Generalized Partial linear varying multi-index coefficient model for gene-environment interactions

Xu Liu[1], Bin Gao[2] and Yuehua Cui[2*]

[1]*School of Statistics and Management, Shanghai University of Finance and Economics, China*

[2]*Department of Statistics and Probability, Michigan State University, USA.*

Summary

This Supplementary file contains two sections. The numeric algorithm is given in the first Section. The second Section gives the detailed proofs of Lemmas and Theorems.

1 Algorithm

In this section, we give the details of our algorithm.

Algorithm.

- **Step 0**. Choose initial values for $\boldsymbol{\beta}_l$, denoted by $\boldsymbol{\beta}_l^{(old)}$, $l = 0, 1$.

- **Step 1**. Calculate $u_l = \mathbf{X}\boldsymbol{\beta}_l^{(old)}$, and $\mathbf{B}_r(u_l)$, $l = 0, 1$.

- **Step 2**. Update $\widehat{\boldsymbol{\alpha}}$ and $\widehat{\lambda}(\boldsymbol{\beta})$ by minimizing:

$$(\widehat{\boldsymbol{\alpha}}^T, \widehat{\lambda}(\boldsymbol{\beta})^T)^T = \mathop{\arg\min}_{\boldsymbol{\alpha} \in \boldsymbol{\Theta}_\alpha, \lambda \in \mathbb{R}^{2J_n}} \ell_n(\boldsymbol{\theta}, \lambda(\boldsymbol{\beta})),$$

  where $\ell_n(\boldsymbol{\theta}, \lambda(\boldsymbol{\beta}))$ is defined in (2.4) in the paper.

- **Step 3**. Update $\boldsymbol{\beta}_l^{(old)}$ by $\boldsymbol{\beta}_l^{(old)} = \boldsymbol{\beta}_l^{(new)} / \|\boldsymbol{\beta}_l^{(new)}\|_2$, $l = 0, 1$, where

$$\widehat{\boldsymbol{\beta}}^{(new)} = \mathop{\arg\min}_{\boldsymbol{\beta} \in \boldsymbol{\Theta}_\beta} \ell_n((\widehat{\boldsymbol{\alpha}}^T, \boldsymbol{\beta}^T)^T, \widehat{\lambda}(\boldsymbol{\beta})).$$

- **Step 4**. Repeat Steps 1-3 until convergence. This gives the whole parametric estimator $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\alpha}}^T, \widehat{\boldsymbol{\beta}}^T)^T$ and $\widehat{\lambda}_l(\widehat{\boldsymbol{\theta}})$, which results in $\tilde{m}_0(\widehat{\boldsymbol{\beta}}_l^T \mathbf{X}, \widehat{\boldsymbol{\beta}})$, $l = 0, 1$.

*To whom correspondence should be addressed: cui@stt.msu.edu

- **Step 5**. Given the parametric estimator $\widehat{\boldsymbol{\theta}}$ in Step 4, calculate $\widehat{\eta}_{-1}(\mathbf{V}_i; a, b) = \widehat{\boldsymbol{\alpha}}_0^T \mathbf{Z}_i + \tilde{m}_1(\widehat{\boldsymbol{\beta}}_0^T \mathbf{X}_i, \widehat{\boldsymbol{\beta}}) + aG_i + b(\widehat{\boldsymbol{\beta}}_1^T \mathbf{X}_i - u_1)G_i$, and then solve following minimization problem to obtain local linear estimator $\widehat{m}_1(u_1, \widehat{\boldsymbol{\beta}}) = \widehat{a}$,

$$(\widehat{a}, \widehat{b}) = \arg\min_{a,b} \sum_{i=1}^{n} Q(g^{-1}\{\widehat{\eta}_{-1}(\mathbf{V}_i; a, b)\}, Y_i) K_{h_1}(\widehat{\boldsymbol{\beta}}_1^T \mathbf{X}_i - u_1).$$

  As the same way, compute $\widehat{m}_1(u_0, \widehat{\boldsymbol{\beta}})$ by calculating $\widehat{\eta}_{-1}(\mathbf{V}_i; a, b) = a + b(\widehat{\boldsymbol{\beta}}_0^T \mathbf{X}_i - u_0) + \widehat{\boldsymbol{\alpha}}_0^T \mathbf{Z}_i + (\tilde{m}_1(\widehat{\boldsymbol{\beta}}_1^T \mathbf{X}_i, \widehat{\boldsymbol{\beta}}) + \widehat{\boldsymbol{\alpha}}_1^T \mathbf{Z}_i)G_i$.

Remark S1: Unlike the algorithms proposed by Carroll et al. (1997) and Cui et al. (2011), we do not need to calculate iteratively parameter and non-parameter based on kernel estimation. Thus, it is much faster than their algorithm. In Step 3, the Newton-Raphson algorithm is used to calculate $\boldsymbol{\beta}$.

