Supplementary Material for "Logarithmic Calibration for Partial Linear Models with Multiplicative Distortion Measurement Errors" [1]

1. CONDITIONS

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

(C1) The distortion functions $\phi(u) > 0$ and $\psi_r(u) > 0$, $r = 1, \ldots, p$, 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 $\phi(u)$ and $\psi_r(u)$'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 $s \geq 4$, $E(|Y|^s) < \infty$, $E(|X_r|^s) < \infty$, $r = 1, \ldots, p$. The matrix $\mathbf{\Sigma}_0$ defined in Theorem 2 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 $K^{(j)}(\pm A) = 0$ with $K^{(j)}(t) = \frac{\mathrm{d}^j K(t)}{\mathrm{d}t^j}$, and $\mu_2 = \int_{-A}^{A} s^2 K(s) ds \neq 0$, $\mu_{K^2} = \int_{-A}^{A} K^2(s) ds > 0$.

(C4) As $n \to \infty$, the bandwidths h and h_1 satisfy $nh^4 \to 0$, $\frac{\log^2 n}{nh^2} \to 0$ and $nh_1^8 \to 0$ and $\frac{\log^2 n}{nh_1^2} \to 0$.

(C5) The density function of Z, $f_Z(z)$ is bounded away from zero on $\mathcal{Z}$, where $\mathcal{Z}$ is a compact support set in $\mathcal{R}^1$. Moreover, $f_Z(z)$, $E(X_s|Z = z)$, $E(Y|Z = z)$ and $g(z)$ have bounded continuous second order derivatives on $\mathcal{Z}$.

(C6) For all ζ_j $j = 1, \ldots, p$, $\zeta_j \to 0$, $\sqrt{n}\zeta_j \to \infty$ as $n \to \infty$,

$$\mathrm{liminf}_{n\to\infty}\mathrm{liminf}_{u\to 0^+} p'_{\zeta_j}(u)/\zeta_j > 0.$$

2. APPENDIX

2.1. A Technical Lemma

Lemma 1 Suppose $E(W|V = v) = w(v)$ and its derivatives up to second order are bounded for all $v \in [\mathcal{V}_L, \mathcal{V}_R]$, where $[\mathcal{V}_L, \mathcal{V}_R]$ denotes the compact support of V. $E|W|^3$ exists and $\sup_v \int |w|^s f(v, w) dw < \infty$ for some $s > 0$, where $f(v, w)$ is the joint density of $(V, W)^{\mathrm{T}}$. Suppose (V_i, W_i), $i = 1, 2, \ldots n$ are independent and identically distributed (i.i.d.) samples from (V, W). If condition (C3) holds true for kernel function $K(v)$, and $n^{2\epsilon-1} h \to \infty$ for $\epsilon < 1 - s^{-1}$, we have

$$\sup_{v\in[\mathcal{V}_L,\mathcal{V}_R]} \left| \frac{1}{n}\sum_{i=1}^{n} K_h(V_i - v)W_i - f_V(v)w(v) - \frac{1}{2}[f_V(v)w(v)]''\mu_2 h^2 \right| = O(\tau_{n,h}), a.s.$$

[1]

where, $f_V(v)$ is the density function of V, 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).

2.2. *Proof of Theorem 1*

Recalling that $\widetilde{Y}_i = \phi(U_i)Y_i = Y_i \exp(\ln(\phi(U_i)))$, we have

$$\begin{aligned}
\hat{Y}_i - Y_i &= \widetilde{Y}_i \exp\left(-\left\{\hat{m}_{\ln(|\widetilde{Y}|)}(U_i) - \overline{\ln(|\widetilde{Y}|)}\right\}\right) - Y_i \\
&= Y_i \left\{\exp\left(\ln(\phi(U_i)) - \left\{\hat{m}_{\ln(|\widetilde{Y}|)}(U_i) - \overline{\ln(|\widetilde{Y}|)}\right\}\right) - 1\right\}. \qquad \text{(A.1)}
\end{aligned}$$

Using Lemma 1, recalling the definition of $\hat{m}_{\ln(|\widetilde{Y}|)}(u)$, we have

$$\begin{aligned}
&\hat{m}_{\ln(|\widetilde{Y}|)}(u) - m_{\ln(|\widetilde{Y}|)}(u) \qquad \text{(A.2)} \\
&= \frac{1}{nf_U(u)} \sum_{j=1}^{n} K_h(U_j - u) \left\{\ln(|\widetilde{Y}_j|) - m_{\ln(|\widetilde{Y}|)}(U_j)\right\}, \\
&\quad + \frac{1}{nf_U(u)} \sum_{j=1}^{n} K_h(U_j - u) \left\{m_{\ln(|\widetilde{Y}|)}(U_j) - m_{\ln(|\widetilde{Y}|)}(u)\right\} \\
&\quad + O_P\left(h^2\sqrt{\frac{\log n}{nh}} + h^4 + \tau_{n,h}^2\right).
\end{aligned}$$

Using Lemma 1, we have

$$\hat{f}_U(u) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{U_i - u}{h}\right) = f_U(u) + \frac{\mu_2 h^2}{2} f_U''(u) + O_P(\tau_{n,h}), \qquad \text{(A.3)}$$

and

$$\begin{aligned}
&\frac{1}{nf_U(u)} \sum_{j=1}^{n} K_h(U_j - u) \left\{m_{\ln(|\widetilde{Y}|)}(U_j) - m_{\ln(|\widetilde{Y}|)}(u)\right\} \qquad \text{(A.4)} \\
&= \frac{h^2\mu_2}{2f_U(u)} \left\{\left[m_{\ln(|\widetilde{Y}|)}(u) f_U(u)\right]'' - m_{\ln(|\widetilde{Y}|)}(u) f_U''(u)\right\} + O_P(\tau_{n,h}).
\end{aligned}$$

Recalling that $m_{\ln(|\widetilde{Y}|)}(u) = \ln(\phi(u)) + E(\ln(|Y|))$, using (A.2), Taylor expansion entails that

$$\begin{aligned}
&\exp\left(\ln(\phi(U_i)) - \left\{\hat{m}_{\ln(|\widetilde{Y}|)}(U_i) - \overline{\ln(|\widetilde{Y}|)}\right\}\right) - 1 \qquad \text{(A.5)} \\
&= \left\{\overline{\ln(|\widetilde{Y}|)} - E(\ln(|Y|))\right\} - \frac{1}{nf_U(U_i)} \sum_{j=1}^{n} K_h(U_j - U_i) \left\{\ln(|\widetilde{Y}_j|) - m_{\ln(|\widetilde{Y}|)}(U_j)\right\} \\
&\quad - \frac{1}{nf_U(U_i)} \sum_{j=1}^{n} K_h(U_j - U_i) \left\{m_{\ln(|\widetilde{Y}|)}(U_j) - m_{\ln(|\widetilde{Y}|)}(U_i)\right\} \\
&\quad + O_P\left(h^2\sqrt{\frac{\log n}{nh}} + h^4 + \tau_{n,h}^2 + n^{-1}\right).
\end{aligned}$$

Let $M(\cdot)$ be a function of $\boldsymbol{W} = (Y, \boldsymbol{X})$, such that $E(M^2(\boldsymbol{W})) < \infty$. Using (A.1) and (A.5), as $h^2 \log n \to 0$, $nh^8 \to 0$ and $\frac{\log^2 n}{nh^2} \to 0$, we have

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)M(\boldsymbol{W}_i) \qquad\qquad (A.6)\\
&= \frac{1}{n}\sum_{i=1}^{n} Y_i M(\boldsymbol{W}_i)\left\{\exp\left(\ln(\phi(U_i)) - \left\{\hat{m}_{\ln(|\widetilde{Y}|)}(U_i) - \overline{\ln(|\widetilde{Y}|)}\right\}\right) - 1\right\}\\
&= \frac{1}{n}\sum_{i=1}^{n} Y_i M(\boldsymbol{W}_i)\left\{\overline{\ln(|\widetilde{Y}|)} - E(\ln(|Y|))\right\}\\
&\quad - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{Y_i M(\boldsymbol{W}_i)}{f_U(U_i)}K_h(U_j - U_i)\left\{\ln(|\widetilde{Y}_j|) - m_{\ln(|\widetilde{Y}|)}(U_j)\right\}\\
&\quad - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{Y_i M(\boldsymbol{W}_i)}{f_U(U_i)}K_h(U_j - U_i)\left\{m_{\ln(|\widetilde{Y}|)}(U_j) - m_{\ln(|\widetilde{Y}|)}(U_i)\right\} + o_P(n^{-1/2})\\
&= \mathcal{V}_{n,1} + \mathcal{V}_{n,2} + \mathcal{V}_{n,3}.
\end{aligned}$$

