[4] By assuming v in Lemma 4.2 is empty.

get that it is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\left[\mathrm{E}\left(Y_1|x_i\right)+\frac{h_{1i}}{\mathrm{E}(g_{1i}|x_i)}+\frac{\mathrm{E}(h_{1i}|x_i,v_i)}{\mathrm{E}(g_{1i}|x_i)}-2\mathrm{E}\left(Y_1\right)\right]+o_p\left(\frac{1}{\sqrt{n}}\right). \tag{56}$$

Apply Lemma 4.1 on the second term in Equation (55), we could get that it is equal to

$$\mathrm{E}\left(Y_1\right)+\frac{1}{n}\sum_{i=1}^{n}\left[\mathrm{E}\left(Y_1|x_i\right)+\frac{h_{1i}}{\mathrm{E}(g_{1i}|x_i)}-2\mathrm{E}\left(Y_1\right)\right]+o_p\left(\frac{1}{\sqrt{n}}\right). \tag{57}$$

From Equation (56), (57), we know that Equation (55) is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\mathrm{E}(h_{1i}|x_i,v_i)}{\mathrm{E}(g_{1i}|x_i)}-\mathrm{E}\left(Y_1\right)\right]+o_p\left(\frac{1}{\sqrt{n}}\right). \tag{58}$$

Combine Equation (53), (54), (58), and (52), we can get that $\frac{1}{n}\sum_{i=1}^{n}\frac{\widehat{\mathrm{E}}\left(\widehat{\widetilde{h}}_{1i}\middle|x_i\right)}{\mathrm{E}(\widetilde{g}_{1i}|x_i)}$ is equal to

$$\mathrm{E}(Y_1)+\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{h_{1i}}{\mathrm{E}\left(g_{1i}|x_i\right)}-\frac{\mathrm{E}\left(h_{1i}|x_i,v_i\right)}{\mathrm{E}\left(g_{1i}|x_i\right)}+2\left[\mathrm{E}\left(Y_1|x_i\right)-\mathrm{E}(Y_1)\right]\right\}+o_p\left(\frac{1}{\sqrt{n}}\right). \tag{59}$$

Use the same strategy on another term in Equation (51),

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\frac{\mathrm{E}\left(\widetilde{h}_{1i}\middle|\,x_i\right)\left[\widehat{\mathrm{E}}\left(\widehat{\widetilde{g}}_{1i}\middle|\,x_i\right)-\mathrm{E}\left(\widetilde{g}_{1i}|\,x_i\right)\right]}{\left[\mathrm{E}\left(\widetilde{g}_{1i}|\,x_i\right)\right]^2}\\
=\;&\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{\mathrm{E}\left(\widetilde{h}_{1i}\middle|\,x_i\right)\widehat{\mathrm{E}}\left(\widehat{\widetilde{g}}_{1i}\middle|\,x_i\right)}{\left[\mathrm{E}\left(\widetilde{g}_{1i}|\,x_i\right)\right]^2}-\frac{\mathrm{E}\left(\widetilde{h}_{1i}\middle|\,x_i\right)}{\mathrm{E}\left(\widetilde{g}_{1i}|\,x_i\right)}\right\}
\end{aligned}$$

from which we could get that it is equal to

$$\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{\mathrm{E}\left(h_{1i}|x_i\right)g_{1i}}{\left[\mathrm{E}\left(g_{1i}|x_i\right)\right]^2}-\frac{\mathrm{E}\left(h_{1i}|x_i\right)\mathrm{E}\left(g_{1i}|x_i,v_i\right)}{\left[\mathrm{E}\left(g_{1i}|x_i\right)\right]^2}+\left[\mathrm{E}\left(Y_1|x_i\right)-\mathrm{E}(Y_1)\right]\right\}+o_p\left(\frac{1}{\sqrt{n}}\right). \tag{60}$$

Sum up the results for $\frac{1}{n}\sum_{i=1}^{n}\widehat{\psi}_1(x_i)$. From Equation (50), (51), (59), (60), we know that

