Supplementary Material for " Multiplicative Distortion Measurement Errors Linear Models with General Moment Identifiability Condition" [1]

Jun Zhang[1], Wangli Xu[2], Yujie Gai[3*]

[1] College of Mathematics and Statistics, Institute of Statistical Sciences, Shenzhen University, 518060, Shenzhen, China.
[2] Center for Applied Statistics and School of Statistics, Renmin University of China, 100872, Beijing, China
[3] School of Statistics and Mathematics, Central University of Finance and Economics, Beijing, 100081, China.

1. Conditions

We now list the assumptions needed in the proof of the following theorems.

(C1) The distortion functions $\phi(u) > 0$ and $\psi_r(u) > 0$, $r = 1, \dots, q$, for all $u \in [\mathcal{U}_L, \mathcal{U}_R]$, where $[\mathcal{U}_L, \mathcal{U}_R]$ denotes the compact support of U. Moreover, the distortion functions $\phi(u)$, $\psi_r(u)$'s, have three continuous derivatives. The density function $f_U(u)$ of the random variable U is bounded away from 0 and satisfies the Lipschitz condition of order 1 on $[\mathcal{U}_L, \mathcal{U}_R]$.

(C2) For some $s \geq 4$, $E\left(|Y|^{\max\{s\rho, s\}}\right) < \infty$, $E\left(|X_r|^{\max\{s\rho, s\}}\right) < \infty$, $r = 1, \dots, q$. The matrices $\mathbf{\Sigma}$ defined in Theorem 1 is a positive-definite matrix.

(C3) The kernel function $K(\cdot)$ is a symmetric bounded density function supported on $[-A, A]$ satisfying a Lipschitz condition. $K(\cdot)$ also has second-order continuous bounded derivatives, satisfying $K^{(j)}(\pm A) = 0$, $K^{(j)}(s) = \frac{d^j K(s)}{ds^j}$ and $\int_{-A}^{A} s^2 K(s) ds \neq 0$.

(C4) As $n \to \infty$, the bandwidth h satisfies $h^3 \log n \to 0$, $\frac{\log^2 n}{nh^2} \to 0$ and $nh^4 \to 0$.

2. A Technical Lemma

Lemma 1 Suppose $E(T|U = u) = m(u)$ and its derivatives up to second order are bounded for all $u \in [a_1, a_2]$. $E|T|^3$ exists and $\sup_u \int |t|^s f(u,t) dt < \infty$, where $f(u,t)$ is the joint density of $(U,T)^\tau$. Suppose (U_i, T_i), $i = 1, 2, \dots n$ are independent and identically distributed (i.i.d.) samples from (U,T). If condition (C3) holds true for kernel function $K(u)$, and $n^{2\epsilon - 1} h \to \infty$ for $\epsilon < 1 - s^{-1}$, we have

$$\sup_{u \in [a_1, a_2]} \left| \frac{1}{n} \sum_{i=1}^{n} K_h(U_i - u) T_i - f_U(u) m(u) - \frac{1}{2} [f_U(u) m(u)]'' \mu_2 h^2 \right| = O(\tau_{n,h}), a.s.$$

where $\mu_2 = \int K(u) u^2 du$, and $\tau_{n,h} = h^3 + \sqrt{\log n / (nh)}$.

Proof Lemma 1 can be immediately proved from the result obtained by Mack and Silverman (1982).

[1] *Corresponding Author: Yujie Gai, email: yujie_gai@126.com.

3. PROOF OF THEOREM 1

Proof Using Lemma 1, we have

$$\hat{f}_U(u) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right) = f_U(u) + \frac{\mu_2 h^2}{2} f_U^{''}(u) + O_P(\tau_{n,h}). \tag{A.1}$$

Define that $m_{|\widetilde{Y}|,\rho}(u) \stackrel{\text{def}}{=} E(|\widetilde{Y}^\rho||U = u) = \phi^\rho(u)E(|Y^\rho|)$, using (A.1) and Lemma 1, it is noted that

$$\begin{aligned} \hat{m}_{|\widetilde{Y}|,\rho}(u) - m_{|\widetilde{Y}|,\rho}(u) &= \frac{1}{nhf_U(u)}\sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right)\left(|\widetilde{Y}_i^\rho| - m_{|\widetilde{Y}|,\rho}(U_i)\right) \\ &\quad + \frac{1}{nhf_U(u)}\sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right)\left(m_{|\widetilde{Y}|,\rho}(U_i) - m_{|\widetilde{Y}|,\rho}(u)\right) \\ &\quad + O_P\left(h^2\sqrt{\frac{\log n}{nh}} + h^4 + \tau_{n,h}^2\right). \end{aligned} \tag{A.2}$$

From (A.2) and $\hat{\phi}(u) = \left(\frac{\hat{m}_{|\widetilde{Y}|,\rho}(u)}{\overline{|\widetilde{Y}^\rho|}}\right)^{1/\rho}$, we have

$$\begin{aligned} \hat{\phi}^\rho(u) - \phi^\rho(u) &= \frac{\hat{m}_{|\widetilde{Y}|,\rho}(u) - \phi^\rho(u)\overline{|\widetilde{Y}^\rho|}}{\overline{|\widetilde{Y}^\rho|}} \\ &= \frac{\hat{m}_{|\widetilde{Y},\rho|}(u) - m_{|\widetilde{Y}|,\rho}(u)}{E|Y^\rho|} - \phi^\rho(u)\frac{\overline{|\widetilde{Y}^\rho|} - E(|Y^\rho|)}{E(|Y^\rho|)} \\ &\quad + O_P(n^{-1} + n^{-1/2}h^2 + n^{-1/2}\tau_{n,h}) \end{aligned} \tag{A.3}$$