For the term $\mathcal{V}_{n,1}$, we have

$$\begin{aligned}
\mathcal{V}_{n,1} &= \frac{1}{n}\sum_{i=1}^{n} Y_i M(\boldsymbol{W}_i)\left\{\overline{\ln(|\widetilde{Y}|)} - E(\ln(|Y|))\right\} \qquad (A.7)\\
&= \frac{E[YM(\boldsymbol{W})]}{n}\sum_{i=1}^{n}\left\{\ln(|\widetilde{Y}_i|) - E(\ln(|Y|))\right\} + o_P(n^{-1/2}).
\end{aligned}$$

For $\mathcal{V}_{n,2}$, as $nh^4 \to 0$, the asymptotic expression of U-statistic (Serfling; 1980) entails that

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

For $\mathcal{V}_{n,3}$, as $nh^4 \to 0$, the asymptotic expression of U-statistic (Serfling; 1980) entails that $\mathcal{V}_{n,3} = O_P(h^2) = o_P(n^{-1/2})$. Together with (A.6)-(A.8), we have

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

Similarly, for $r = 1, \ldots, p$, we have

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}(\hat{X}_{ri} - X_{ri})M(\boldsymbol{W}_i) \qquad (A.10)\\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{\ln(|\widetilde{X}_{ri}|) - \ln(|X_{ri}|)\right\}E(X_r M(\boldsymbol{W})) + o_P(n^{-1/2})\\
&= \frac{1}{n}\sum_{i=1}^{n}\ln(\psi_r(U_i))E(X_r M(\boldsymbol{W})) + o_P(n^{-1/2}).
\end{aligned}$$

We complete the proof of Theorem 1.

2.3. Proof of Theorem 2

Recalling that

$$\begin{aligned}\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0 &= \left\{\frac{1}{n}\sum_{i=1}^{n}\left[\hat{\boldsymbol{X}}_i-\hat{S}_{\boldsymbol{X}}(Z_i)\right]^{\otimes 2}\right\}^{-1} \qquad \text{(B.1)}\\ &\quad \times\frac{1}{n}\sum_{i=1}^{n}\left\{\hat{\boldsymbol{X}}_i-\hat{S}_{\boldsymbol{X}}(Z_i)\right\}\left\{\hat{Y}_i-\hat{S}_Y(Z_i)-\hat{\boldsymbol{X}}_i^{\mathrm{T}}\boldsymbol{\beta}_0+\hat{S}_{\boldsymbol{X}}^{\mathrm{T}}(Z_i)\boldsymbol{\beta}_0\right\}\\ &= \left\{\frac{1}{n}\sum_{i=1}^{n}\left[\hat{\boldsymbol{X}}_i-\hat{S}_{\boldsymbol{X}}(Z_i)\right]^{\otimes 2}\right\}^{-1}\left[\mathbb{D}_{n1}+\mathbb{D}_{n2}+\mathbb{D}_{n3}\right],\end{aligned}$$

where

$$\mathbb{D}_{n1} = \frac{1}{n}\sum_{i=1}^{n}\left\{\hat{\boldsymbol{X}}_i-\hat{S}_{\boldsymbol{X}}(Z_i)\right\}\epsilon_i, \qquad \text{(B.2)}$$

$$\mathbb{D}_{n2} = \frac{1}{n}\sum_{i=1}^{n}\left\{\hat{\boldsymbol{X}}_i-\hat{S}_{\boldsymbol{X}}(Z_i)\right\}\left\{\hat{Y}_i-Y_i-(\hat{\boldsymbol{X}}_i-\boldsymbol{X}_i)^{\mathrm{T}}\boldsymbol{\beta}_0\right\} \qquad \text{(B.3)}$$

$$\begin{aligned}\mathbb{D}_{n3} &= \frac{1}{n}\sum_{i=1}^{n}\left\{\hat{\boldsymbol{X}}_i-\hat{S}_{\boldsymbol{X}}(Z_i)\right\} \qquad \text{(B.4)}\\ &\quad \times\left\{S_Y(Z_i)-\hat{S}_Y(Z_i)-(S_{\boldsymbol{X}}(Z_i)-\hat{S}_{\boldsymbol{X}}(Z_i))^{\mathrm{T}}\boldsymbol{\beta}_0\right\}.\end{aligned}$$

Step 2.1 For the expression $\mathbb{D}_{n1}$, we have

$$\begin{aligned}\mathbb{D}_{n1} &= \frac{1}{n}\sum_{i=1}^{n}\left\{\hat{\boldsymbol{X}}_i-\boldsymbol{X}_i\right\}\epsilon_i+\frac{1}{n}\sum_{i=1}^{n}\left\{\boldsymbol{X}_i-S_{\boldsymbol{X}}(Z_i)\right\}\epsilon_i \qquad \text{(B.5)}\\ &\quad \frac{1}{n}\sum_{i=1}^{n}\left\{S_{\boldsymbol{X}}(Z_i)-\hat{S}_{\boldsymbol{X}}(Z_i)\right\}\epsilon_i\\ &\stackrel{\text{def}}{=} \mathbb{D}_{n1}[1]+\mathbb{D}_{n1}[2]+\mathbb{D}_{n1}[3].\end{aligned}$$

Recalling $\epsilon_i = Y_i-\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta}_0-g(Z_i)$ and $E(\epsilon_i|\boldsymbol{X}_i,Z_i)=0$. Using the asymptotic results of Theorem 1, we have

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}\left\{\hat{X}_{ri}-X_{ri}\right\}\epsilon_i \qquad \text{(B.6)}\\ &= \frac{1}{n}\sum_{i=1}^{n}\ln(\psi_r(U_i))E\{X_r[Y-\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\beta}_0-g(Z)]\}+o_P(n^{-1/2})=o_P(n^{-1/2}).\end{aligned}$$

Based on (B.6), we have $\mathbb{D}_{n1}[1]=o_P(n^{-1/2})$.

Step 2.2 In the following, we define

$$M^{\Delta}_{n\delta,\hat{W}}(z)=\frac{1}{nh_1}\sum_{i=1}^{n}\left(\frac{Z_i-z}{h_1}\right)^{\delta}K\left(\frac{Z_i-z}{h_1}\right)(\hat{W}_i-W_i), \qquad \text{(B.7)}$$

where, $\hat{W}_i=\hat{Y}_i$, $W_i=Y_i$ and $\hat{W}_i=\hat{X}_{ri}$, $W_i=X_{ri}$ for $\delta=0,1$, $r=1,\ldots,p$ and $i=1,\ldots,n$.

For $\delta = 0$, similar to (A.1) and (A.5), we have

$$\begin{aligned}
M^{\Delta}_{n0,\hat{X}_r}(z) &= \frac{1}{nh_1}\sum_{i=1}^{n} K\left(\frac{Z_i - z}{h_1}\right)(\hat{X}_{ri} - X_{ri}) \qquad \text{(B.8)}\\
&= \frac{1}{nh_1}\sum_{i=1}^{n} K\left(\frac{Z_i - z}{h_1}\right) X_{ri}\left\{\overline{\ln(|\widetilde{X}_r|)} - E(\ln(|X_r|))\right\}\\
&\quad -\frac{1}{n^2h_1h}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{X_{ri}}{f_U(U_i)}K\left(\frac{Z_i - z}{h_1}\right)K\left(\frac{U_j - U_i}{h}\right)\\
&\qquad\qquad \times\left\{\ln(|\widetilde{X}_{rj}|) - m_{\ln(|\widetilde{X}_r|)}(U_j)\right\}\\
&\quad -\frac{1}{n^2h_1h}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{X_{ri}}{f_U(U_i)}K\left(\frac{Z_i - z}{h_1}\right)K\left(\frac{U_j - U_i}{h}\right)\\
&\qquad\qquad \times\left\{m_{\ln(|\widetilde{Y}|)}(U_j) - m_{\ln(|\widetilde{Y}|)}(U_i)\right\}\\
&\quad +O_P\left(h^2\sqrt{\frac{\log n}{nh}} + h^4 + \tau^2_{n,h} + n^{-1}\right).
\end{aligned}$$