2 Proofs

Let $m(\mathbf{V}_i, \boldsymbol{\beta}) = m_0(\mathbf{V}_i, \boldsymbol{\beta}_0) + m_1(\mathbf{V}_i, \boldsymbol{\beta}_1)G_i$ and $\mathbf{m} = (m(\mathbf{V}_1, \boldsymbol{\beta}), \cdots, m(\mathbf{V}_n, \boldsymbol{\beta}))^T$. Denote $\mathbb{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)^T$, $\mathbb{Z} = (\mathbf{Z}_1, \cdots, \mathbf{Z}_n)^T$, $\mathbf{G} = (G_1, \cdots, G_n)^T$, and $\mathbb{G} = (\mathbf{1}_n, \mathbf{G})$. Let $\boldsymbol{\Theta}$ be the parametric space of $\boldsymbol{\theta}$. Let $q_j(x) = (\partial^j/\partial x^j)Q\{g^{-1}(x), y\}$, $j = 1, 2, 3$. Then $q_1(x) = \{y - g^{-1}(x)\}\rho_1(x)$ and $q_2(x) = \{y - g^{-1}(x)\}\rho_1'(x) - \rho_2(x)$, and $\rho_j(x) = \{dg^{-1}(x)/dx\}^j / V\{g^{-1}(x)\}$. Denote $q_k(\tilde{\eta}_i)$ by $q_k\{\tilde{\eta}(\mathbf{V}_i; \boldsymbol{\theta}_0, \lambda(\boldsymbol{\beta}^0))\}$, $k = 1, 2$, $i = 1, \cdots, n$. Let $\mathbf{q}_k = (q_k(\tilde{\eta}_1), \cdots, q_k(\tilde{\eta}_n))^T$ and $\mathbf{W}_{q_2}$ be a diagonal matrix with diagonal elements $\mathbf{q}_2\{\tilde{\eta}\}$. Define

$$\begin{aligned} \mathbf{U}(\boldsymbol{\beta}) &= E[q_2(\tilde{\eta}_i) D_i(\boldsymbol{\beta}) D_i(\boldsymbol{\beta})^T], \quad \widehat{\mathbf{U}}(\boldsymbol{\beta}) = \frac{1}{n}\mathbf{D}(\boldsymbol{\beta})^T \mathbf{W}_{q_2} \mathbf{D}(\boldsymbol{\beta}), \\ \mathbf{U}(\tilde{\mathbf{Z}}, \boldsymbol{\beta}) &= E[q_2(\tilde{\eta}_i) D_i(\tilde{\mathbf{Z}}, \boldsymbol{\beta}) D_i(\tilde{\mathbf{Z}}, \boldsymbol{\beta})^T], \quad \widehat{\mathbf{U}}(\tilde{\mathbf{Z}}, \boldsymbol{\beta}) = \frac{1}{n}\mathbf{D}(\tilde{\mathbf{Z}}, \boldsymbol{\beta})^T \mathbf{W}_{q_2} \mathbf{D}(\tilde{\mathbf{Z}}, \boldsymbol{\beta}) \end{aligned} \tag{S.1}$$

where $D_i(\boldsymbol{\beta}) = (D_{i,sl}(\boldsymbol{\beta}_l)\mathbb{G}_{il}, 1 \le s \le J_n, l = 0, 1)^T$ and $\mathbf{D}(\boldsymbol{\beta}) = (D_1(\boldsymbol{\beta}), \cdots, D_n(\boldsymbol{\beta}))^T$, which is a $n \times 2J_n$ matrix.

Lemma S.1 *If assumptions* (A1), (A4), and (A5) *hold, for any vector* $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_0^T, \boldsymbol{\zeta}_1^T)^T$ *with* $\boldsymbol{\zeta}_l = (\boldsymbol{\zeta}_{s,l} : 1 \le s \le J_n)^T$ *and* $||\boldsymbol{\zeta}_l|| = 1$, $l = 0, 1$, *there exists constants* $0 < c_U < C_U < \infty$, *such that for* *any* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ *and for large enough* n,

$$c_U J_n^{-1} \leq \boldsymbol{\zeta}^T \mathbf{U}(\boldsymbol{\beta})\boldsymbol{\zeta} \leq C_U J_n^{-1}, \quad \textit{and} \quad C_U^{-1} J_n \leq \boldsymbol{\zeta}^T \mathbf{U}(\boldsymbol{\beta})^{-1}\boldsymbol{\zeta} \leq c_U^{-1} J_n, \tag{S.2}$$

$$\begin{aligned}\sup_{1\leq s,s'\leq J_n, 0\leq l\leq 1} &\left| n^{-1}\sum_{i=1}^{n} D_{i,sl}(\boldsymbol{\beta}_l)D_{i,s'l}(\boldsymbol{\beta}_l) - E[D_{i,sl}(\boldsymbol{\beta}_l)D_{i,s'l}(\boldsymbol{\beta}_l)]\right| \\ &= O\left(\sqrt{J_n^{-1}n^{-1}\log n}\right), \quad a.s.,\end{aligned} \tag{S.3}$$

$$\begin{aligned}\sup_{1\leq s,s'\leq J_n, l\neq l'} &\left| n^{-1}\sum_{i=1}^{n} D_{i,sl}(\boldsymbol{\beta}_l)D_{i,s'l'}(\boldsymbol{\beta}_l) - E[D_{i,sl}(\boldsymbol{\beta}_l)D_{i,s'l'}(\boldsymbol{\beta}_l)]\right| \\ &= O\left(J_n^{-1}\sqrt{n^{-1}\log n}\right), \quad a.s.,\end{aligned} \tag{S.4}$$

and with probability approaching 1,

$$c_U J_n^{-1} \leq \boldsymbol{\zeta}^T \widehat{\mathbf{U}}(\boldsymbol{\beta})\boldsymbol{\zeta} \leq C_U J_n^{-1}, \quad \textit{and} \quad C_U^{-1} J_n \leq \boldsymbol{\zeta}^T \widehat{\mathbf{U}}(\boldsymbol{\beta})\boldsymbol{\zeta} \leq c_U^{-1} J_n. \tag{S.5}$$

Proof of Lemma S.1: Lemma S.1 can be shown along the lines of Liu et al. (2016).