According to (A.3) and Taylor expansion, we have

$$\begin{aligned} \hat{\phi}(u) - \phi(u) &= \frac{\phi^{1-\rho}(u)}{\rho}\left(\hat{\phi}^\rho(u) - \phi^\rho(u)\right) + O_P\left(|\hat{\phi}^\rho(u) - \phi^\rho(u)|^2\right) \\ &= \phi^{1-\rho}(u)\frac{\hat{m}_{|\widetilde{Y},\rho|}(u) - m_{|\widetilde{Y}|,\rho}(u)}{\rho E|Y^\rho|} - \phi(u)\frac{\overline{|\widetilde{Y}^\rho|} - E(|Y^\rho|)}{\rho E(|Y^\rho|)} \\ &\quad + O_P(n^{-1} + n^{-1/2}h^2 + n^{-1/2}\tau_{n,h}) \end{aligned} \tag{A.4}$$

Similar to the proof of Lemma B.2 in Zhang et al. (2012), together with (A.4) and (A.2), we have

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n}\left(\hat{Y}_i - Y_i\right)M(\boldsymbol{W}_i) \\ &= -\frac{1}{n}\sum_{i=1}^{n}\frac{Y_i M(\boldsymbol{W}_i)}{\phi(U_i)}\left(\hat{\phi}(U_i) - \phi(U_i)\right) + o_P(n^{-1/2}) \\ &= -\frac{1}{n}\sum_{i=1}^{n}\frac{Y_i M(\boldsymbol{W}_i)}{\phi^\rho(U_i)\rho E|Y^\rho|}\left(\hat{m}_{|\widetilde{Y},\rho|}(U_i) - m_{|\widetilde{Y}|,\rho}(U_i)\right) \\ &\quad + \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i M(\boldsymbol{W}_i)}{\rho E|Y^\rho|}\left(\overline{|\widetilde{Y}^\rho|} - E(|Y^\rho|)\right) + o_P(n^{-1/2}) \\ &\stackrel{\text{def}}{=} \mathcal{A}_{n1} + \mathcal{A}_{n2}. \end{aligned} \tag{A.5}$$

Recalling that $\overline{|\widetilde{Y}^\rho|} = \frac{1}{n}\sum_{i=1}^n |\widetilde{Y}_i^\rho|$, we have

$$\mathcal{A}_{n2} = \frac{E[YM(\boldsymbol{W})]}{\rho E|Y^\rho|}\frac{1}{n}\sum_{i=1}^n \left(|\widetilde{Y}_i^\rho| - E|Y^\rho|\right) + o_P(n^{-1/2}). \tag{A.6}$$

For $\mathcal{A}_{n1}$, using asymptotic expression (A.2), we have

$$\begin{aligned}
&\mathcal{A}_{n1} \qquad\qquad\qquad (A.7)\\
&= -\frac{1}{\rho E|Y^\rho|n^2h}\sum_{i=1}^n\sum_{j=1}^n \frac{Y_iM(\boldsymbol{W}_i)}{\phi^\rho(U_i)f_U(U_i)}K\left(\frac{U_i-U_j}{h}\right)\left(|\widetilde{Y}_j^\rho| - m_{|\widetilde{Y}|,\rho}(U_j)\right)\\
&-\frac{1}{\rho E|Y^\rho|n^2h}\sum_{i=1}^n\sum_{j=1}^n \frac{Y_iM(\boldsymbol{W}_i)}{\phi^\rho(U_i)f_U(U_i)}K\left(\frac{U_i-U_j}{h}\right)\left(m_{|\widetilde{Y}|,\rho}(U_j) - m_{|\widetilde{Y}|,\rho}(U_i)\right)\\
&\stackrel{\text{def}}{=} \mathcal{A}_{n1}[1] + \mathcal{A}_{n1}[2].
\end{aligned}$$

For $\mathcal{A}_{n1}[1]$, as $nh^4 \to 0$, the projection of U-statistic (Serfling; 1980) entails that

$$\begin{aligned}
\mathcal{A}_{n1}[1] &= -\frac{E[YM(\boldsymbol{W})]}{\rho E|Y^\rho|n}\sum_{i=1}^n \frac{|\widetilde{Y}_i^\rho| - m_{|\widetilde{Y}|,\rho}(U_i)}{\phi^\rho(U_i)} + O_P(h^2) \qquad (A.8)\\
&= -\frac{E[YM(\boldsymbol{W})]}{\rho E|Y^\rho|}\frac{1}{n}\sum_{i=1}^n (|Y_i^\rho| - E|Y^\rho|) + o_P(n^{-1/2}).
\end{aligned}$$

Similar to (A.8), the the projection of U-statistic (Serfling; 1980) entails $\mathcal{A}_{n1}[2] = O_P(h^2) = o_P(n^{-1/2})$ as $nh^4 \to 0$. Together with (A.5)-(A.8), we have

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \left(\hat{Y}_i - Y_i\right)M(\boldsymbol{W}_i) \qquad\qquad (A.9)\\
&= \frac{1}{n}\sum_{i=1}^n \left(|\widetilde{Y}_i^\rho| - |Y_i^\rho|\right)\frac{E[YM(\boldsymbol{W})]}{\rho E|Y^\rho|} + o_P(n^{-1/2})\\
&= \frac{1}{n}\sum_{i=1}^n \frac{|Y_i^\rho|}{\rho E|Y^\rho|}\left[\phi^\rho(U_i) - 1\right]E[YM(\boldsymbol{W})] + o_P(n^{-1/2}).
\end{aligned}$$

4. PROOF OF THEOREM 2

Define that $m_{|\widetilde{X}_r|,\rho}(u) \stackrel{\text{def}}{=} E(|\widetilde{X}_r^\rho||U=u) = \psi_r^\rho(u)E(|X_r^\rho|)$. Similar to (A.2)-(A.4), for $s = 1,\ldots,p$, we have

$$\begin{aligned}
\hat{m}_{|\widetilde{X}_r|,\rho}(u) - m_{|\widetilde{X}_r|,\rho}(u) &= \frac{1}{nhf_U(u)}\sum_{i=1}^n K\left(\frac{U_i-u}{h}\right)\left(|\widetilde{X}_{ri}^\rho| - m_{|\widetilde{X}_r|,\rho}(U_i)\right) \qquad (A.10)\\
&+\frac{1}{nhf_U(u)}\sum_{i=1}^n K\left(\frac{U_i-u}{h}\right)\left(m_{|\widetilde{X}_r|,\rho}(U_i) - m_{|\widetilde{X}_r|,\rho}(u)\right)\\
&+O_P\left(h^2\sqrt{\frac{\log n}{nh}} + h^4 + \tau_{n,h}^2\right),
\end{aligned}$$