Recalling that $m_{\ln(|\widetilde{X}_r|)}(u) = \ln(\psi_r(u)) + E(\ln(|X_r|))$, the asymptotic expression of U-statistic (Serfling; 1980) entails that

$$\begin{aligned}
&\frac{1}{n^2h_1h}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{X_{ri}}{f_U(U_i)}K\left(\frac{Z_i - z}{h_1}\right)K\left(\frac{U_j - U_i}{h}\right) \qquad \text{(B.9)}\\
&\qquad \times\left\{\ln(|\widetilde{X}_{rj}|) - m_{\ln(|\widetilde{X}_r|)}(U_j)\right\}\\
&= s_{X_r}(z)f_Z(z)\frac{1}{n}\sum_{i=1}^{n}\{\ln(|X_{ri}|) - E(\ln(|X_r|))\} + o_P(n^{-1/2}) + O_P(n^{-1/2}h_1^2).
\end{aligned}$$

Similar to (B.8), we have

$$\begin{aligned}
&\frac{1}{n^2h_1h}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{X_{ri}}{f_U(U_i)}K\left(\frac{Z_i - z}{h_1}\right)K\left(\frac{U_j - U_i}{h}\right) \qquad \text{(B.10)}\\
&\qquad \times\left\{m_{\ln(|\widetilde{Y}|)}(U_j) - m_{\ln(|\widetilde{Y}|)}(U_i)\right\} = O_P(h^2 + h_1^2h^2)
\end{aligned}$$

Together with (B.8)-(B.10), as $nh^4 \to 0$, $\frac{\log n}{nh_1} \to 0$, we have

$$\begin{aligned}
M^{\Delta}_{n0,\hat{X}_r}(z) &= \frac{1}{nh_1}\sum_{i=1}^{n} K\left(\frac{Z_i - z}{h_1}\right) X_{ri}\left\{\overline{\ln(|\widetilde{X}_r|)} - E(\ln(|X_r|))\right\} \qquad \text{(B.11)}\\
&\quad -s_{X_r}(z)f_Z(z)\frac{1}{n}\sum_{i=1}^{n}\{\ln(|X_{ri}|) - E(\ln(|X_r|))\}\\
&\quad +O_P\left(h^2 + h_1^2h^2 + h^2\sqrt{\frac{\log n}{nh}} + h^4 + \tau^2_{n,h} + n^{-1}\right)\\
&= s_{X_r}(z)f_Z(z)\frac{1}{n}\sum_{i=1}^{n}\ln(\psi_r(U_i)) + o_P(n^{-1/2}).
\end{aligned}$$

Similar to (B.11), we have

$$M^{\Delta}_{n1,\hat{X}_r}(z) = h_1[s_{X_r}(z)f_Z(z)]'\frac{1}{n}\sum_{i=1}^{n}\ln(\psi_r(U_i)) + o_P(n^{-1/2}). \qquad \text{(B.12)}$$

Thus, using Lemma 1 and (B.12), we have

$$\begin{aligned}\hat{s}_{X_r}(z) &= \frac{Q_{n2}(z)M_{n0,\hat{X}_r}(z) - Q_{n1}(z)M_{n1,X_r}(z)}{Q_{n2}(z)Q_{n0}(z) - [Q_{n1}(z)]^2} \\ &\quad + \frac{Q_{n2}(z)M^{\Delta}_{n0,\hat{X}_r}(z) - Q_{n1}(z)M^{\Delta}_{n1,\hat{X}_r}(z)}{Q_{n2}(z)Q_{n0}(z) - [Q_{n1}(z)]^2} \\ &\stackrel{\text{def}}{=} \hat{s}^*_{X_r}(z) + s_{X_r}(z)\frac{1}{n}\sum_{i=1}^{n}\ln(\psi_r(U_i)) + o_P(n^{-1/2}).\end{aligned} \tag{B.13}$$

Directly using Lemma A.1 in Liang and Li (2009) and similar to the proof of Theorem 1 in Liang and Li (2009), we have

$$\frac{1}{n}\sum_{i=1}^{n}\left\{S_{\boldsymbol{X}}(Z_i) - \hat{S}^*_{\boldsymbol{X}}(Z_i)\right\}\epsilon_i = o_P(n^{-1/2}) \tag{B.14}$$

where, $\hat{S}^*_{\boldsymbol{X}}(Z_i) = (\hat{s}^*_{X_1}(Z_i), \ldots, \hat{s}^*_{X_r}(Z_i))^{\mathrm{T}}$.

Appealing to (B.13)-(B.14), we obtain

$$\begin{aligned}&\mathbb{D}_{n1}[3] \\ &= \frac{1}{n}\sum_{i=1}^{n}\left\{S_{\boldsymbol{X}}(Z_i) - \hat{S}^*_{\boldsymbol{X}}(Z_i)\right\}\epsilon_i - \frac{1}{n}\sum_{i=1}^{n}S_{\boldsymbol{X}}(Z_i)\epsilon_i\left\{\frac{1}{n}\sum_{i=1}^{n}\ln(\psi_r(U_i))\right\} \\ &\quad + o_P(n^{-1/2}) = O_P(n^{-1}) + o_P(n^{-1/2}) = o_P(n^{-1/2}).\end{aligned} \tag{B.15}$$

Thus, according to (B.5)-(B.6) and (B.15), we obtain that

$$\mathbb{D}_{n1} = \frac{1}{n}\sum_{i=1}^{n}\left\{\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z_i)\right\}\epsilon_i + o_P(n^{-1/2}). \tag{B.16}$$

Step 2.3 For the argument $\mathbb{D}_{n2}$, we have

$$\begin{aligned}\mathbb{D}_{n2} &= \frac{1}{n}\sum_{i=1}^{n}\left\{\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i\right\}\left\{\hat{Y}_i - Y_i - (\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i)^{\mathrm{T}}\boldsymbol{\beta}_0\right\} \\ &\quad + \frac{1}{n}\sum_{i=1}^{n}\left\{\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z_i)\right\}\left\{\hat{Y}_i - Y_i - (\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i)^{\mathrm{T}}\boldsymbol{\beta}_0\right\} \\ &\quad + \frac{1}{n}\sum_{i=1}^{n}\left\{S_{\boldsymbol{X}}(Z_i) - \hat{S}_{\boldsymbol{X}}(Z_i)\right\}\left\{\hat{Y}_i - Y_i - (\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i)^{\mathrm{T}}\boldsymbol{\beta}_0\right\} \\ &\stackrel{\text{def}}{=} \mathbb{D}_{n2}[1] + \mathbb{D}_{n2}[2] + \mathbb{D}_{n2}[3].\end{aligned} \tag{B.17}$$

Let $\hat{V}_i = \hat{Y}_i$, or $\hat{V}_i = \hat{X}_{ri}$, and $\hat{D}_i = \hat{Y}_i$, or $\hat{D}_i = \hat{X}_{ri}$, accordingly, $V_i = Y_i$, or $V_i = X_{ri}$ or $V_i = Z_i$, and $D_i = Y_i$, or $D_i = X_{ri}$ or $D_i = Z_i$. Based on (A.5), as $nh^8 \to 0$ and $\frac{\log^2 n}{nh^2} \to 0$, we have

$$\frac{1}{n}\sum_{i=1}^{n}(\hat{V}_i - V_i)(\hat{D}_i - D_i) = O_P((n^{-1/2} + h^2 + \tau_{n,h})^2) = o_P(n^{-1/2}). \tag{B.18}$$

Using (B.18), we have $\mathbb{D}_{n2}[1] = o_P(n^{-1/2})$. For $\mathbb{D}_{n2}[2]$, using $E[\boldsymbol{X} - S_{\boldsymbol{X}}(Z)|Z] = 0$, and

$\mathrm{Cov}(Y, \boldsymbol{X} - S_{\boldsymbol{X}}(Z)) = \boldsymbol{\Sigma}_0\boldsymbol{\beta}_0$, Theorem 1 entails that