□

Lemma S.2 *Let assumptions* (A1) and (A3)-(A5) *be satisfied. For any* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $||n^{-1}\mathbf{D}(\boldsymbol{\beta})^T\mathbf{q}_1\{\tilde{\eta}\}||_2 = O_p(n^{-1/2})$.

Proof of Lemma S.2: By the law of large numbers, we have

$$\begin{aligned}||n^{-1}\mathbf{D}(\boldsymbol{\beta})^T\mathbf{q}_1\{\tilde{\eta}\}||_2^2 =& \sum_{1\leq s\leq J_n, l=0,1}\left[n^{-1}\sum_{i=1}^{n} B_{s,q}(\boldsymbol{\beta}_l^T\mathbf{X}_i)\mathbb{G}_{il}q_1(\tilde{\eta}_i)\right]^2 \\ =& n^{-1}\sum_{1\leq s\leq J_n, l=0,1} E\left[B_{s,q}(\boldsymbol{\beta}_l^T\mathbf{X}_i)\mathbb{G}_{il}q_1(\tilde{\eta}_i)\right]^2 + o_p(n^{-1}) \\ =& O_p(n^{-1}).\end{aligned}$$

□

Lemma S.3 *If assumptions* (A1), (A4), and (A5) *hold, there exists a constant* $0 \leq c_D \leq \infty$*, such that for any* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ *and for large enough* n*,*

$$\|n^{-1}\widehat{\mathbf{U}}(\tilde{\mathbf{Z}}, \boldsymbol{\beta})^{-1}\mathbf{D}(\tilde{\mathbf{Z}}, \boldsymbol{\beta})^T\mathbf{W}_{q_2}\mathbf{1}_n\|_\infty \leq C_D.$$

Proof of Lemma S.3: According to Assumptions (A4) and (A5), the proof of Lemma S.3 can be completed as the same as Liu et al. (2016).

□

Lemma S.4 *If assumptions* (A1)-(A5) *hold, we have*

(a) $N \to \infty$ *and* $nN^{-1} \to \infty$*, as* $n \to \infty$*,*

$$|\tilde{m}_0(u_l, \boldsymbol{\beta}^0) - m_l(u_l)| = O_p(\sqrt{N/n} + N^{-r})$$

uniformly for any $u_l \in [a_l, b_l]$*;*

(b) under $N \to \infty$ *and* $nN^{-3} \to \infty$*, as* $n \to \infty$*,*

$$|\tilde{m}_0'(u_l, \boldsymbol{\beta}^0) - m_l'(u_l)| = O_p(\sqrt{N^3/n} + N^{1-r})$$

uniformly for any $u_l \in [a_l, b_l]$*.*

Proof of Lemma S.4: This Lemma can be proved by the similar argue of Lemma S.4 in Liu et al. (2016).

□

Let $\widehat{\boldsymbol{\zeta}}_k = n^{-1}\widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1}\mathbf{D}(\boldsymbol{\beta}^0)^T\mathbf{W}_{q_2}\bar{\mathbf{X}}_k$, which is

$$\widehat{\boldsymbol{\zeta}}_k^X = \underset{\boldsymbol{\zeta}^X \in \mathcal{R}^{2J_n}}{\arg\min} \sum_{i=1}^n q_2\{\tilde{\eta}(\mathbf{V}_i; \boldsymbol{\theta}_0, \lambda(\boldsymbol{\beta}^0))\}(X_{ik} - D_i(\boldsymbol{\beta}^0)\boldsymbol{\zeta}_k^X)^2,$$

where $\bar{\mathbf{X}}_k = (X_{1k}, \cdots, X_{nk})^T$. Let $\mathbf{P}_n(\mathbf{X}_i) = D_i(\boldsymbol{\beta}^0)^T\widehat{\boldsymbol{\zeta}}^X$, where $\widehat{\boldsymbol{\zeta}}^X = (\widehat{\boldsymbol{\zeta}}_1^X, \cdots, \widehat{\boldsymbol{\zeta}}_p^X)$ be a $2J_n \times p$

matrix. Let $\mathbf{P}_n(\mathbf{Z}_i) = D_i(\boldsymbol{\beta}^0)^T\widehat{\boldsymbol{\zeta}}^Z$ with $\widehat{\boldsymbol{\zeta}}^Z = (\widehat{\boldsymbol{\zeta}}_1^Z, \cdots, \widehat{\boldsymbol{\zeta}}_q^Z)$, where

$$\widehat{\boldsymbol{\zeta}}_k^Z = \underset{\boldsymbol{\zeta}^Z\in\mathcal{R}^{2J_n}}{\arg\min}\sum_{i=1}^n q_2\{\tilde{\eta}(\mathbf{V}_i;\boldsymbol{\theta}_0,\lambda(\boldsymbol{\beta}^0))\}(Z_{ik} - D_i(\boldsymbol{\beta}^0)\boldsymbol{\zeta}_k^Z)^2.$$