$\frac{1}{n}\sum_{i=1}^{n}\widehat{\psi}_1(x_i)$ is equal to the difference between Equation (59) and Equation(60),

$$\mathrm{E}(Y_1)+\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{h_{1i}}{\mathrm{E}\left(g_{1i}|x_i\right)}-\frac{\mathrm{E}\left(h_{1i}|x_i,v_i\right)}{\mathrm{E}\left(g_{1i}|x_i\right)}-\frac{\mathrm{E}\left(h_{1i}|x_i\right)g_{1i}}{\left[\mathrm{E}\left(g_{1i}|x_i\right)\right]^2}\right.$$
$$\left.+\frac{\mathrm{E}\left(h_{1i}|x_i\right)\mathrm{E}\left(g_{1i}|x_i,v_i\right)}{\left[\mathrm{E}\left(g_{1i}|x_i\right)\right]^2}+\left[\mathrm{E}\left(Y_1|x_i\right)-\mathrm{E}(Y_1)\right]\right\}+o_p\left(\frac{1}{\sqrt{n}}\right). \tag{61}$$

Equation (61) is very similar to Equation (9). Compared to Equation (9), the additional term $\mathrm{E}(h_{1i}|x_i,v_i)$ exists is because we also estimate f_{xv} nonparametrically.

Do the similar analysis for $\frac{1}{n}\sum_{i=1}^{n}\widehat{\psi}_2(x_i)$, it is equal to

$$\mathrm{E}(Y_0)+\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{h_{2i}}{\mathrm{E}\left(g_{2i}|x_i\right)}-\frac{\mathrm{E}\left(h_{2i}|x_i,v_i\right)}{\mathrm{E}\left(g_{2i}|x_i\right)}-\frac{\mathrm{E}\left(h_{2i}|x_i\right)g_{2i}}{\left[\mathrm{E}\left(g_{2i}|x_i\right)\right]^2}\right.$$
$$\left.+\frac{\mathrm{E}\left(h_{2i}|x_i\right)\mathrm{E}\left(g_{2i}|x_i,v_i\right)}{\left[\mathrm{E}\left(g_{2i}|x_i\right)\right]^2}+\left[\mathrm{E}\left(Y_0|x_i\right)-\mathrm{E}(Y_0)\right]\right\}+o_p\left(\frac{1}{\sqrt{n}}\right). \tag{62}$$

From Equation (61) and (62), the conclusion is proved. ■

Lemma 4.5 *a_i, b_t are random vectors that satisfy Assumption 2.10. w_{it} are random vectors and $w_{it}\perp w_{it'}|a_i$, for $t\neq t'$, $w_{it}\perp w_{i't}|b_t$ for $i\neq i'$, $w_{it}\perp w_{i't'}$ for $i\neq i'$, $t\neq t'$. $h(a_i,b_t,w_{it})$ are a real function that the first and second moment exist, and $E\left[h(a_i,b_t,w_{it})^2\right]=o(n)$. $E[h(a_i,b_t,w_{it})]=E[h(a_{i'},b_{t'},w_{i't'})]$ for any i,t,i',t'. $T\to\infty$ as $n\to\infty$. Then*

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}h(a_i,b_t,w_{it})$$

is equal to

$$E[h(a_i,b_t,w_{it})]+\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[E[h(a_i,b_t,w_{it})|a_i]+E[h(a_i,b_t,w_{it})|b_t]-2E[h(a_i,b_t,w_{it})]\right]+o_p\left(\frac{1}{\sqrt{T}}\right).$$

w_{it} are heterogeneous across t, but $\mathrm{E}(h)$ are assumed the same across t. It is not strange as it looks; one typical case that satisfies those is that $\mathrm{E}(h)=0$ for any i,t.

Proof of Lemma 4.5. Let

$$Q = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[h(a_i,b_t,w_{it}) - \mathrm{E}(h(a_i,b_t,w_{it})|a_i) - \mathrm{E}(h(a_i,b_t,w_{it})|b_t) + \mathrm{E}(h(a_i,b_t,w_{it}))\right], \quad (63)$$

then it is equivalent to show that $Q = o_p(\frac{1}{\sqrt{T}})$.

$$\mathrm{E}[Q^2] = \frac{1}{n^2T^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{i'=1}^{n}\sum_{t'=1}^{T}\mathrm{E}\left[\left(h - \mathrm{E}(h|a_i) - \mathrm{E}(h|b_t) + \mathrm{E}(h)\right)\left(h - \mathrm{E}(h|a_{i'}) - \mathrm{E}(h|b_{t'}) + \mathrm{E}(h)\right)\right].$$