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\{\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z_i)\}\left\{\hat{Y}_i - Y_i\right\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\ln(\phi(U_i))E[Y(\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z))] + o_P(n^{-1/2}) \\
&= \frac{1}{n}\sum_{i=1}^{n}\ln(\phi(U_i))\mathrm{Cov}(Y, \boldsymbol{X} - S_{\boldsymbol{X}}(Z)) + o_P(n^{-1/2}) \\
&= \frac{1}{n}\sum_{i=1}^{n}\ln(\phi(U_i))\boldsymbol{\Sigma}_0\boldsymbol{\beta}_0 + o_P(n^{-1/2}).
\end{aligned} \tag{B.19}
$$

Similarly, we have

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\{\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z_i)\}(\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i)^{\mathrm{T}}\boldsymbol{\beta}_0 \\
&= \sum_{r=1}^{p}\left\{\frac{1}{n}\sum_{i=1}^{n}\{\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z_i)\}(\hat{X}_{ri} - X_{ri})\beta_{0r}\right\} \\
&= \frac{1}{n}\sum_{r=1}^{p}\sum_{i=1}^{n}\ln(\psi_r(U_i))E[X_r(\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z))]\beta_{0r} + o_P(n^{-1/2}) \\
&= \frac{1}{n}\sum_{r=1}^{p}\sum_{i=1}^{n}\ln(\psi_r(U_i))E[(\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z))^{\otimes 2}]\boldsymbol{e}_r\boldsymbol{e}_r^{\mathrm{T}}\boldsymbol{\beta}_0 + o_P(n^{-1/2}). \\
&= \frac{1}{n}\sum_{r=1}^{p}\sum_{i=1}^{n}\ln(\psi_r(U_i))\boldsymbol{\Sigma}_0\boldsymbol{e}_r\boldsymbol{e}_r^{\mathrm{T}}\boldsymbol{\beta}_0 + o_P(n^{-1/2}).
\end{aligned} \tag{B.20}
$$

Together with (B.19) and (B.20), we have

$$
\begin{aligned}
\mathbb{D}_{n2}[2] \quad = \quad & \frac{1}{n}\sum_{i=1}^{n}\ln(\phi(U_i))\boldsymbol{\Sigma}_0\boldsymbol{\beta}_0 \\
& -\frac{1}{n}\sum_{r=1}^{p}\sum_{i=1}^{n}\ln(\psi_r(U_i))\boldsymbol{\Sigma}_0\boldsymbol{e}_r\boldsymbol{e}_r^{\mathrm{T}}\boldsymbol{\beta}_0 + o_P(n^{-1/2}).
\end{aligned} \tag{B.21}
$$

Under the condition $nh_1^8 \to 0$ and $\frac{\log n}{nh_1^2} \to 0$, the conclusion of (A.1) in Liang and Li (2009) entails that $\sup_{z\in\mathcal{Z}}|\hat{s}^*_{X_r}(z) - s_{X_r}(z)| = o_P(n^{-1/4})$, $r = 1, \ldots, p$.

According to the proof of Theorem 1 in Zhang et al. (2016), using (B.13), we have

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\{s_{X_l}(Z_i) - \hat{s}_{X_l}(Z_i)\}(\hat{X}_{ri} - X_{ri}) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{s_{X_l}(Z_i) - \hat{s}^*_{X_l}(Z_i) - s_{X_l}(Z_i)\frac{1}{n}\sum_{i=1}^{n}\psi_l(U_i)\right\}(\hat{X}_{ri} - X_{ri}) + o_P(n^{-1/2}) \\
&= O_P(n^{-1/2}h_1^2 + n^{-1}) + o_P(n^{-1/2}) = o_P(n^{-1/2}).
\end{aligned} \tag{B.22}
$$

Similar to (B.22), $\mathbb{D}_{n2}[3] = o_P(n^{-1/2})$, and also $\mathbb{D}_{n3} = o_P(n^{-1/2})$. Moreover,

$$
\frac{1}{n}\sum_{i=1}^{n}\left[\hat{\boldsymbol{X}}_i - \hat{S}_{\boldsymbol{X}}(Z_i)\right]^{\otimes 2} \xrightarrow{P} \boldsymbol{\Sigma}_0. \tag{B.23}
$$

Thus, together with (B.16), (B.21) and (B.23), we obtain that

$$\begin{aligned}\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0 &= \boldsymbol{\Sigma}_0^{-1}\left(\mathbb{D}_{n1}+\mathbb{D}_{n2}+\mathbb{D}_{n3}\right)+o_P(n^{-1/2}) \\ &= \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\Sigma}_0^{-1}\left\{\boldsymbol{X}_i-S_{\boldsymbol{X}}(Z_i)\right\}\epsilon_i \\ &\quad +\frac{1}{n}\sum_{i=1}^{n}\left\{\ln(\phi(U_i))-\sum_{r=1}^{p}\ln(\psi_r(U_i))\boldsymbol{e}_r\boldsymbol{e}_r^{\mathrm{T}}\right\}\boldsymbol{\beta}_0+o_P(n^{-1/2}).\end{aligned} \tag{B.24}$$

We have completed the proof of Theorem 2.

2.4. Proof of Theorem 3

Note that

$$\begin{aligned}\hat{g}(z)-g(z) &= \frac{T_{n2}(z)\hat{V}_{n0}(z)-T_{n1}(z)\hat{V}_{n1}(z)}{T_{n2}(z)T_{n0}(z)-\left[T_{n1}(z)\right]^2}-g(z) \\ &= \frac{T_{n2}(z)[\hat{V}_{n0}(z)-T_{n0}(z)g(z)]}{T_{n2}(z)T_{n0}(z)-\left[T_{n1}(z)\right]^2}-\frac{T_{n1}(z)[\hat{V}_{n1}(z)-T_{n1}(z)g(z)]}{T_{n2}(z)T_{n0}(z)-\left[T_{n1}(z)\right]^2} \\ &= S_{n1}(z)-S_{n2}(z)\end{aligned} \tag{C.1}$$

For the term $S_{n1}(z)$, we have

$$\begin{aligned} S_{n1}(z) &= \frac{1}{T_{n0}(z)-\left[T_{n1}(z)\right]^2/T_{n2}(z)}\frac{1}{nh_2}\sum_{i=1}^{n}K\left(\frac{Z_i-z}{h_2}\right)\epsilon_i \\ &+\frac{1}{T_{n0}(z)-\left[T_{n1}(z)\right]^2/T_{n2}(z)}\frac{1}{nh_2}\sum_{i=1}^{n}K\left(\frac{Z_i-z}{h_2}\right)(g(Z_i)-g(z)) \\ &+\frac{1}{T_{n0}(z)-\left[T_{n1}(z)\right]^2/T_{n2}(z)}\frac{1}{nh_2}\sum_{i=1}^{n}K\left(\frac{Z_i-z}{h_2}\right)\boldsymbol{X}_i^{\mathrm{T}}\left(\boldsymbol{\beta}_0-\hat{\boldsymbol{\beta}}\right) \\ &+\frac{1}{T_{n0}(z)-\left[T_{n1}(z)\right]^2/T_{n2}(z)}\frac{1}{nh_2}\sum_{i=1}^{n}K\left(\frac{Z_i-z}{h_2}\right)\left(\hat{Y}_i-Y_i-(\hat{\boldsymbol{X}}_i-\boldsymbol{X}_i)^{\mathrm{T}}\hat{\boldsymbol{\beta}}\right) \\ &\stackrel{\text{def}}{=} S_{n1,[1]}(z)+S_{n1,[2]}(z)+S_{n1,[3]}(z)+S_{n1,[4]}(z).\end{aligned} \tag{C.2}$$

Directly using Lemma 1, we have

$$S_{n1,[1]}(z) = \frac{1}{nh_2f_Z(z)}\sum_{i=1}^{n}K\left(\frac{Z_i-z}{h_2}\right)\epsilon_i+O_P\left(h_2^2\sqrt{\frac{\log n}{nh_2}}+\frac{\log n}{nh_2}\right), \tag{C.3}$$

$$S_{n1,[2]}(z) = \frac{h_2^2\mu_2}{2}g''(z)+h_2^2\mu_2\frac{g'(z)f_Z'(z)}{f_Z(z)}+O_P\left(h_2^4+\frac{\log n}{nh_2}\right). \tag{C.4}$$