Furthermore, we define

$$\begin{aligned}
\mathbf{X}_{m,i} =&(m_0'(\mathbf{X}_i^T\boldsymbol{\beta}_0^0)\mathbf{X}_i^T, m_1'(\mathbf{X}_i^T\boldsymbol{\beta}_1^0)G_i\mathbf{X}_i^T)^T,\\
\mathbf{X}_{\widehat{m},i} =&(\mathbf{B}_r'(\mathbf{X}_i^T\widehat{\boldsymbol{\beta}}_0)^T\widehat{\lambda}_0(\widehat{\boldsymbol{\theta}})\mathbf{X}_i^T, \mathbf{B}_r'(\mathbf{X}_i^T\widehat{\boldsymbol{\beta}}_1)^T\widehat{\lambda}_1(\widehat{\boldsymbol{\theta}})G_i\mathbf{X}_i^T)^T,\\
\mathbb{X}_m =&(\mathbf{X}_{m,1}^T, \cdots, \mathbf{X}_{m,n}^T)^T,\\
\mathbb{X}_{\widehat{m}} =&(\mathbf{X}_{\widehat{m},1}^T, \cdots, \mathbf{X}_{\widehat{m},n}^T)^T.
\end{aligned}$$

and

$$\begin{aligned}
\tilde{\mathbf{P}}_n(\mathbf{X}_i) =&(m_0'(\mathbf{X}_i^T\boldsymbol{\beta}_0^0)\mathbf{P}_n(\mathbf{X}_i)^T, m_1'(\mathbf{X}_i^T\boldsymbol{\beta}_0^0)G_i\mathbf{P}_n(\mathbf{X}_i)^T)^T,\\
\tilde{\mathbf{P}}_n(\mathbf{Z}_i) =&(\mathbf{P}_n(\mathbf{Z}_i)^T, G_i\mathbf{P}_n(\mathbf{Z}_i)^T)^T,\\
\tilde{\mathbf{P}}(\mathbf{X}_i) =&(m_0'(\mathbf{X}_i^T\boldsymbol{\beta}_0^0)\mathbf{P}(\mathbf{X}_i)^T, m_1'(\mathbf{X}_i^T\boldsymbol{\beta}_0^0)G_i\mathbf{P}(\mathbf{X}_i)^T)^T,\\
\tilde{\mathbf{P}}(\mathbf{Z}_i) =&(\mathbf{P}(\mathbf{Z}_i)^T, G_i\mathbf{P}(\mathbf{Z}_i)^T)^T.
\end{aligned}$$

Lemma S.5 *If assumptions* (A1)-(A6) *hold, and* $nN^{-4}\to\infty$ *and* $nN^{-2r-2}\to 0$, *as* $n\to\infty$. *We have*

$$\begin{aligned}
D_i(\boldsymbol{\beta}^0)^T\{\widehat{\lambda}(\widehat{\boldsymbol{\beta}}) - \lambda\} =&n^{-1}D_i(\boldsymbol{\beta}^0)^T\widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1}\mathbf{D}(\boldsymbol{\beta}^0)^T\mathbf{q}_1\{\tilde{\eta}\}\\
&-\tilde{\mathbf{P}}(\mathbf{X}_i)^T(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) - \tilde{\mathbf{P}}(\mathbf{Z}_i)^T(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^0)\\
&+o_p(\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0\|_2) + o_p(\|\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^0\|_2) + o_p(n^{-1/2}).
\end{aligned}\tag{S.6}$$

Proof of Lemma S.5: $\lambda(\boldsymbol{\theta})$ can be solved by equation

$$0 = n^{-1}\sum_{i=1}^n q_1\{\tilde{\eta}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}},\widehat{\lambda}(\widehat{\boldsymbol{\beta}}))\}D_i(\widehat{\boldsymbol{\beta}}).$$

Because $\widehat{\mathbf{U}}(\boldsymbol{\beta}^0) = O_p(J_n^{-1})$ and $nN^{-2r-2} \to 0$, we have

$$
\begin{aligned}
0 =& n^{-1}\sum_{i=1}^{n} D_i(\boldsymbol{\beta}^0) q_1\{\tilde{\eta}_i\} \\
& - n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\} D_i(\boldsymbol{\beta}^0) D_i(\boldsymbol{\beta}^0)^T \{(\widehat{\lambda}(\widehat{\boldsymbol{\beta}}) - \lambda(\boldsymbol{\beta}^0)) + o_p(\sqrt{N/n} + N^{-r})\} \\
& - n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\} D_i(\boldsymbol{\beta}^0) X_{\widehat{m},i}^T \{(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) + o_p(\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0\|_2)\} \\
& - n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\} D_i(\boldsymbol{\beta}^0) \tilde{\mathbf{Z}}_i^T \{(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^0) + o_p(\|\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^0\|_2)\} \\
=& n^{-1}\mathbf{D}(\boldsymbol{\beta}^0)^T \mathbf{q}_1\{\tilde{\eta}\} - \widehat{\mathbf{U}}(\boldsymbol{\beta}^0)(\widehat{\lambda}(\widehat{\boldsymbol{\beta}}) - \lambda(\boldsymbol{\beta}^0)) + o_p(\sqrt{N/n} + N^{-r}) \\
& - n^{-1}\mathbf{D}(\boldsymbol{\beta}^0)^T \mathbf{W}_{q_2} \mathbb{X}_{\widehat{m}} \{(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) + o_p(\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0\|_2\} \\
& - n^{-1}\mathbf{D}(\boldsymbol{\beta}^0)^T \mathbf{W}_{q_2} \tilde{\mathbf{Z}} \{(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^0) + o_p(\|\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^0\|_2)\} + o_p(n^{-1/2}),
\end{aligned}
$$