and

$$\begin{aligned}
&\hat{\psi}_r(u) - \psi_r(u) \qquad (A.11)\\
&= \psi_r^{1-\rho}(u)\frac{\hat{m}_{|\widetilde{X}_r|,\rho}(u) - m_{|\widetilde{X}_r|,\rho}(u)}{\rho E|X_r^\rho|} - \psi_r(u)\frac{\overline{|\widetilde{X}_r^\rho|} - E(|X_r^\rho|)}{\rho E(|X_r^\rho|)}\\
&+O_P(n^{-1} + n^{-1/2}h^2 + n^{-1/2}\tau_{n,h}).
\end{aligned}$$

Using (A.4) and (A.11), as $nh^8 \to 0$ and $\frac{\log^2 n}{nh^2} \to 0$, we have

$$\frac{1}{n}\sum_{i=1}^n(\hat{X}_{ri} - X_{ri})(\hat{X}_{si} - X_{si}) = O_P((n^{-1/2} + h^2 + \tau_{n,h})^2) = o_P(n^{-1/2}), \qquad (A.12)$$

$$\frac{1}{n}\sum_{i=1}^n(\hat{X}_{ri} - X_{ri})(\hat{Y}_i - Y_i) = O_P((n^{-1/2} + h^2 + \tau_{n,h})^2) = o_P(n^{-1/2}). \qquad (A.13)$$

Using (A.11)-(A.12) and Theorem 1, we have

$$\begin{aligned}
\boldsymbol{B}_n &\overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^n[(1, \hat{\boldsymbol{X}}_i^{\mathrm{T}})^{\mathrm{T}}]^{\otimes 2} = \frac{1}{n}\sum_{i=1}^n[(1, \boldsymbol{X}_i^{\mathrm{T}})^{\mathrm{T}}]^{\otimes 2} \qquad (A.14)\\
&+\frac{1}{n}\sum_{i=1}^n(0, \hat{\boldsymbol{X}}_i^{\mathrm{T}} - \boldsymbol{X}_i^{\mathrm{T}})^{\mathrm{T}}(1, \boldsymbol{X}_i^{\mathrm{T}}) + \frac{1}{n}\sum_{i=1}^n(1, \boldsymbol{X}_i^{\mathrm{T}})^{\mathrm{T}}(0, \hat{\boldsymbol{X}}_i^{\mathrm{T}} - \boldsymbol{X}_i^{\mathrm{T}})\\
&+O_P((n^{-1/2} + h^2 + \tau_{n,h})^2)\\
&= \boldsymbol{\Sigma} + o_P(1).
\end{aligned}$$

Using Theorem 1, $\varepsilon_i = Y_i - \boldsymbol{X}^{\mathrm{T}}\boldsymbol{\beta}_0$ and $E(\varepsilon|\boldsymbol{X}) = 0$,

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n\varepsilon_i(\hat{X}_{ri} - X_{ri}) = \frac{1}{n}\sum_{i=1}^n(Y_i - \boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)(\hat{X}_{ri} - X_{ri}) \qquad (A.15)\\
&= \frac{1}{n}\sum_{i=1}^n\left(|\widetilde{X}_{ri}^\rho| - |X_{ri}^\rho|\right)\frac{E\left[X_r(Y - \boldsymbol{X}^{\mathrm{T}}\boldsymbol{\beta}_0)\right]}{\rho E(|X_r^\rho|)} + o_P(n^{-1/2})\\
&= \frac{1}{n}\sum_{i=1}^n\left(|\widetilde{X}_{ri}^\rho| - |X_{ri}^\rho|\right)\frac{E\left[X_r\varepsilon\right]}{\rho E(|X_r^\rho|)} + o_P(n^{-1/2}) = o_P(n^{-1/2}).
\end{aligned}$$

It is easily seen that

$$\begin{aligned}
\begin{pmatrix}\hat{\alpha}\\ \hat{\boldsymbol{\beta}}\end{pmatrix} - \begin{pmatrix}\alpha_0\\ \boldsymbol{\beta}_0\end{pmatrix} &= \boldsymbol{B}_n^{-1}\frac{1}{n}\sum_{i=1}^n(1, \hat{\boldsymbol{X}}_i^{\mathrm{T}})^{\mathrm{T}}\left\{\hat{Y}_i - \alpha_0 - \boldsymbol{\beta}_0^{\mathrm{T}}\hat{\boldsymbol{X}}_i\right\} \qquad (A.16)\\
&= T_{n,1} + T_{n,2} - T_{n,3}.
\end{aligned}$$

Here, using (A.12)-(A.15), we have

$$\begin{aligned}
T_{n,1} &\overset{\text{def}}{=} \boldsymbol{B}_n^{-1}\frac{1}{n}\sum_{i=1}^n(1, \boldsymbol{X}_i^{\mathrm{T}})^{\mathrm{T}}\varepsilon_i + \boldsymbol{B}_n^{-1}\frac{1}{n}\sum_{i=1}^n(0, \hat{\boldsymbol{X}}_i^{\mathrm{T}} - \boldsymbol{X}_i^{\mathrm{T}})^{\mathrm{T}}\varepsilon_i \qquad (A.17)\\
&= \boldsymbol{\Sigma}^{-1}\frac{1}{n}\sum_{i=1}^n(1, \boldsymbol{X}_i^{\mathrm{T}})^{\mathrm{T}}\varepsilon_i + o_P(n^{-1/2}).
\end{aligned}$$