By using Theorem 2, we obtain that $\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0=O_P(n^{-1/2})$, and we can have that

$$S_{n1,[3]}(z)=O_P(n^{-1/2})=o_P((nh_2)^{-1/2}). \tag{C.5}$$

Using (A.1) and (A.5), similar to (B.8), we have

$$\begin{aligned}S_{n1,[4]}(z) &= s_Y(z)\frac{1}{n}\sum_{i=1}^{n}\ln(\phi(U_i)) \\ &\quad -\frac{1}{n}\sum_{i=1}^{n}\sum_{r=1}^{p}s_{X_r}(z)\beta_{0r}\ln(\psi_r(U_i))+o_P(n^{-1/2}) \\ &= O_P(n^{-1/2})=o_P((nh_2)^{-1/2}).\end{aligned} \tag{C.6}$$

Similar to the analysis of (C.2)-(C.6), we have

$$S_{n2}(z) = \frac{g'(z)f_Z'(z)}{f_Z(z)}h_2^2\mu_2 + o_P(h_2^2 + 1/\sqrt{nh_2}). \tag{C.7}$$

Together with (C.2) and (C.7), we have

$$\begin{aligned}
&\hat{g}(z) - g(z) - \frac{\mu_2 h_2^2}{2}g''(z) \\
&= \frac{1}{f_Z(z)nh_2}\sum_{i=1}^{n} K\left(\frac{Z_i - z}{h_2}\right)\epsilon_i + o_P(h_2^2 + 1/\sqrt{nh_2}).
\end{aligned} \tag{C.8}$$

The asymptotic result of Theorem is directly obtained from (C.8), we have completed the proof of Theorem 3.

2.5. Proof of Theorem 4

We first consider the conditional mean calibration. For $1 \le r \le p$, let $\hat{\wp}_{n,i}^{[r]}(\boldsymbol{\beta}_0)$ be the r-component of $\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)$. We decompose $\hat{\wp}_{n,i}^{[r]}(\boldsymbol{\beta}_0)$ into following terms:

$$\hat{\wp}_{n,i}^{[r]}(\boldsymbol{\beta}_0) = (Y_i - S_Y(Z_i) - [\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z_i)]^{\mathrm{T}}\boldsymbol{\beta}_0)[X_{ri} - s_{X_r}(Z_i)] + \sum_{t=1}^{8} R_{n,it}^{[r]},$$

where,

$$\begin{aligned}
R_{n,i1}^{[r]} &= \{\hat{Y}_i - Y_i - [\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i]^{\mathrm{T}}\boldsymbol{\beta}_0\}[X_{ri} - s_{X_r}(Z_i)], \\
R_{n,i2}^{[r]} &= \{\hat{Y}_i - Y_i - [\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i]^{\mathrm{T}}\boldsymbol{\beta}_0\}[\hat{X}_{ri} - X_{ri}], \\
R_{n,i3}^{[r]} &= \{\hat{Y}_i - Y_i - [\hat{\boldsymbol{X}}_i - \boldsymbol{X}_i]^{\mathrm{T}}\boldsymbol{\beta}_0\}[s_{X_r}(Z_i) - \hat{s}_{X_r}(Z_i)], \\
R_{n,i4}^{[r]} &= \{Y_i - S_Y(Z_i) - [\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z_i)]^{\mathrm{T}}\boldsymbol{\beta}_0\}[s_{X_r}(Z_i) - \hat{s}_{X_r}(Z_i)], \\
R_{n,i5}^{[r]} &= \{Y_i - S_Y(Z_i) - [\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z_i)]^{\mathrm{T}}\boldsymbol{\beta}_0\}[\hat{X}_{ri} - X_{ri}], \\
R_{n,i6}^{[r]} &= \{S_Y(Z_i) - \hat{S}_Y(Z_i) - [S_{\boldsymbol{X}}(Z_i) - \hat{S}_{\boldsymbol{X}}(Z_i)]^{\mathrm{T}}\boldsymbol{\beta}_0\}[s_{X_r}(Z_i) - \hat{s}_{X_r}(Z_i)], \\
R_{n,i7}^{[r]} &= \{S_Y(Z_i) - \hat{S}_Y(Z_i) - [S_{\boldsymbol{X}}(Z_i) - \hat{S}_{\boldsymbol{X}}(Z_i)]^{\mathrm{T}}\boldsymbol{\beta}_0\}[\hat{X}_{ri} - X_{ri}], \\
R_{n,i8}^{[r]} &= \{S_Y(Z_i) - \hat{S}_Y(Z_i) - [S_{\boldsymbol{X}}(Z_i) - \hat{S}_{\boldsymbol{X}}(Z_i)]^{\mathrm{T}}\boldsymbol{\beta}_0\}[X_{ri} - s_{X_r}(Z_i)].
\end{aligned}$$

To prove Theorem 4, we need to show that

$$\max_{1\le i\le n} |\hat{\wp}_{n,it}^{[r]}| = o_P(n^{1/2}), \quad t = 1, \ldots, 8.$$

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

$$\max_{1\le i\le n} \left|(Y_i - S_Y(Z_i) - [\boldsymbol{X}_i - S_{\boldsymbol{X}}(Z_i)]^{\mathrm{T}}\boldsymbol{\beta}_0)[X_{ri} - s_{X_r}(Z_i)]\right| = o_P(n^{1/2}).$$

Next, for $R^{[r]}_{n,i1}$, according to (A.1) and (A.5),

$$\begin{aligned}
&\max_{1\le i\le n}|\{\hat{Y}_i-Y_i\}[X_{ri}-s_{X_r}(Z_i)]| \qquad (D.1)\\
&\le \max_{1\le i\le n}\left|\left\{\exp\left(\ln(\phi(U_i))-\left\{\hat{m}_{\ln(|\widetilde{Y}|)}(U_i)-\overline{\ln(|\widetilde{Y}|)}\right\}\right)-1\right\}\right|\\
&\qquad\times|Y_i[X_{ri}-s_{X_r}(Z_i)]|\\
&\le \max_{1\le i\le n}|Y_i[X_{ri}-s_{X_r}(Z_i)]|\left|\overline{\ln(|\widetilde{Y}|)}-E(\ln(|Y|))\right|\\
&\quad+\max_{1\le i\le n}|Y_i[X_{ri}-s_{X_r}(Z_i)]|\,O_P\left(h^2+\sqrt{\frac{\log n}{nh}}\right)+O_P\left(h^4+\frac{\log n}{nh}\right)O_P(n^{1/2})\\
&=o_P(n^{1/2}).
\end{aligned}$$

Similar to (D.1), we have

$$\max_{1\le i\le n}|R^{[r]}_{n,i1}|=o_P(n^{1/2}),\quad \max_{1\le i\le n}|R^{[r]}_{n,i5}|=o_P(n^{1/2}). \qquad (D.2)$$

For $R^{[r]}_{n,i2}$, similar to (D.1), we have

$$\begin{aligned}
&\max_{1\le i\le n}|\{\hat{Y}_i-Y_i\}[\hat{X}_{ri}-X_{ri}]| \qquad (D.3)\\
&\le \max_{1\le i\le n}|Y_iX_{ri}|\max_{1\le i\le n}\left|\left\{\exp\left(\ln(\phi(U_i))-\left\{\hat{m}_{\ln(|\widetilde{Y}|)}(U_i)-\overline{\ln(|\widetilde{Y}|)}\right\}\right)-1\right\}\right|\\
&\quad\times\max_{1\le i\le n}\left|\left\{\exp\left(\ln(\psi_r(U_i))-\left\{\hat{m}_{\ln(|\widetilde{X}_r|)}(U_i)-\overline{\ln(|\widetilde{X}_r|)}\right\}\right)-1\right\}\right|\\
&=O_P\left(h^4+\frac{\log n}{nh}+n^{-1}\right)O_P(n^{1/2})=o_P(n^{1/2})
\end{aligned}$$