where $\mathbf{W}_{q_2}$ is a diagonal matrix with diagonal elements $\mathbf{q}_2\{\tilde{\eta}\}$. Thus, we have

$$
\begin{aligned}
& D_i(\boldsymbol{\beta}^0)^T \{\widehat{\lambda}(\widehat{\boldsymbol{\beta}}) - \lambda(\boldsymbol{\beta}^0)\} \\
=& n^{-1} D_i(\boldsymbol{\beta}^0)^T \widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1} \mathbf{D}(\boldsymbol{\beta}^0)^T \mathbf{q}_1\{\tilde{\eta}\} \\
& - n^{-1} D_i(\boldsymbol{\beta}^0)^T \widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1} \mathbf{D}(\boldsymbol{\beta}^0)^T \mathbf{W}_{q_2} \mathbb{X}_{\widehat{m}} \{(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0) + o_p(\|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^0\|_2\} \\
& - n^{-1} D_i(\boldsymbol{\beta}^0)^T \widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1} \mathbf{D}(\boldsymbol{\beta}^0)^T \mathbf{W}_{q_2} \tilde{\mathbf{Z}} \{(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^0) + o_p(\|\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^0\|_2\} + o_p(n^{-1/2}).
\end{aligned} \tag{S.7}
$$

By the same arguments of Liu et al. (2016), we have

$$
\begin{aligned}
n^{-1} D_i(\boldsymbol{\beta}^0)^T \widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1} \mathbf{D}(\boldsymbol{\beta}^0)^T \mathbf{W}_{q_2} \{m_0'(\mathbf{X}_i^T \boldsymbol{\beta}_0^0) X_i^T\}_{i=1}^n &= m_0'(\mathbf{X}_i^T \boldsymbol{\beta}_0^0) \mathbf{P}_n(\mathbf{X}_i) + O_p(J_n^{1-r}), \\
n^{-1} D_i(\boldsymbol{\beta}^0)^T \widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1} \mathbf{D}(\boldsymbol{\beta}^0)^T \mathbf{W}_{q_2} \{m_1'(\mathbf{X}_i^T \boldsymbol{\beta}_1^0) X_i^T G_i\}_{i=1}^n &= m_1'(\mathbf{X}_i^T \boldsymbol{\beta}_1^0) G_i \mathbf{P}_n(\mathbf{X}_i) + O_p(J_n^{1-r}),
\end{aligned}
$$

which are followed by

$$
D_i(\boldsymbol{\beta}^0)^T \widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1} \mathbf{D}(\boldsymbol{\beta}^0)^T \mathbf{W}_{q_2} \mathbb{X}_{\widehat{m}} = \tilde{\mathbf{P}}_n(\mathbf{X}_i) + O_p(J_n^{1-r}).
$$

By the same argument, we have

$$D_i(\boldsymbol{\beta}^0)^T\widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1}\mathbf{D}(\boldsymbol{\beta}^0)^T\mathbf{W}_{q_2}\tilde{\mathbf{Z}} = \tilde{\mathbf{P}}_n(\mathbf{Z}_i) + O_p(J_n^{1-r}).$$

We can show that $\|\mathbf{P}_n(\mathbf{X}_i) - \mathbf{P}(\mathbf{X}_i)\|_\infty = O_p(\sqrt{N/n} + N^{-r})$ and $\|\mathbf{P}_n(\mathbf{Z}_i) - \mathbf{P}(\mathbf{Z}_i)\|_\infty = O_p(\sqrt{N/n} + N^{-r})$ which implies that $\tilde{\mathbf{P}}_n(\mathbf{X}_i) = \tilde{\mathbf{P}}(\mathbf{X}_i) + O_p(\sqrt{N/n} + N^{-r})$ and $\tilde{\mathbf{P}}_n(\mathbf{Z}_i) = \tilde{\mathbf{P}}(\mathbf{Z}_i) + O_p(\sqrt{N/n} + N^{-r})$ Therefore, with (S.7), this arrives at the result of Lemma S.5.

□

Lemma S.6 *If assumptions* (A1)-(A6) *hold, and* $nN^{-4} \to \infty$ *and* $nN^{-2r-2} \to 0$, *as* $n \to \infty$. *We have*

$$\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0 = \{E[\phi(\mathbf{V}, \boldsymbol{\beta}^0)^{\otimes 2}]\}^{-1} n^{-1/2} \sum_{i=1}^{n} \boldsymbol{\Psi}_i(\mathbf{Z}, \mathbf{X}) q_1\{\tilde{\eta}_i\}(1 + O_p(J_n^{1-r})) + o_p(n^{-1/2}) \tag{S.8}$$

where

$$\boldsymbol{\Psi}_i(\mathbf{Z}, \mathbf{X}) = q_2\{\tilde{\eta}_i\} \begin{pmatrix} \tilde{Z}_i - \tilde{\mathbf{P}}(\mathbf{Z}_i) \\ X_{\widehat{m},i} - \tilde{\mathbf{P}}(\mathbf{X}_i) \end{pmatrix}.$$