Then, using (A.13), Theorem 1 and $E[(1,\boldsymbol{X}^{\mathrm{T}})^{\mathrm{T}}Y]=\boldsymbol{\Sigma}(\alpha_0,\boldsymbol{\beta}_0^{\mathrm{T}})^{\mathrm{T}}$, we have

$$\begin{aligned} T_{n,2} &\stackrel{\text{def}}{=} \boldsymbol{B}_n^{-1}\frac{1}{n}\sum_{i=1}^{n}(1,\hat{\boldsymbol{X}}_i^{\mathrm{T}})^{\mathrm{T}}\{\hat{Y}_i-Y_i\} \qquad (A.18)\\ &= \boldsymbol{B}_n^{-1}\frac{1}{n}\sum_{i=1}^{n}(1,\boldsymbol{X}_i^{\mathrm{T}})^{\mathrm{T}}\{\hat{Y}_i-Y_i\}+o_P(n^{-1/2})\\ &= \boldsymbol{\Sigma}^{-1}\frac{1}{n}\sum_{i=1}^{n}(|\widetilde{Y}_i^{\rho}|-|Y_i^{\rho}|)\frac{E[(1,\boldsymbol{X}^{\mathrm{T}})^{\mathrm{T}}Y]}{\rho E(|Y^{\rho}|)}+o_P(n^{-1/2})\\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{|\widetilde{Y}_i^{\rho}|-|Y_i^{\rho}|}{\rho E(|Y^{\rho}|)}(\alpha_0,\boldsymbol{\beta}_0^{\mathrm{T}})^{\mathrm{T}}+o_P(n^{-1/2})\\ &= \frac{1}{n}\sum_{i=1}^{n}(\phi^{\rho}(U_i)-1)\frac{|Y_i^{\rho}|}{\rho E(|Y^{\rho}|)}(\alpha_0,\boldsymbol{\beta}_0^{\mathrm{T}})^{\mathrm{T}}+o_P(n^{-1/2}). \end{aligned}$$

Moreover, e_{s+1} is a $(p+1)$-dimensional vector with 1 in the $s+1$-th position and 0's elsewhere, $s=1,\ldots,p$,

$$\begin{aligned} &T_{n,3} \qquad (A.19)\\ &\stackrel{\text{def}}{=} \boldsymbol{B}_n^{-1}\frac{1}{n}\sum_{i=1}^{n}(1,\hat{\boldsymbol{X}}_i^{\mathrm{T}})^{\mathrm{T}}(\hat{\boldsymbol{X}}_i-\boldsymbol{X}_i)^{\mathrm{T}}\boldsymbol{\beta}_0\\ &= \boldsymbol{B}_n^{-1}\frac{1}{n}\sum_{i=1}^{n}(1,\boldsymbol{X}_i^{\mathrm{T}})^{\mathrm{T}}(\hat{\boldsymbol{X}}_i-\boldsymbol{X}_i)^{\mathrm{T}}\boldsymbol{\beta}_0+o_P(n^{-1/2})\\ &= \boldsymbol{\Sigma}^{-1}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{p}(|\widetilde{X}_{si}^{\rho}|-|X_{si}^{\rho}|)\frac{E[(1,\boldsymbol{X}^{\mathrm{T}})^{\mathrm{T}}X_s]\beta_{0s}}{\rho E(|X_s^{\rho}|)}+o_P(n^{-1/2})\\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{p}\frac{(|\widetilde{X}_{si}^{\rho}|-|X_{si}^{\rho}|)}{\rho E(|X_s^{\rho}|)}e_{s+1}e_{s+1}^{\mathrm{T}}(\alpha_0,\boldsymbol{\beta}_0^{\mathrm{T}})^{\mathrm{T}}+o_P(n^{-1/2})\\ &= \frac{1}{n}\sum_{i=1}^{n}\mathrm{diag}\left(0,(\psi_1^{\rho}(U_i)-1)\frac{|X_1^{\rho}|}{E(\rho|X_1^{\rho}|)},\ldots,(\psi_p^{\rho}(U_i)-1)\frac{|X_p^{\rho}|}{\rho E(|X_p^{\rho}|)}\right)(\alpha_0,\boldsymbol{\beta}_0^{\mathrm{T}})^{\mathrm{T}}\\ &\quad +o_P(n^{-1/2}). \end{aligned}$$

From (A.16)-(A.19), it is easily seen that

$$\begin{aligned} &\sqrt{n}\left\{\begin{pmatrix}\hat{\alpha}\\ \hat{\boldsymbol{\beta}}\end{pmatrix}-\begin{pmatrix}\alpha_0\\ \boldsymbol{\beta}_0\end{pmatrix}\right\} \qquad (A.20)\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{\boldsymbol{\Sigma}^{-1}\begin{pmatrix}1\\ \boldsymbol{X}_i\end{pmatrix}\varepsilon_i+\frac{(\phi^{\rho}(U_i)-1)|Y_i^{\rho}|}{\rho E(|Y^{\rho}|)}\begin{pmatrix}\alpha_0\\ \boldsymbol{\beta}_0\end{pmatrix}\right\}\\ &\quad -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathrm{diag}\left(0,\frac{(\psi_1^{\rho}(U_i)-1)|X_{1i}^{\rho}|}{\rho E(|X_1^{\rho}|)},\ldots,\frac{(\psi_p^{\rho}(U_i)-1)|X_{pi}^{\rho}|}{\rho E(|X_p^{\rho}|)}\right)\begin{pmatrix}\alpha_0\\ \boldsymbol{\beta}_0\end{pmatrix}\\ &\quad +o_P(1). \end{aligned}$$

According to the asymptotic expression (A.20), we have completed the proof of Theorem 2.