Thus, according to (D.3), we show that

$$\max_{1\le i\le n}|R^{[r]}_{n,i2}|=o_P(n^{1/2}). \qquad (D.4)$$

The conclusion of (A.1) in Liang and Li (2009) entails that $\sup_{z\in\mathcal{Z}}|\hat{S}^*_Y(z)-S_Y(z)|=o_P(n^{-1/4})$, and $\sup_{z\in\mathcal{Z}}|\hat{s}^*_{X_r}(z)-s_{X_r}(z)|=o_P(n^{-1/4})$, $r=1,\dots,p$. Similar to (B.22), we have

$$\begin{aligned}
&\max_{1\le i\le n}|\{\hat{Y}_i-Y_i\}[s_{X_r}(Z_i)-\hat{s}_{X_r}(Z_i)]| \qquad (D.5)\\
&\le \max_{1\le i\le n}|Y_i|\max_{1\le i\le n}|s_{X_r}(Z_i)-\hat{s}_{X_r}(Z_i)|\\
&\quad\times\max_{1\le i\le n}\left|\left\{\exp\left(\ln(\phi(U_i))-\left\{\hat{m}_{\ln(|\widetilde{Y}|)}(U_i)-\overline{\ln(|\widetilde{Y}|)}\right\}\right)-1\right\}\right|=o_P(n^{1/2}).
\end{aligned}$$

Similar to (D.5), we show that

$$\max_{1\le i\le n}|R^{[r]}_{n,i3}|=o_P(n^{1/2}). \qquad (D.6)$$

Similar to the proofs of $|R^{[r]}_{n,it}|$, $t=1,2,3,5$, we have $\max_{1\le i\le n}|R^{[r]}_{n,it}|=o_P(n^{1/2})$ for $t=4,6,7,8$. We omit the details. Followed the same argument in the proof (2.14) in Owen (1991), we have $\hat{\lambda}=O_P(n^{1/2})$. Thus, $\max_{1\le i\le n}|\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)|=o_P(1)$. Note that $\log(1+t)\approx t-\frac{1}{2}t^2$ for t sufficiently small, we have

$$\hat{l}(\boldsymbol{\beta}_0)=2\sum_{i=1}^n\left(\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)-\frac{1}{2}\{\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\}^2\right)+o_P(1). \qquad (D.7)$$

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

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

Further,

$$\begin{aligned}\mathbf{0} &= \frac{1}{n}\sum_{i=1}^{n}\frac{\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)}{1+\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)}\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)-\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)^{\mathrm{T}}\hat{\lambda}\\ &\quad+\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\{\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\}^2}{1+\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)}. \end{aligned}\tag{D.8}$$

Above equation (D.8) and $\max\limits_{1\le i\le n}|\hat{\lambda}^{\mathrm{T}}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)|=o_P(1)$ entail that

$$\hat{\lambda}=\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)^{\mathrm{T}}\Big)^{-1}\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)+o_P(n^{-1/2}). \tag{D.9}$$

Plugging the asymptotic expressions (D.7)-(D.9), we have

$$\begin{aligned}&\hat{l}(\boldsymbol{\beta}_0)\\ &=n\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\Big)^{\mathrm{T}}\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)^{\mathrm{T}}\Big)^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n}\hat{\wp}_{n,i}(\boldsymbol{\beta}_0)\Big)\\ &\quad+o_P(1).\end{aligned}\tag{D.10}$$

According the proof Theorem 2, we can 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 $\kappa_{n,i}(\boldsymbol{\beta}_0)=\{Y_i-S_Y(Z_i)-[\boldsymbol{X}_i-S_{\boldsymbol{X}}(Z_i)]^{\mathrm{T}}\boldsymbol{\beta}_0\}[\boldsymbol{X}_i-S_{\boldsymbol{X}}(Z_i)]$ is independent and identically distributed p-dimensional random vector with zero mean. Theorem 4 for $\hat{l}(\boldsymbol{\beta}_0)$ follows from the central limit theorem and the Slutsky theorem.

3. Proof of Theorem 5 and Theorem 6

Step 1 Note that

$$\hat{\boldsymbol{\beta}}_{\mathrm{R}}=\hat{\boldsymbol{\beta}}-\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\left\{\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\right\}^{-1}\left[\boldsymbol{A}\hat{\boldsymbol{\beta}}-\boldsymbol{b}\right]. \tag{E.1}$$

Under the null hypothesis $\mathcal{H}_0$, we have $\boldsymbol{A}\boldsymbol{\beta}_0=\boldsymbol{b}$. Using (E.1), it is seen that

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{\mathrm{R}}-\boldsymbol{\beta}_0 &= \left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0\right)-\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\left\{\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\right\}^{-1}\left[\boldsymbol{A}\hat{\boldsymbol{\beta}}-\boldsymbol{A}\boldsymbol{\beta}_0\right]\\ &= \left[\boldsymbol{I}_p-\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\left\{\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\right\}^{-1}\boldsymbol{A}\right]\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0\right).\end{aligned}\tag{E.2}$$

Together with (B.23) and (B.24), the equation (E.2) can be expressed as

$$\hat{\boldsymbol{\beta}}_{\mathrm{R}}-\boldsymbol{\beta}_0=\left[\boldsymbol{I}_p-\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}}\left\{\boldsymbol{A}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}}\right\}^{-1}\boldsymbol{A}\right]\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0\right)+o_P(n^{-1/2}). \tag{E.3}$$

Define $\boldsymbol{\Omega}_{\boldsymbol{A}}=\boldsymbol{I}_p-\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}}\left\{\boldsymbol{A}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}}\right\}^{-1}\boldsymbol{A}$, the expression (E.3) entails that

$$\sqrt{n}\left(\hat{\boldsymbol{\beta}}_{\mathrm{R}}-\boldsymbol{\beta}_0\right)\xrightarrow{L}N(\mathbf{0},\boldsymbol{\Omega}_{\boldsymbol{A}}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{0\epsilon}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Omega}_{\boldsymbol{A}}^{\mathrm{T}}+\boldsymbol{\Omega}_{\boldsymbol{A}}\boldsymbol{\Sigma}_{\phi,\boldsymbol{\psi}}\boldsymbol{\Omega}_{\boldsymbol{A}}^{\mathrm{T}}).$$

We have completed the proof of Theorem 5.

Step 2 Under the null hypothesis $\mathcal{H}_0 : \boldsymbol{A}\boldsymbol{\beta}_0 = \boldsymbol{b}$, using (B.24) and Theorem 1, we have

$$\begin{aligned} \sqrt{n}\left(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}\right) &= \sqrt{n}\boldsymbol{A}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right) \\ &\xrightarrow{L} N\left(\mathbf{0}, \boldsymbol{A}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{0\epsilon}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{A}\boldsymbol{\Sigma}_{\phi,\boldsymbol{\psi}}\boldsymbol{A}^{\mathrm{T}}\right). \end{aligned} \tag{E.4}$$

Similar to the analysis of (B.23), we have

$$\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}_\epsilon\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{A}\hat{\boldsymbol{\Sigma}}_{\phi,\boldsymbol{\psi}}\boldsymbol{A}^{\mathrm{T}} \xrightarrow{P} \boldsymbol{A}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{0\epsilon}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{A}\boldsymbol{\Sigma}_{\phi,\boldsymbol{\psi}}\boldsymbol{A}^{\mathrm{T}}. \tag{E.5}$$

The Slutsky theorem entails that

$$\begin{aligned} &\left[\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}_\epsilon\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{A}\hat{\boldsymbol{\Sigma}}_{\phi,\boldsymbol{\psi}}\boldsymbol{A}^{\mathrm{T}}\right]^{-1/2}\left[\sqrt{n}\left(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}\right)\right] \\ &\xrightarrow{L} N(\mathbf{0}, \boldsymbol{I}_k), \end{aligned} \tag{E.6}$$

where $\boldsymbol{I}_k$ is a $k \times k$ dimensional identity matrix. Using (E.6), the continuous mapping theorem entails that

$$\begin{aligned} \mathcal{T}_n &= n\left(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}\right)^{\mathrm{T}}\left[\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}_\epsilon\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{A}\hat{\boldsymbol{\Sigma}}_{\phi,\boldsymbol{\psi}}\boldsymbol{A}^{\mathrm{T}}\right]^{-1}\left(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}\right) \\ &\xrightarrow{L} \chi_k^2, \end{aligned} \tag{E.7}$$

where χ_k^2 is the centered chi-squared distribution with degree of freedom k We have completed the proof of Theorem 6.