Proof of Lemma S.6: Denote by τ_0 and τ_1 Lagrange multipliers. The estimates of $\boldsymbol{\beta}$ can be obtaining by solving

$$0 = \begin{pmatrix} \tau_0 \widehat{\boldsymbol{\beta}}_0 \\ \tau_1 \widehat{\boldsymbol{\beta}}_1 \end{pmatrix} + n^{-1} \sum_{i=1}^{n} q_1\{\tilde{\eta}(\mathbf{V}_i; \widehat{\boldsymbol{\theta}}, \widehat{\lambda}(\widehat{\boldsymbol{\beta}}))\} \begin{pmatrix} D_i'(\widehat{\boldsymbol{\beta}})^T \widehat{\lambda}_0(\widehat{\boldsymbol{\beta}}) \mathbf{X}_i \\ D_i'(\widehat{\boldsymbol{\beta}})^T \widehat{\lambda}_1(\widehat{\boldsymbol{\beta}}) G_i \mathbf{X}_i \end{pmatrix}.$$

Denote $q_j\{\tilde\eta_i\}$ by $q_j\{\tilde\eta(\mathbf{V}_i;\boldsymbol\theta^0,\lambda(\boldsymbol\beta^0))\}$, $j=1,2$, $i=1,\cdots,n$. We have

$$
\begin{aligned}
0 =& \begin{pmatrix}\tau_0\widehat{\boldsymbol\beta}_0\\ \tau_1\widehat{\boldsymbol\beta}_1\end{pmatrix} + n^{-1}\sum_{i=1}^n q_1\{\tilde\eta_i\}\mathbf{X}_{\widehat m,i}\\
&- n^{-1}\mathbb{X}_{\widehat m}^T\mathbf{W}_{q_2}\mathbb{X}_m\{(\widehat{\boldsymbol\beta}-\boldsymbol\beta^0)+o_p(\|\widehat{\boldsymbol\beta}-\boldsymbol\beta^0\|_2)\}\\
&- n^{-1}\mathbb{X}_{\widehat m}^T\mathbf{W}_{q_2}\tilde{\mathbf{Z}}\{(\widehat{\boldsymbol\alpha}-\boldsymbol\alpha^0)+o_p(\|\widehat{\boldsymbol\alpha}-\boldsymbol\alpha^0\|_2)\}\\
&- n^{-1}\sum_{i=1}^n q_2\{\tilde\eta_i\}X_{\widehat m,i}D_i(\boldsymbol\beta^0)^T\{\widehat\lambda(\widehat{\boldsymbol\beta})-\lambda(\boldsymbol\theta^0)+o_p(\sqrt{N/n}+N^{-r})\}\\
=& \begin{pmatrix}\tau_0\widehat{\boldsymbol\beta}_0\\ \tau_1\widehat{\boldsymbol\beta}_1\end{pmatrix} + n^{-1}\mathbb{X}_{\widehat m}^T\mathbf{q}_1\{\tilde\eta\} - n^{-1}\mathbb{X}_{\widehat m}^T\mathbf{W}_{q_2}\mathbb{X}_m\{(\widehat{\boldsymbol\beta}-\boldsymbol\beta^0)+o_p(\|\widehat{\boldsymbol\beta}-\boldsymbol\beta^0\|_2)\}\\
&- n^{-1}\mathbb{X}_{\widehat m}^T\mathbf{W}_{q_2}\tilde{\mathbf{Z}}\{(\widehat{\boldsymbol\alpha}-\boldsymbol\alpha^0)+o_p(\|\widehat{\boldsymbol\alpha}-\boldsymbol\alpha^0\|_2)\}\\
&- n^{-2}\sum_{i=1}^n q_2\{\tilde\eta_i\}X_{\widehat m,i}D_i(\boldsymbol\beta^0)^T\widehat{\mathbf{U}}(\boldsymbol\beta^0)^{-1}\mathbf{D}(\boldsymbol\beta^0)^T\mathbf{q}_1\{\tilde\eta\}\\
&+ n^{-1}\sum_{i=1}^n q_2\{\tilde\eta_i\}X_{\widehat m,i}\tilde{\mathbf{P}}(\mathbf{X}_i)^T\{(\widehat{\boldsymbol\beta}-\boldsymbol\beta^0)+o_p(\|\widehat{\boldsymbol\beta}-\boldsymbol\beta^0\|_2)\}\\
&+ n^{-1}\sum_{i=1}^n q_2\{\tilde\eta_i\}X_{\widehat m,i}\tilde{\mathbf{P}}(\mathbf{Z}_i)^T\{(\widehat{\boldsymbol\alpha}-\boldsymbol\alpha^0)+o_p(\|\widehat{\boldsymbol\alpha}-\boldsymbol\alpha^0\|_2)\}\\
=& \begin{pmatrix}\tau_0\widehat{\boldsymbol\beta}_0\\ \tau_1\widehat{\boldsymbol\beta}_1\end{pmatrix} - n^{-1}\sum_{i=1}^n q_2\{\tilde\eta_i\}X_{\widehat m,i}\left[X_{m,i}-\tilde{\mathbf{P}}(\mathbf{X}_i)\right]^T\{(\widehat{\boldsymbol\beta}-\boldsymbol\beta^0)+o_p(\|\widehat{\boldsymbol\beta}-\boldsymbol\beta^0\|_2)\}\\
&- n^{-1}\sum_{i=1}^n q_2\{\tilde\eta_i\}X_{\widehat m,i}\left[\tilde Z_i-\tilde{\mathbf{P}}(\mathbf{Z}_i)\right]^T\{(\widehat{\boldsymbol\alpha}-\boldsymbol\alpha^0)+o_p(\|\widehat{\boldsymbol\alpha}-\boldsymbol\alpha^0\|_2)\}\\
&+ n^{-1}\mathbb{X}_{\widehat m}^T\mathbf{q}_1\{\tilde\eta\} - n^{-2}\sum_{j=1}^n q_1\{\tilde\eta_j\}D_j(\boldsymbol\beta^0)\widehat{\mathbf{U}}(\boldsymbol\beta^0)^{-1}\sum_{i=1}^n q_2\{\tilde\eta_i\}D_i(\boldsymbol\beta^0)^T X_{\widehat m,i} + o_p(n^{-1/2}),\\
=& \begin{pmatrix}\tau_0\widehat{\boldsymbol\beta}_0\\ \tau_1\widehat{\boldsymbol\beta}_1\end{pmatrix} - n^{-1}\sum_{i=1}^n q_2\{\tilde\eta_i\}X_{\widehat m,i}\left[X_{m,i}-\tilde{\mathbf{P}}(\mathbf{X}_i)\right]^T\{(\widehat{\boldsymbol\beta}-\boldsymbol\beta^0)+o_p(\|\widehat{\boldsymbol\beta}-\boldsymbol\beta^0\|_2)\}\\
&- n^{-1}\sum_{i=1}^n q_2\{\tilde\eta_i\}X_{\widehat m,i}\left[\tilde Z_i-\tilde{\mathbf{P}}(\mathbf{Z}_i)\right]^T\{(\widehat{\boldsymbol\alpha}-\boldsymbol\alpha^0)+o_p(\|\widehat{\boldsymbol\alpha}-\boldsymbol\alpha^0\|_2)\}\\
&+ n^{-1}\sum_{i=1}^n\left[X_{\widehat m,i}-\tilde{\mathbf{P}}(\mathbf{X}_i)\right]q_1\{\tilde\eta_i\}\{1+O_p(J_n^{1-r})\}+o_p(n^{-1/2}).
\end{aligned}
$$