5. PROOF OF THEOREM 3

For $1 \leq s \leq q$, let $\hat{\wp}_{n,i}^{[s]}(\boldsymbol{\beta}_0)$ be the s-component of $\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)$. We decompose $\hat{\wp}_{n,i}^{[s]}(\boldsymbol{\beta}_0)$ into following terms:

$$\hat{\wp}_{n,i}^{[s]}(\boldsymbol{\beta}_0) = (Y_i - \alpha_0 - \boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)X_{si} + \sum_{t=1}^{7} Q_{n,it}^{[s]},$$

where,

$$\begin{aligned}
Q_{n,i1}^{[s]} &= (\hat{Y}_i - Y_i)X_{si}, \quad Q_{n,i2}^{[s]} = (\alpha_0 - \hat{\alpha})X_{si}, \\
Q_{n,i3}^{[s]} &= (Y_i - \alpha_0 - \boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)\Big(\hat{X}_{si} - X_{si}\Big), \\
Q_{n,i4}^{[s]} &= \boldsymbol{\beta}_0^{\mathrm{T}}\left(\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i\right)X_{si}, \quad Q_{n,i5}^{[s]} = (\alpha_0 - \hat{\alpha})\Big(\hat{X}_{si} - X_{si}\Big) \\
Q_{n,i6}^{[s]} &= (\hat{Y}_i - Y_i)\Big(\hat{X}_{si} - X_{si}\Big), \\
Q_{n,i7}^{[s]} &= \boldsymbol{\beta}_0^{\mathrm{T}}\left(\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i\right)\Big(\hat{X}_{si} - X_{si}\Big).
\end{aligned}$$

To prove Theorem 3, we need to show that

$$\max_{1\leq i\leq n} |\hat{\wp}_{n,it}^{[s]}| = o_P(n^{1/2}), \quad t = 1, \ldots, 7.$$

It is noted that for any sequence of $i.i.d$ random $\{V_i, 1 \leq i \leq n\}$ and $E[V^2] \leq \infty$, we have $\max_{1\leq i\leq n} \frac{|V_i|}{\sqrt{n}} \to 0, a.s.$. Then,

$$\max_{1\leq i\leq n} \Big|(Y_i - \alpha_0 - \boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)X_{si}\Big| = o_P(n^{1/2}).$$

Next, for $Q_{n,i1}^{[s]}$, directly using (A.2) and (A.4), we have

$$\begin{aligned}
\max_{1\leq i\leq n} |Q_{n,i1}^{[s]}| &\leq \max_{1\leq i\leq n} \Big|\phi(U_i) - \hat{\phi}_M(U_i)\Big| \max_{1\leq i\leq n} \left|\frac{Y_i X_{si}}{\phi(U_i)}\right| \\
&\quad + O_P\left(h^4 + \frac{\log n}{nh}\right) O_P(n^{1/2}) = o_P(n^{1/2}).
\end{aligned}$$

Using Theorem 2, we have $(\alpha_0 - \hat{\alpha}) = O_P(n^{-1/2})$, thus,

$$\max_{1\leq i\leq n} |Q_{n,i2}^{[s]}| = |\alpha_0 - \hat{\alpha}| \max_{1\leq i\leq n} |X_{si}| = o_P(n^{1/2}).$$

Similar to the proofs of $|Q_{n,i1}^{[s]}(\boldsymbol{\beta}_0)|$ and $|Q_{n,i2}^{[s]}(\boldsymbol{\beta}_0)|$, we have $\max_{1\leq i\leq n} |Q_{n,ij}^{[s]}(\boldsymbol{\beta}_0)| = o_P(n^{1/2})$ for $j = 3, \ldots, 7$. Followed the same argument in the proof (2.14) in Owen (1991), we have $\hat{\lambda} = O_P(n^{1/2})$. Thus, $\max_{1\leq i\leq n} |\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)| = o_P(1)$. Using $\log(1+t) \approx t - \frac{1}{2}t^2$ for t sufficiently small, we have

$$\hat{l}(\boldsymbol{\beta}_0) = 2\sum_{i=1}^{n}\left(\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0) - \frac{1}{2}\{\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\}^2\right) + o_P(1). \tag{A.21}$$

Note that $\hat{\lambda}$ satisfies the following equation,

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)}{1 + \hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)} = \mathbf{0}.$$

Moreover,

$$\begin{aligned}\mathbf{0} &= \frac{1}{n}\sum_{i=1}^{n}\frac{\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)}{1+\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)}\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)-\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)^{\mathrm{T}}\hat{\lambda}\\ &\quad+\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\{\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\}^2}{1+\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)}. \qquad \text{(A.22)}\end{aligned}$$

The equation (A.22) and $\max_{1\le i\le n}|\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)|=o_P(1)$ entail that

$$\hat{\lambda}=\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)^{\mathrm{T}}\Big)^{-1}\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)+o_P(n^{-1/2}). \qquad \text{(A.23)}$$

Plugging the asymptotic expression (A.23) to (A.21), we have

$$\begin{aligned}\hat{l}(\boldsymbol{\beta}_0) &= n\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\Big)^{\mathrm{T}}\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)^{\mathrm{T}}\Big)^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\Big)\\ &\quad+o_P(1).\end{aligned}$$

We obtain that

$$\hat{l}(\boldsymbol{\beta}_0)=n\Big(\frac{1}{n}\sum_{i=1}^{n}\kappa_{n,i}(\boldsymbol{\beta}_0)\Big)^{\mathrm{T}}\Big(\frac{1}{n}\sum_{i=1}^{n}\kappa_{n,i}(\boldsymbol{\beta}_0)\kappa_{n,i}(\boldsymbol{\beta}_0)^{\mathrm{T}}\Big)^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n}\kappa_{n,i}(\boldsymbol{\beta}_0)\Big)+o_P(1),$$

where $\kappa_{n,i}(\boldsymbol{\beta}_0)=(Y_i-\alpha_0-\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0)\boldsymbol{X}_i$ is independent and identically distributed p-dimensional random vector with zero mean.