For $i \neq i', t \neq t'$, the term inside summation is zero. For the case when only one index is equal to the other one, i.e., $i = i'$, $t \neq t'$, since

$$\begin{aligned}\mathrm{E}\left[h(a_i,b_t,w_{it})h(a_i,b_{t'},w_{it'})\right] &= \mathrm{E}\left[\mathrm{E}\left[h(a_i,b_t,w_{it})h(a_i,b_{t'},w_{it'})|a_i\right]\right] \\ &= \mathrm{E}\left[\mathrm{E}\left[h(a_i,b_t,w_{it})|a_i\right]\mathrm{E}[h(a_i,b_{t'},w_{it'})|a_i]\right],\end{aligned}$$

the term inside summation is zero again. So we can rewrite $\mathrm{E}[Q^2]$ as

$$\mathrm{E}[Q^2] = \frac{1}{n^2T^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\mathrm{E}\left[\left(h - \mathrm{E}(h|a_i) - \mathrm{E}(h|b_t) + \mathrm{E}(h)\right)^2\right].$$

By assumption $\mathrm{E}\big(h^2\big) = o_p(n)$, so $\mathrm{E}[Q^2] = o_p\left(\frac{1}{T}\right)$, which implies $Q = o_p\left(\frac{1}{\sqrt{T}}\right)$. ∎

Lemma 4.6 *Under the same assumptions in Lemma 4.5 and Assumption 2.9. Further assume* $var(E\left[h(a_i,b_t,w_{it})|a_i\right]) \leq M$, *for all* i, *where* M *is a finite positive number. Then*

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[E\left[h(a_i,b_t,w_{it})|a_i\right] + E\left[h(a_i,b_t,w_{it})|b_t\right] - 2E\left[h(a_i,b_t,w_{it})\right]\right]$$

is equal to $\frac{1}{T}\sum_{t=1}^{T}\left[E\left[h(a_i,b_t,w_{it})|b_t\right] - E\left[h(a_i,b_t,w_{it})\right]\right] + o_p\left(\frac{1}{\sqrt{T}}\right)$.

Proof of Lemma 4.6.

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\mathrm{E}\left[h(a_i,b_t,w_{it})|a_i\right] + \mathrm{E}\left[h(a_i,b_t,w_{it})|b_t\right] - 2\mathrm{E}(h(a_i,b_t,w_{it}))\right]$$

First by assumption that $\omega_{it}|b_t$ is i.i.d across i, we know that

$$\mathrm{E}\left[h(a_i, b_t, w_{it})|b_t\right] = \mathrm{E}\left[h(a_{i'}, b_t, w_{i't})|b_t\right],$$

which gives

$$\begin{aligned} & \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\mathrm{E}\left[h(a_i, b_t, w_{it})|b_t\right] - \mathrm{E}(h(a_i, b_t, w_{it}))\right] \\ = \; & \frac{1}{T}\sum_{t=1}^{T}\left[\mathrm{E}\left[h(a_i, b_t, w_{it})|b_t\right] - \mathrm{E}(h(a_i, b_t, w_{it}))\right] \end{aligned} \tag{64}$$

For the other part, note that

$$\begin{aligned} & \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\mathrm{E}\left[h(a_i, b_t, w_{it})|a_i\right] - \mathrm{E}(h(a_i, b_t, w_{it}))\right] \\ = \; & \frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{n}\sum_{i=1}^{n}\left[\mathrm{E}\left[h(a_i, b_t, w_{it})|a_i\right] - \mathrm{E}(h(a_i, b_t, w_{it}))\right]\right], \end{aligned}$$

where $\mathrm{E}[h(a_i, b_t, w_{it})|a_i]$ is independent across i.

$$\begin{aligned} & \mathrm{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\left[\mathrm{E}\left[h(a_i, b_t, w_{it})|a_i\right] - \mathrm{E}(h(a_i, b_t, w_{it}))\right]\right)^2\right] \\ = \; & \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}\left[\left(\left[\mathrm{E}\left[h(a_i, b_t, w_{it})|a_i\right] - \mathrm{E}(h(a_i, b_t, w_{it}))\right]\right)^2\right] \leq \frac{M}{n}, \end{aligned}$$

by Markov's inequality,

$$\frac{1}{n}\sum_{i=1}^{n}\left[\mathrm{E}\left[h(a_i, b_t, w_{it})|a_i\right] - \mathrm{E}(h(a_i, b_t, w_{it}))\right] = O_p(\frac{1}{\sqrt{n}}),$$

which gives that

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\mathrm{E}\left[h(a_i, b_t, w_{it})|a_i\right] - \mathrm{E}(h(a_i, b_t, w_{it}))\right] = O_p(\frac{1}{\sqrt{n}}). \tag{65}$$