By the same way, the estimates of $\boldsymbol{\alpha}$ solve

$$0 = n^{-1}\sum_{i=1}^{n} q_1\{\tilde{\eta}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}},\widehat{\lambda}(\widehat{\boldsymbol{\beta}}))\}\tilde{\mathbf{Z}}_i.$$

And then, we obtain

$$\begin{aligned}
0 =& n^{-1}\sum_{i=1}^{n} q_1\{\tilde{\eta}_i\}\tilde{\mathbf{Z}}_i \\
& - n^{-1}\tilde{\mathbf{Z}}^T\mathbf{W}_{q_2}\mathbb{X}_{\widehat{m}}\{(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0)+o_p(\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0\|_2)\} \\
& - n^{-1}\tilde{\mathbf{Z}}^T\mathbf{W}_{q_2}\tilde{\mathbf{Z}}\{(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^0)+o_p(\|\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^0\|_2)\} \\
& - n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\}\tilde{Z}_i D_i(\boldsymbol{\beta}^0)^T\{\widehat{\lambda}(\widehat{\boldsymbol{\beta}})-\lambda+o_p(\sqrt{N/n}+N^{-r})\} \\
=& n^{-1}\tilde{\mathbf{Z}}^T\mathbf{q}_1\{\tilde{\eta}\} - n^{-1}\tilde{\mathbf{Z}}^T\mathbf{W}_{q_2}\mathbb{X}_{\widehat{m}}\{(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0)+o_p(\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0\|_2)\} \\
& - n^{-1}\tilde{\mathbf{Z}}^T\mathbf{W}_{q_2}\tilde{\mathbf{Z}}\{(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^0)+o_p(\|\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^0\|_2)\} \\
& - n^{-2}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\}\tilde{Z}_i D_i(\boldsymbol{\beta}^0)^T\widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1}\mathbf{D}(\boldsymbol{\beta}^0)^T\mathbf{e} \\
& + n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\}\tilde{Z}_i\tilde{\mathbf{P}}(\mathbf{X}_i)^T\{(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0)+o_p(\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0\|_2)\} \\
& + n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\}\tilde{Z}_i\tilde{\mathbf{P}}(\mathbf{Z}_i)^T\{(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^0)+o_p(\|\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^0\|_2)\} \\
=& - n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\}\tilde{Z}_i\left[X_{\widehat{m},i}-\tilde{\mathbf{P}}(\mathbf{X}_i)\right]^T\{(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0)+o_p(\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0\|_2)\} \\
& - n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\}\tilde{Z}_i\left[\tilde{Z}_i-\tilde{\mathbf{P}}(\mathbf{Z}_i)\right]^T\{(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^0)+o_p(\|\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^0\|_2)\} \\
& + n^{-1}\tilde{\mathbf{Z}}^T\mathbf{q}_1\{\tilde{\eta}\} - n^{-2}\sum_{j=1}^{n} q_1\{\tilde{\eta}_j\}D_j(\boldsymbol{\beta}^0)\widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\}D_i(\boldsymbol{\beta}^0)^T\tilde{Z}_i + o_p(n^{-1/2}), \\
=& - n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\}\tilde{Z}_i\left[X_{\widehat{m},i}-\tilde{\mathbf{P}}(\mathbf{X}_i)\right]^T\{(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0)+o_p(\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}^0\|_2)\} \\
& - n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\}\tilde{Z}_i\left[\tilde{Z}_i-\tilde{\mathbf{P}}(\mathbf{Z}_i)\right]^T\{(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^0)+o_p(\|\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^0\|_2)\} \\
& + n^{-1}\sum_{i=1}^{n}\left[\tilde{Z}_i-\tilde{\mathbf{P}}(\mathbf{Z}_i)\right]q_1\{\tilde{\eta}_i\}\{1+O_p(J_n^{1-r})\}+o_p(n^{-1/2}),
\end{aligned}$$