6. Proof of Theorem 4

The proof of Theorem 4 is similar to the proof of Theorem 2 in Zhang et al. (2013). We present the main step in the following. Let $K_{h,ij}=K_h(U_i-U_j)$, $\beta_{00}(U_i)=\alpha_0(U_i)$, $\widetilde{X}_{0i}\equiv 1$, $i,j=1,\ldots,n$. Define

$$\begin{aligned}\boldsymbol{\gamma}(U_i) &= (\beta_{00}(U_i),\beta_{01}(U_i),\ldots,\beta_{0p}(U_i))^{\mathrm{T}},\\ G_{n,i} &= \frac{1}{n}\sum_{j=1}^{n}\Big[(1,\widetilde{\boldsymbol{X}}_j^{\mathrm{T}})^{\mathrm{T}}\Big]^{\otimes 2}K_{h,ij},\\ G_{\widetilde{\boldsymbol{X}}}(u) &= E\left\{\Big[(1,\widetilde{\boldsymbol{X}}^{\mathrm{T}})^{\mathrm{T}}\Big]^{\otimes 2}|U=u\right\}.\end{aligned}$$

Using Lemma 1, we have

$$G_{n,i}=G_{\widetilde{\boldsymbol{X}}}(U_i)f_U(U_i)+\frac{h^2[G_{\widetilde{\boldsymbol{X}}}(U_i)f_U(U_i)]''\mu_2}{2}+O_P(\tau_{n,h}), \qquad \text{(A.24)}$$

where $[G_{\widetilde{\boldsymbol{X}}}(u)f_U(u)]''$ is a $(p+1)\times(p+1)$ matrix, and its (s,t)-th position element is defined as $\frac{d^2\{e_s^{\mathrm{T}}G_{\widetilde{\boldsymbol{X}}}(u)e_tf_U(u)\}}{du^2}$. Recalling that for $i,j=1,\ldots,n$, we have

$$\widetilde{Y}_j=[1,\widetilde{\boldsymbol{X}}_j^{\mathrm{T}}](\boldsymbol{\gamma}(U_j)-\boldsymbol{\gamma}(U_i))+[1,\widetilde{\boldsymbol{X}}_j^{\mathrm{T}}]\boldsymbol{\gamma}(U_i)+\epsilon(U_j).$$

For $r = 0, 1, \ldots, p$, we have

$$
\begin{aligned}
& e_{r+1}^{\mathrm{T}} G_{n,i}^{-1} \frac{1}{n} \sum_{j=1}^{n} (1, \widetilde{\boldsymbol{X}}_j^{\mathrm{T}})^{\mathrm{T}} \widetilde{Y}_j K_{h,ij} \qquad \text{(A.25)} \\
&= e_{r+1}^{\mathrm{T}} \boldsymbol{\gamma}(U_i) + e_{r+1}^{\mathrm{T}} G_{n,i}^{-1} \frac{1}{n} \sum_{j=1}^{n} [(1, \widetilde{\boldsymbol{X}}_j^{\mathrm{T}})^{\mathrm{T}}]^{\otimes 2} K_{h,ij} (\boldsymbol{\gamma}(U_j) - \boldsymbol{\gamma}(U_i)) \\
&\quad + e_{r+1}^{\mathrm{T}} G_{n,i}^{-1} \frac{1}{n} \sum_{j=1}^{n} (1, \widetilde{\boldsymbol{X}}_j^{\mathrm{T}})^{\mathrm{T}} K_{h,ij} \phi(U_j) \epsilon_j \\
&= e_{r+1}^{\mathrm{T}} \boldsymbol{\gamma}(U_i) + \mathcal{V}_{n,i,r}[1] + \mathcal{V}_{n,i,r}[2].
\end{aligned}
$$

Using Lemma 1, we have

$$
\begin{aligned}
\mathcal{V}_{n,i,r}[1] &= \frac{h^2 \mu_2}{2} e_{r+1}^{\mathrm{T}} \left[\boldsymbol{\gamma}''(U_i) + 2 G'_{\widetilde{\boldsymbol{X}}}(U_i) \boldsymbol{\gamma}'(U_i) f_U(U_i) + 2 G_{\widetilde{\boldsymbol{X}}}(U_i) \boldsymbol{\gamma}'(U_i) f'_U(U_i) \right] \\
&\quad + O_P(h^4 + h^2 \tau_{n,h}), \qquad \text{(A.26)} \\
\mathcal{V}_{n,i,r}[2] &= e_{r+1}^{\mathrm{T}} f_U^{-1}(U_i) G_{\widetilde{\boldsymbol{X}}}^{-1}(U_i) \frac{1}{n} \sum_{j=1}^{n} (1, \widetilde{\boldsymbol{X}}_j^{\mathrm{T}})^{\mathrm{T}} K_{h,ij} \phi(U_j) \epsilon_j + O_P(h^2 \tau_{n,h}). \qquad \text{(A.27)}
\end{aligned}
$$

Using (A.26)-(A.27), as $nh^8 \to 0$, $\frac{\log n}{nh^2} \to 0$, Taylor expansion entails that

$$
\begin{aligned}
& \frac{1}{n \overline{|\widetilde{X}_r^{\rho}|}} \sum_{i=1}^{n} |\widetilde{X}_{ri}^{\rho}| \left[e_{r+1}^{\mathrm{T}} G_{n,i}^{-1} \frac{1}{n} \sum_{j=1}^{n} (1, \widetilde{\boldsymbol{X}}_j^{\mathrm{T}})^{\mathrm{T}} \widetilde{Y}_j K_{h,ij} \right]^{\rho} \qquad \text{(A.28)} \\
&= \frac{1}{n \overline{|\widetilde{X}_r^{\rho}|}} \sum_{i=1}^{n} |\widetilde{X}_{ri}^{\rho}| \left[e_{r+1}^{\mathrm{T}} \boldsymbol{\gamma}(U_i) + \mathcal{V}_{n,i,r}[1] + \mathcal{V}_{n,i,r}[2] \right]^{\rho} \\
&= \frac{1}{n \overline{|\widetilde{X}_r^{\rho}|}} \sum_{i=1}^{n} |\widetilde{X}_{ri}^{\rho}| \left[e_{r+1}^{\mathrm{T}} \boldsymbol{\gamma}(U_i) \right]^{\rho} + \frac{\rho}{n \overline{|\widetilde{X}_r^{\rho}|}} \sum_{i=1}^{n} |\widetilde{X}_{ri}^{\rho}| \left[e_{r+1}^{\mathrm{T}} \boldsymbol{\gamma}(U_i) \right]^{\rho-1} \mathcal{V}_{n,i,r}[1] \\
&\quad + \frac{\rho}{n \overline{|\widetilde{X}_r^{\rho}|}} \sum_{i=1}^{n} |\widetilde{X}_{ri}^{\rho}| \left[e_{r+1}^{\mathrm{T}} \boldsymbol{\gamma}(U_i) \right]^{\rho-1} \mathcal{V}_{n,i,r}[2] + o_P(n^{-1/2}) \\
&\stackrel{\mathrm{def}}{=} \mathcal{S}_{n1,r} + \mathcal{S}_{n2,r} + \mathcal{S}_{n3,r}.
\end{aligned}
$$