Combining Equation (64) and Equation (65), the lemma follows. ■

Lemma 4.7 *Denote* $\zeta_n = (A_{1n}, B_{1n}, A_{2n}, B_{2n})'$, *a 4-by-1 vector, where* $A_{1n}, B_{1n}, A_{2n}, B_{2n}$ *are* *random variables that evolve as n goes to infinity. Assume that ζ_n converge in probability to $\overline{\zeta} = (0, \overline{B}_1, 0, \overline{B}_2)'$, where $\overline{B}_1 \neq 0, \overline{B}_2 \neq 0$, and*

$$\sqrt{n}\,[\zeta_n - \zeta] \xrightarrow{d} N(\mathbf{0}, \mathbf{\Omega}),$$

where $\mathbf{\Omega}$ is a positive definite matrix

$$\Omega = \begin{pmatrix} \sigma^2_{A_1} & \sigma_{A_1B_1} & \sigma_{A_1A_2} & \sigma_{A_1B_2} \\ . & \sigma^2_{B_1} & \sigma_{B_1A_2} & \sigma_{B_1B_2} \\ . & . & \sigma^2_{A_2} & \sigma_{A_2B_2} \\ . & . & . & \sigma^2_{B_2} \end{pmatrix}.$$

Then

$$\sqrt{n}\left(\frac{A_{1n}}{B_{1n}} - \frac{A_{2n}}{B_{2n}}\right) \xrightarrow{d} N\left(0, \frac{\sigma^2_{A_1}}{\overline{B}_1^2} - \frac{2\sigma_{A_1A_2}}{\overline{B}_1\overline{B}_2} + \frac{\sigma^2_{A_2}}{\overline{B}_2^2}\right).$$

Proof. The Lemma follows immediately after delta method. ■

Proof of Theorem 2.8. First note that

$$\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{1it} \xrightarrow{p} \overline{\Pi}_1,$$

$$\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{2it} \xrightarrow{p} \overline{\Pi}_2.$$

we will show the Equation (3) is equal to

$$\frac{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Lambda_{1it}}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{1it}} - \frac{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Lambda_{2it}}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{2it}} + o_p(\frac{1}{\sqrt{n}}).$$

It is enough to show that

$$\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{D_{it}\left(Y_{it} - E(\widetilde{a}_i + \widetilde{b}_t + Y_1)\right)}{\widehat{f}_{v_t}(v_{it})} = \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Lambda_{1it} + o_p(\frac{1}{\sqrt{n}}).$$

To this end,

$$
\begin{aligned}
& \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{D_{it}\left(Y_{it}-\mathrm{E}(\widetilde{a}_i+\widetilde{b}_t+Y_1)\right)}{\widehat{f}_{v_t}(v_{it})} \qquad (66)\\
= & \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{D_{it}\left(Y_{it}-\mathrm{E}(\widetilde{a}_i+\widetilde{b}_t+Y_1)\right)}{f_{v_t}(v_{it})}\\
& -\frac{D_{it}\left(Y_{it}-\mathrm{E}(\widetilde{a}_i+\widetilde{b}_t+Y_1)\right)}{f_{v_t}^2(v_{it})}\left(\widehat{f}_{v_t}(v_{it})-f_{v_t}(v_{it})\right)+R_{nit},
\end{aligned}
$$

where

$$
R_{nit}=\frac{D_{it}\left(Y_{it}-E(\widetilde{a}_i+\widetilde{b}_t+Y_1)\right)}{f_{v_t}^2(v_{it})\widehat{f}_{v_t}(v_{it})}\left(\widehat{f}_{v_t}(v_{it})-f_{v_t}(v_{it})\right)^2.
$$

Again, by Silverman (1978), $\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}|R_{nit}|=o_p\left(\frac{1}{n^{1-\varepsilon}h^2}\right)$, which is $o_p\left(\frac{1}{\sqrt{n}}\right)$, by $n^{1-\varepsilon}h^4\to\infty$.