4. Proof of Theorem 7

Step 1 It is noted that $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\beta}_0 - n^{-1/2}\boldsymbol{c}$ under the null hypothesis $\mathcal{H}_{1n}$, from (E.1) and we have

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{\mathrm{R}} &= \hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\left\{\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\right\}^{-1}\left[\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}\right] \\ &= \hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\left\{\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\right\}^{-1}\left[\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{A}\boldsymbol{\beta}_0 + n^{-1/2}\boldsymbol{c}\right] \\ &= \hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\left\{\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\right\}^{-1}\boldsymbol{A}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right) \\ &\quad - n^{-1/2}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\left\{\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}\right\}^{-1}\boldsymbol{c}. \end{aligned} \tag{F.1}$$

Using (E.2)-(E.3) and (F.1), we have

$$\hat{\boldsymbol{\beta}}_{\mathrm{R}} - \boldsymbol{\beta}_0 = \boldsymbol{\Omega}_{\boldsymbol{A}}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right) - n^{-1/2}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}}\left\{\boldsymbol{A}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}}\right\}^{-1}\boldsymbol{c} + o_P(n^{-1/2}). \tag{F.2}$$

According to Theorem 1, we have

$$\begin{aligned} &\sqrt{n}\left(\hat{\boldsymbol{\beta}}_{\mathrm{R}} - \boldsymbol{\beta}_0\right) \\ &\xrightarrow{L} N(-\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}}\left\{\boldsymbol{A}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}}\right\}^{-1}\boldsymbol{c}, \boldsymbol{\Omega}_{\boldsymbol{A}}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_\epsilon\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Omega}_{\boldsymbol{A}}^{\mathrm{T}} + \boldsymbol{\Omega}_{\boldsymbol{A}}\boldsymbol{\Sigma}_{\phi,\boldsymbol{\psi}}\boldsymbol{\Omega}_{\boldsymbol{A}}^{\mathrm{T}}). \end{aligned} \tag{F.3}$$

Step 2 Under the local alternative hypothesis $\mathcal{H}_{1n} : \boldsymbol{A}\boldsymbol{\beta}_0 = \boldsymbol{b} + n^{-1/2}\boldsymbol{c}$, using Theorem 1, we have

$$\begin{aligned} \sqrt{n}\left(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}\right) &= \sqrt{n}\left(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{A}\boldsymbol{\beta}_0 + n^{-1/2}\boldsymbol{c}\right) \\ &= \sqrt{n}\boldsymbol{A}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right) + \boldsymbol{c} \\ &\xrightarrow{L} N\left(\boldsymbol{c}, \boldsymbol{A}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{0\epsilon}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{A}\boldsymbol{\Sigma}_{\phi,\boldsymbol{\psi}}\boldsymbol{A}^{\mathrm{T}}\right). \end{aligned} \tag{F.4}$$

Using (E.5)-(E.6) and (F.4), we have

$$\left[\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}_{\epsilon}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{A}\hat{\boldsymbol{\Sigma}}_{\phi_M,\boldsymbol{\psi}_M}\boldsymbol{A}^{\mathrm{T}}\right]^{-1/2}\left[\sqrt{n}\left(\boldsymbol{A}\hat{\boldsymbol{\beta}}-\boldsymbol{b}\right)\right] \tag{F.5}$$
$$\xrightarrow{L} N\left(\left[\boldsymbol{A}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{0\epsilon}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{A}\boldsymbol{\Sigma}_{\phi,\boldsymbol{\psi}}\boldsymbol{A}^{\mathrm{T}}\right]^{-1/2}\boldsymbol{c},\boldsymbol{I}_k\right).$$

Then, according to (F.5), the continuous mapping theorem entails that

$$\begin{aligned}
&\mathcal{T}_n \\
&= n\left(\boldsymbol{A}\hat{\boldsymbol{\beta}}-\boldsymbol{b}\right)^{\mathrm{T}}\left[\boldsymbol{A}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\Sigma}}_{\epsilon}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{A}\hat{\boldsymbol{\Sigma}}_{\phi,\boldsymbol{\psi}}\boldsymbol{A}^{\mathrm{T}}\right]^{-1}\left(\boldsymbol{A}\hat{\boldsymbol{\beta}}-\boldsymbol{b}\right) \\
&\xrightarrow{L} \chi_k^2(\pi_{\boldsymbol{c}}),
\end{aligned} \tag{F.6}$$

where $\chi_k^2(\pi_{\boldsymbol{c}})$ is the noncentral chi-squared distribution with degree of freedom k, and $\pi_{\boldsymbol{c}}$ is the noncentrality parameter, defined as $\pi_{\boldsymbol{c}}=\boldsymbol{c}^{\mathrm{T}}\left[\boldsymbol{A}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{0\epsilon}\boldsymbol{\Sigma}_0^{-1}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{A}\boldsymbol{\Sigma}_{\phi,\boldsymbol{\psi}}\boldsymbol{A}^{\mathrm{T}}\right]^{-1}\boldsymbol{c}$. We have completed the proof of Theorem 7.

5. Proof of Theorem 8

Step 1 In this step, we establish the asymptotic order of minimizer estimator $\hat{\boldsymbol{\beta}}_P$. Define

$$\mathcal{L}_P(\boldsymbol{\beta})=\frac{1}{2}\sum_{i=1}^n\left\{\hat{Y}_i-\hat{S}_Y(Z_i)-\left[\hat{\boldsymbol{X}}_i-\hat{S}_{\boldsymbol{X}}(Z_i)\right]^{\mathrm{T}}\boldsymbol{\beta}\right\}^2+n\sum_{s=1}^p p_{\zeta_s}(|\beta_s|).$$

Let $\kappa_n=n^{-1/2}+a_n^*$ with $a_n^*=\max_{1\le j\le p}\{p'_{\zeta_j}(|\beta_{0j}|),\beta_{0j}\neq 0\}$, and $\boldsymbol{s}=(s_1,\ldots,s_p)^{\mathrm{T}}$ with $\|\boldsymbol{s}\|=C_0$. Moreover, we define $\boldsymbol{\beta}(n)=\boldsymbol{\beta}_0+\kappa_n\boldsymbol{s}$ and

$$\begin{aligned}
\mathcal{F}_{n,1} &= \frac{1}{2}\sum_{i=1}^n\left\{\hat{Y}_i-\hat{S}_Y(Z_i)-\left[\hat{\boldsymbol{X}}_i-\hat{S}_{\boldsymbol{X}}(Z_i)\right]^{\mathrm{T}}\boldsymbol{\beta}(n)\right\}^2 \\
&\quad -\frac{1}{2}\sum_{i=1}^n\left\{\hat{Y}_i-\hat{S}_Y(Z_i)-\left[\hat{\boldsymbol{X}}_i-\hat{S}_{\boldsymbol{X}}(Z_i)\right]^{\mathrm{T}}\boldsymbol{\beta}_0\right\}^2 \\
\mathcal{F}_{n,2} &= -n\sum_{j=1}^{p_0}\{p_{\zeta_j}(|\beta_{0j}+\kappa_n s_j|)-p_{\zeta_j}(|\beta_{0j}|)\}.
\end{aligned}$$

Using (B.23)-(B.24), we have

$$\begin{aligned}
\mathcal{F}_{n,1} &= \frac{1}{2}\kappa_n^2\sum_{i=1}^n\boldsymbol{s}^{\mathrm{T}}\left[\hat{\boldsymbol{X}}_i-\hat{S}_{\boldsymbol{X}}(Z_i)\right]^{\otimes 2}\boldsymbol{s} \\
&\quad -\kappa_n\sum_{i=1}^n\boldsymbol{s}^{\mathrm{T}}\left[\hat{\boldsymbol{X}}_i-\hat{S}_{\boldsymbol{X}}(Z_i)\right]^{\mathrm{T}}\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0\right) \\
&= \frac{n}{2}\kappa_n^2\boldsymbol{s}^{\mathrm{T}}\boldsymbol{\Sigma}_0\boldsymbol{s}-n\kappa_n\boldsymbol{s}^{\mathrm{T}}\boldsymbol{\Sigma}_0\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0\right) \\
&\quad +o_P(n\kappa_n^2C_0^2)+o_P(n^{1/2}\kappa_nC_0)
\end{aligned} \tag{G.1}$$