For $r = 0$, recalling that $\widetilde{X}_{0i} \equiv 1$ and $e_1^{\mathrm{T}} \boldsymbol{\gamma}(U_i) = \alpha_0 \phi(U_i)$, as $nh^4 \to 0$, we have

$$
\begin{aligned}
& \hat{\alpha}_{\mathrm{Vc}}^{\rho} - \alpha_0^{\rho} = \mathcal{S}_{n1,0} + \mathcal{S}_{n2,0} + \mathcal{S}_{n3,0} - \alpha_0^{\rho} \qquad \text{(A.29)} \\
&= \alpha_0^{\rho} \left(\frac{1}{n} \sum_{i=1}^{n} \phi^{\rho}(U_i) - 1 \right) + \frac{\rho \alpha_0^{\rho-1} \mu_2 h^2}{2n} \sum_{i=1}^{n} \phi^{\rho-1}(U_i) \phi''(U_i) \\
&\quad + \frac{\rho \alpha_0^{\rho-1} \mu_2 h^2}{n} \sum_{i=1}^{n} \phi^{\rho-1}(U_i) e_1^{\mathrm{T}} \left[G'_{\widetilde{\boldsymbol{X}}}(U_i) \boldsymbol{\gamma}'(U_i) f_U(U_i) + G_{\widetilde{\boldsymbol{X}}}(U_i) \boldsymbol{\gamma}'(U_i) f'_U(U_i) \right] \\
&\quad + \frac{\rho \alpha_0^{\rho-1}}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\phi^{\rho-1}(U_i)}{f_U(U_i)} e_1^{\mathrm{T}} G_{\widetilde{\boldsymbol{X}}}^{-1}(U_i) (1, \widetilde{\boldsymbol{X}}_j^{\mathrm{T}})^{\mathrm{T}} K_{h,ij} \phi(U_j) \epsilon_j \\
&\quad + O_P(h^4 + (\log n)/(nh)) \\
&= \alpha_0^{\rho} \left(\frac{1}{n} \sum_{i=1}^{n} \phi^{\rho}(U_i) - 1 \right) + \frac{\rho \alpha_0^{\rho-1}}{n} \sum_{i=1}^{n} \phi^{\rho}(U_i) e_1^{\mathrm{T}} G_{\widetilde{\boldsymbol{X}}}^{-1}(U_i) (1, \widetilde{\boldsymbol{X}}_i^{\mathrm{T}})^{\mathrm{T}} \epsilon_i + o_P(n^{-1/2}).
\end{aligned}
$$

For $r \geq 1$, we choose $r = 1$ as an illustration. Note that $\widetilde{X}_{1i} = \psi_1(U_i)X_{1i}$ and $e_2^{\mathrm{T}}\boldsymbol{\gamma}(U_i) = \beta_{01}(U_i) = \beta_{01}\frac{\phi(U_i)}{\psi_1(U_i)}$, as $nh^4 \to 0$, we have

$$\begin{aligned}
\mathcal{S}_{n1,1} &= \frac{1}{n\overline{|\widetilde{X}_1^\rho|}}\sum_{i=1}^n |\widetilde{X}_{1i}^\rho| \left[e_2^{\mathrm{T}}\boldsymbol{\gamma}(U_i)\right]^\rho = \frac{\beta_{01}^\rho}{n\overline{|\widetilde{X}_1^\rho|}}\sum_{i=1}^n |X_{1i}^\rho|\phi^\rho(U_i) \qquad \text{(A.30)}\\
&= \frac{\beta_{01}^\rho}{n}\sum_{i=1}^n \frac{|X_{1i}^\rho|}{E|X_1^\rho|}\phi^\rho(U_i) - \frac{\beta_{01}^\rho}{n}\sum_{i=1}^n\left(\frac{|\widetilde{X}_{1i}^\rho|}{E|X_1^\rho|} - 1\right) + o_P(n^{-1/2})\\
&= \frac{\beta_{01}^\rho}{n}\sum_{i=1}^n \frac{|X_{1i}^\rho|}{E|X_1^\rho|}\left[\phi^\rho(U_i) - \psi_1^\rho(U_i)\right] + \beta_{01}^\rho + o_P(n^{-1/2}).
\end{aligned}$$

Similar to the analysis of (A.29), we have $\mathcal{S}_{n1,2} = O_P(h^2) = o_P(n^{-1/2})$ as $nh^4 \to 0$. For the term $\mathcal{S}_{n1,3}$, we have

$$\begin{aligned}
\mathcal{S}_{n1,3} &= \frac{\rho}{n\overline{|\widetilde{X}_1^\rho|}}\sum_{i=1}^n |\widetilde{X}_{1i}^\rho| \left[e_2^{\mathrm{T}}\boldsymbol{\gamma}(U_i)\right]^{\rho-1}\mathcal{V}_{n,i,1}[2] \qquad \text{(A.31)}\\
&= \frac{\rho\beta_{01}^{\rho-1}}{n^2\overline{|\widetilde{X}_1^\rho|}}\sum_{i=1}^n\sum_{j=1}^n |X_{1i}^\rho|\psi_1(U_i)\frac{\phi^{\rho-1}(U_i)}{f_U(U_i)}e_2^{\mathrm{T}}G_{\widetilde{\boldsymbol{X}}}^{-1}(U_i)(1,\widetilde{\boldsymbol{X}}_j^{\mathrm{T}})^{\mathrm{T}}K_{h,ij}\phi(U_j)\epsilon_j\\
&\quad + o_P(n^{-1/2})\\
&= \frac{\rho\beta_{01}^{\rho-1}}{n}\sum_{i=1}^n \psi_1(U_i)\phi^\rho(U_i)e_2^{\mathrm{T}}G_{\widetilde{\boldsymbol{X}}}^{-1}(U_i)(1,\widetilde{\boldsymbol{X}}_i^{\mathrm{T}})^{\mathrm{T}}\epsilon_i + o_P(n^{-1/2}).
\end{aligned}$$