Generalize Lemma 4.1 a little bit, by Assumption 4.6, not hard to see that,

$$
\frac{1}{n}\sum_{i=1}^{n}\frac{D_{it}\left(Y_{it}-\mathrm{E}(\widetilde{a}_i+\widetilde{b}_t+Y_1)\right)}{f_{v_t}^2(v_{it})}\left(\widehat{f}_{v_t}(v_{it})-f_{v_t}(v_{it})\right)
$$

is equal to

$$
\frac{1}{n}\sum_{i=1}^{n}\frac{\mathrm{E}\left[\left(Y_{it}-E(\widetilde{a}_i+\widetilde{b}_t+Y_1)\right)D_{it}\Big| v_{it}\right]}{f_{v_t}(v_{it})}+o_p\left(\frac{1}{\sqrt{n}}\right).
$$

Substitute this back to Equation (66), we could get $\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{D_{it}\left(Y_{it}-E(\widetilde{a}_i+\widetilde{b}_t+Y_1)\right)}{\widehat{f}_{v_t}(v_{it})}$ is equal to

$$
\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{\left(Y_{it}-E(\widetilde{a}_i+\widetilde{b}_t+Y_1)\right)D_{it}-\mathrm{E}\left[\left(Y_{it}-E(\widetilde{a}_i+\widetilde{b}_t+Y_1)\right)D_{it}|v_{it}\right]}{f_{v_t}(v_{it})}+o_p\left(\frac{1}{\sqrt{n}}\right),
$$

which is $\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Lambda_{1it}+o_p\left(\frac{1}{\sqrt{n}}\right)$. For the same reason

$$
\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{(1-D_{it})Y_{it}}{\widehat{f}_{v_t}(v_{it})}=\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Lambda_{2it}+o_p\left(\frac{1}{\sqrt{n}}\right).
$$

Therefore, we know the Equation (23) is equal to

$$\frac{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Lambda_{1it}}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{1it}} - \frac{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Lambda_{2it}}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{2it}} + o_p\left(\frac{1}{\sqrt{n}}\right).$$

Applying Lemma 4.6 on this expression, it is equivalent to

$$\frac{\frac{1}{T}\sum_{t=1}^{T}\mathrm{E}\left[\Lambda_{1it}|\, b_t, \widetilde{b}_t\right]}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{1it}} - \frac{\frac{1}{T}\sum_{t=1}^{T}\mathrm{E}\left[\Lambda_{2it}|\, b_t, \widetilde{b}_t\right]}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{2it}} + o_p\left(\frac{1}{\sqrt{T}}\right),$$

we have now shown that

$$\begin{aligned}
&\frac{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{D_{it}Y_{it}}{\widehat{f}_{v_t}(v_{it})}}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{D_{it}}{\widehat{f}_{v_t}(v_{it})}} - \frac{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{(1-D_{it})Y_{it}}{\widehat{f}_{v_t}(v_{it})}}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\frac{(1-D_{it})}{\widehat{f}_{v_t}(v_{it})}} - E(\widetilde{a}_i + \widetilde{b}_t + Y_1) + E(\widetilde{a}_i + \widetilde{b}_t + Y_0) \\
= &\frac{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Lambda_{1it}}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{1it}} - \frac{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Lambda_{2it}}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{2it}} + o_p\left(\frac{1}{\sqrt{n}}\right) \\
= &\frac{\frac{1}{T}\sum_{t=1}^{T}\mathrm{E}\left[\Lambda_{1it}|\, b_t, \widetilde{b}_t\right]}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{1it}} - \frac{\frac{1}{T}\sum_{t=1}^{T}\mathrm{E}\left[\Lambda_{2it}|\, b_t, \widetilde{b}_t\right]}{\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\Pi_{2it}} + o_p\left(\frac{1}{\sqrt{T}}\right),
\end{aligned}$$

which gives the conclusion by applying Lemma 4.7. ■

References

[1] Li, Q., and J. S. Racine (2007), "Nonparametric Econometrics: Theory and Practise," Princeton University Press.

[2] Powell, J. L., J. H. Stock, and T. M. Stoker (1989), "Semiparametric Estimation of Index Coeffients," Econometrica, 57, 1403-1430.

[3] Silverman, B. W. (1978), "Weak and Strong Uniform Consistency of the Kernel Estimate of a Density Function and its Derivatives," Annals of Statistics, 6, 177-184.