As $a_n^*=O_P(n^{-1/2})$, we have $\kappa_n=O_P(n^{-1/2})$ and the asymptotic expression (G.1) entails that the first argument of $\mathcal{D}_{n,1}$ is positive and dominated by $\frac{n}{2}\kappa_n^2C_0^2$ in probability and the second argument of is dominated by $C_0O_P(1)$. Taylor expansion and Cauchy-Schawz inequality entail that

$$|\mathcal{F}_{n,2}|\le n\sqrt{p_0}\kappa_na_n^*\|\boldsymbol{s}\|+n\kappa_n^2a_n^{**}\|\boldsymbol{s}\|^2\le C_0n\kappa_n^2\{\sqrt{p_0}+a_n^{**}C_0\}.$$

where $a_n^{**} = \max_{1\le j\le p}\{p''_{\zeta_j}(|\beta_{0j}|), \beta_{0j} \neq 0\}$. Furthermore, $\mathcal{D}_{n,2}$ is bounded by $n\kappa_n^2 C_0^2$ in probability. Thus, as a_n^{**}, b_n^{**} tend to 0 and C_0 sufficiently large, $\mathcal{D}_{n,1}$ dominates $\mathcal{D}_{n,2}$. As a consequence, for any given $\delta > 0$, there exists a large constant C_0 such that

$$P\left\{ \inf_{\mathcal{S}} \mathcal{L}_P(\boldsymbol{\beta}(n)) > \mathcal{L}_P(\boldsymbol{\beta}_0) \right\} \ge 1 - \delta,$$

where $\mathcal{S} = \{\boldsymbol{s} : \|\boldsymbol{s}\| = C_0\}$. We conclude that $\hat{\boldsymbol{b}}_P$ is $O_P(n^{-1/2})$.

Step 2. Let $\boldsymbol{\beta}_{\mathbf{1}}^*$ satisfies $\|\boldsymbol{\beta}_{\mathbf{1}}^* - \boldsymbol{\beta}_{0\mathbf{1}}\| = O_P(n^{-1/2})$. Similar to the proof of Lemma 1 in Fan and Li (2001), we can show that

$$\mathcal{L}_P\left((\boldsymbol{\beta}_{\mathbf{1}}^{*\mathrm{T}}, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}} \right) = \min_{\mathcal{L}^*} \mathcal{L}_P\left((\boldsymbol{\beta}_{\mathbf{1}}^{*\mathrm{T}}, \boldsymbol{\beta}_{\mathbf{2}}^{*\mathrm{T}})^{\mathrm{T}} \right), \tag{G.2}$$

where, $\mathcal{L}^* = \{\|\boldsymbol{\beta}_{\mathbf{2}}^*\| \le L^* n^{-1/2}\}$ and L^* is a positive constant. We omit the details for the proof in this step.

Step 3. Denote that $\hat{\boldsymbol{\beta}}_{P,\mathbf{1}}$ is the penalized least squares estimator of $\boldsymbol{\beta}_{0,\mathbf{1}}$. In addition, we denote that $\hat{\boldsymbol{X}}_{i,\mathbf{1}}$ and $\hat{S}_{\boldsymbol{X},\mathbf{1}}(Z_i)$ consist of the first p_0 components of $\hat{\boldsymbol{X}}_i$ and $\hat{S}_{\boldsymbol{X}}(Z_i)$, respectively. Define $\mathcal{L}_P^*(\boldsymbol{\beta}_{\mathbf{1}}) = \mathcal{L}_P\left((\boldsymbol{\beta}_{\mathbf{1}}^{\mathrm{T}}, \mathbf{0}^{\mathrm{T}})^{\mathrm{T}} \right)$. Taylor expansion entails that

$$\begin{aligned}
\mathbf{0} &= \left.\frac{\partial \mathcal{L}_P^*(\boldsymbol{\beta}_{\mathbf{1}})}{\partial \boldsymbol{\beta}_{\mathbf{1}}}\right|_{\boldsymbol{\beta}_{\mathbf{1}} = \hat{\boldsymbol{\beta}}_{P,\mathbf{1}}} \qquad \text{(G.3)}\\
&= -\sum_{i=1}^n \left[\hat{\boldsymbol{X}}_{i,\mathbf{1}} - \hat{S}_{\boldsymbol{X},\mathbf{1}}(Z_i) \right] \left\{ \hat{Y}_i - \left[\hat{\boldsymbol{X}}_{i,\mathbf{1}} - \hat{S}_{\boldsymbol{X},\mathbf{1}}(Z_i) \right]^{\mathrm{T}} \boldsymbol{\beta}_{0,\mathbf{1}} \right\} \\
&\quad + n\mathcal{R}_{\boldsymbol{\zeta}_{\mathbf{1}}} + \left(\sum_{i=1}^n \left[\hat{\boldsymbol{X}}_{i,\mathbf{1}} - \hat{S}_{\boldsymbol{X},\mathbf{1}}(Z_i) \right]^{\otimes 2} + n\boldsymbol{\Sigma}_{\boldsymbol{\zeta}_{\mathbf{1}}} \right) \left(\hat{\boldsymbol{\beta}}_{P,\mathbf{1}} - \boldsymbol{\beta}_{0,\mathbf{1}} \right) + O_P(\delta_n),
\end{aligned}$$

where $\delta_n = n\|\hat{\boldsymbol{\beta}}_{P,\mathbf{1}} - \boldsymbol{\beta}_{0\mathbf{1}}\|^2$. Similar to (B.24), we have that

$$\begin{aligned}
&\frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\hat{\boldsymbol{X}}_{i,\mathbf{1}} - \hat{S}_{\boldsymbol{X},\mathbf{1}}(Z_i) \right] \left\{ \hat{Y}_i - \left[\hat{\boldsymbol{X}}_{i,\mathbf{1}} - \hat{S}_{\boldsymbol{X},\mathbf{1}}(Z_i) \right]^{\mathrm{T}} \boldsymbol{\beta}_{0,\mathbf{1}} \right\} \qquad \text{(G.4)}\\
&\xrightarrow{\mathcal{L}} N\left(\mathbf{0}_{p_0}, \boldsymbol{\Sigma}_{0\epsilon,\mathbf{1}} + \boldsymbol{\Sigma}_{0,\mathbf{1}} \boldsymbol{\Sigma}_{\phi,\boldsymbol{\psi}_{\mathbf{1}}} \boldsymbol{\Sigma}_{0,\mathbf{1}} \right),
\end{aligned}$$

where $\boldsymbol{\Sigma}_{0\epsilon,\mathbf{1}}$, $\boldsymbol{\Sigma}_{0,\mathbf{1}}$ and $\boldsymbol{\Sigma}_{\phi,\boldsymbol{\psi}_{\mathbf{1}}}$ are defined in Theorem 1. The asymptotic expression (G.2) and (G.4) entail that

$$\begin{aligned}
&\sqrt{n}\left(\boldsymbol{\Sigma}_{0,\mathbf{1}} + \boldsymbol{\Sigma}_{\boldsymbol{\zeta}_{\mathbf{1}}} \right) \left\{ \left(\hat{\boldsymbol{\beta}}_{P,\mathbf{1}} - \boldsymbol{\beta}_{0,\mathbf{1}} \right) + \left(\boldsymbol{\Sigma}_{0,\mathbf{1}} + \boldsymbol{\Sigma}_{\boldsymbol{\zeta}_{\mathbf{1}}} \right)^{-1} \mathcal{R}_{\boldsymbol{\zeta}_{\mathbf{1}}} \right\} \\
&\xrightarrow{\mathcal{L}} N\left(\mathbf{0}_{p_0}, \boldsymbol{\Sigma}_{0\epsilon,\mathbf{1}} + \boldsymbol{\Sigma}_{0,\mathbf{1}} \boldsymbol{\Sigma}_{\phi,\boldsymbol{\psi}_{\mathbf{1}}} \boldsymbol{\Sigma}_{0,\mathbf{1}} \right).
\end{aligned}$$

We have completed the proof of Theorem 8.

References

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties, *Journal of the American Statistical Association* **96**(456): 1348–1360.

Liang, H. and Li, R. (2009). Variable selection for partially linear models with measurement errors, *Journal of the American Statistical Association* **104**: 234–248.

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., Zhou, N., Sun, Z., Li, G. and Wei, Z. (2016). Statistical inference on restricted partial linear regression models with partial distortion measurement errors, *Statistica Neerlandica. Journal of the Netherlands Society for Statistics and Operations Research* **70**(4): 304–331.