According to (A.30)-(A.31), we have

$$\begin{aligned}
&\hat{\beta}_{\mathrm{Vc},1}^\rho - \beta_{01}^\rho \qquad \text{(A.32)}\\
&= \frac{\beta_{01}^\rho}{n}\sum_{i=1}^n \frac{|X_{1i}^\rho|}{E|X_1^\rho|}\left[\phi^\rho(U_i) - \psi_1^\rho(U_i)\right]\\
&\quad + \frac{\rho\beta_{01}^{\rho-1}}{n}\sum_{i=1}^n \psi_1(U_i)\phi^\rho(U_i)e_2^{\mathrm{T}}G_{\widetilde{\boldsymbol{X}}}^{-1}(U_i)(1,\widetilde{\boldsymbol{X}}_i^{\mathrm{T}})^{\mathrm{T}}\epsilon_i + o_P(n^{-1/2}).
\end{aligned}$$

Using Taylor expansion and Delta theorem, we have

$$\begin{aligned}
&\hat{\alpha}_{\mathrm{Vc}} - \alpha_0 = \frac{\alpha_0^{1-\rho}}{\rho}(\hat{\alpha}_{\mathrm{Vc}}^\rho - \alpha_0^\rho) + o_P(n^{-1/2}) \qquad \text{(A.33)}\\
&= \frac{\alpha_0}{\rho}\left(\frac{1}{n}\sum_{i=1}^n \phi^\rho(U_i) - 1\right) + \frac{1}{n}\sum_{i=1}^n \phi^\rho(U_i)e_1^{\mathrm{T}}G_{\widetilde{\boldsymbol{X}}}^{-1}(U_i)(1,\widetilde{\boldsymbol{X}}_i^{\mathrm{T}})^{\mathrm{T}}\epsilon_i + o_P(n^{-1/2}),
\end{aligned}$$

and similarly,

$$\begin{aligned}
&\hat{\beta}_{\mathrm{Vc},1} - \beta_{01} = \frac{\beta_{01}^{1-\rho}}{\rho}\left(\hat{\beta}_{\mathrm{Vc},1}^\rho - \beta_{01}^\rho\right) + o_P(n^{-1/2}) \qquad \text{(A.34)}\\
&= \frac{\beta_{01}}{\rho}\frac{1}{n}\sum_{i=1}^n \frac{|X_{1i}^\rho|}{E|X_1^\rho|}\left[\phi^\rho(U_i) - \psi_1^\rho(U_i)\right]\\
&\quad + \frac{1}{n}\sum_{i=1}^n \psi_1(U_i)\phi^\rho(U_i)e_2^{\mathrm{T}}G_{\widetilde{\boldsymbol{X}}}^{-1}(U_i)(1,\widetilde{\boldsymbol{X}}_i^{\mathrm{T}})^{\mathrm{T}}\epsilon_i + o_P(n^{-1/2})
\end{aligned}$$

Recalling that $G_{\widetilde{\boldsymbol{X}}}(u) = \breve{\boldsymbol{\psi}}(u)\boldsymbol{\Sigma}\breve{\boldsymbol{\psi}}(u)$, where $\breve{\boldsymbol{\psi}}(u) = \mathrm{diag}(1, \psi_1(u), \ldots, \psi_p(u))$, together with (A.33)-(A.34), we have $G^{-1}_{\widetilde{\boldsymbol{X}}}(u) = \breve{\boldsymbol{\psi}}^{-1}(u)\boldsymbol{\Sigma}^{-1}\breve{\boldsymbol{\psi}}^{-1}(u)$ and

$$\begin{aligned}
&\begin{pmatrix} \hat{\alpha}_{\mathrm{Vc}} - \alpha_0 \\ \hat{\boldsymbol{\beta}}_{\mathrm{Vc}} - \boldsymbol{\beta}_0 \end{pmatrix} \qquad \text{(A.35)} \\
&= \frac{1}{n}\sum_{i=1}^{n}[\phi^{\rho}(U_i) - 1]\mathrm{diag}\left(\frac{1}{\rho}, \frac{|X_{1i}^{\rho}|}{\rho E|X_1^{\rho}|}, \ldots, \frac{|X_{pi}^{\rho}|}{\rho E|X_p^{\rho}|}\right)\begin{pmatrix} \alpha_0 \\ \boldsymbol{\beta}_0 \end{pmatrix} \\
&\quad - \frac{1}{n}\sum_{i=1}^{n}\left(0, \frac{[\psi_1^{\rho}(U_i) - 1]|X_{1i}^{\rho}|}{\rho E|X_1^{\rho}|}, \ldots, \frac{[\psi_p^{\rho}(U_i) - 1]|X_{pi}^{\rho}|}{\rho E|X_p^{\rho}|}\right)\begin{pmatrix} \alpha_0 \\ \boldsymbol{\beta}_0 \end{pmatrix} \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\phi^{\rho}(U_i)\boldsymbol{\Sigma}^{-1}\begin{pmatrix} 1 \\ \boldsymbol{X}_i \end{pmatrix}\epsilon_i + o_P(n^{-1/2}).
\end{aligned}$$

The asymptotic normality of Theorem 4 is obtained from (A.35).

References

Mack, Y. P. and Silverman, B. W. (1982). Weak and strong uniform consistency of kernel regression estimates, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **61**(3): 405–415.

Owen, A. (1991). Empirical likelihood for linear models, *The Annals of Statistics* **19**: 1725–1747.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons Inc., New York.

Zhang, J., Gai, Y. and Wu, P. (2013). Estimation in linear regression models with measurement errors subject to single-indexed distortion, *Computational Statistics & Data Analysis* **59**: 103–120.

Zhang, J., Zhu, L. and Liang, H. (2012). Nonlinear models with measurement errors subject to single-indexed distortion, *Journal of Multivariate Analysis* **112**: 1–23.