which arrives

$$\begin{pmatrix} \mathbf{0} \\ \tau_0\widehat{\boldsymbol{\beta}}_0 \\ \tau_1\widehat{\boldsymbol{\beta}}_1 \end{pmatrix} + n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\} \begin{pmatrix} \tilde{Z}_i \\ X_{\widehat{m},i} \end{pmatrix} \begin{pmatrix} \tilde{Z}_i - \tilde{\mathbf{P}}(\mathbf{Z}_i) \\ X_{\widehat{m},i} - \tilde{\mathbf{P}}(\mathbf{X}_i) \end{pmatrix}^T \{(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) + o_p(\|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2)\}$$

$$= n^{-1}\sum_{i=1}^{n} \begin{pmatrix} \tilde{Z}_i - \tilde{\mathbf{P}}(\mathbf{Z}_i) \\ X_{\widehat{m},i} - \tilde{\mathbf{P}}(\mathbf{X}_i) \end{pmatrix} q_1\{\tilde{\eta}_i\}\{1 + O_p(J_n^{1-r})\} + o_p(n^{-1/2}).$$

Let

$$P_{\boldsymbol{\beta}} \equiv \begin{pmatrix} I & 0 & 0 \\ 0 & I - \widehat{\boldsymbol{\beta}}_0\widehat{\boldsymbol{\beta}}_0^T & 0 \\ 0 & 0 & I - \widehat{\boldsymbol{\beta}}_1\widehat{\boldsymbol{\beta}}_1^T \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ 0 & I - \boldsymbol{\beta}_0^0(\boldsymbol{\beta}_0^0)^T & 0 \\ 0 & 0 & I - \boldsymbol{\beta}_1^0(\boldsymbol{\beta}_1^0)^T \end{pmatrix} + o_p(1).$$

We have, by the constraints $\|\boldsymbol{\beta}_0\|_2 = 1$ and $\|\boldsymbol{\beta}_1\|_2 = 1$,

$$n^{-1}P_{\boldsymbol{\beta}}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\} \begin{pmatrix} \tilde{Z}_i \\ X_{\widehat{m},i} \end{pmatrix} \begin{pmatrix} \tilde{Z}_i - \tilde{\mathbf{P}}(\mathbf{Z}_i) \\ X_{\widehat{m},i} - \tilde{\mathbf{P}}(\mathbf{X}_i) \end{pmatrix}^T \{(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) + o_p(\|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|_2)\}$$

$$= n^{-1}P_{\boldsymbol{\beta}}\sum_{i=1}^{n} \begin{pmatrix} \tilde{Z}_i - \tilde{\mathbf{P}}(\mathbf{Z}_i) \\ X_{m,i} - \tilde{\mathbf{P}}(\mathbf{X}_i) \end{pmatrix} q_1\{\tilde{\eta}_i\}\{1 + O_p(J_n^{1-r})\} + o_p(n^{-1/2}).$$

Because $\|X_{\widehat{m},i} - X_{m,i}\| = O_p(\sqrt{N^3/n} + N^{1-r})$, we have, by the law of large numbers,

$$n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\} \begin{pmatrix} \tilde{Z}_i \\ X_{\widehat{m},i} \end{pmatrix} \begin{pmatrix} \tilde{Z}_i - \tilde{\mathbf{P}}(\mathbf{Z}_i) \\ X_{m,i} - \tilde{\mathbf{P}}(\mathbf{X}_i) \end{pmatrix}^T = E[\rho_2\{\eta(\mathbf{V}, \boldsymbol{\theta}^0)\}\boldsymbol{\phi}(\mathbf{V}, \boldsymbol{\beta})^{\otimes 2}] + O_p(\sqrt{N^3/n} + N^{1-r}) + O_p(n^{-2}).$$

This completes the proof of Lemma S.6.

□

Proof of Theorem 1: Due to independence of the observations $\mathbf{V}_1, \cdots, \mathbf{V}_n$, we have, by Lindeberg-

Feller central limit theorem,

$$n^{-1/2}\sum_{i=1}^{n}\begin{pmatrix}\tilde{Z}_i-\tilde{\mathbf{P}}(\mathbf{Z}_i)\\ X_{m,i}-\tilde{\mathbf{P}}(\mathbf{X}_i)\end{pmatrix}q_1\{\tilde{\eta}_i\}\xrightarrow{\mathcal{L}}N(\mathbf{0},\tilde{\Sigma}),$$

where $\Sigma = E\left[\rho_2\{\eta(\mathbf{V},\boldsymbol{\theta}^0)\}\boldsymbol{\phi}(\mathbf{V},\boldsymbol{\beta}^0)^{\otimes 2}\right]$. Then combining Lemma S.6 and Slusky's theorem, we can complete the proof of Theorem 1. □

Proof of Theorem 2: Because for any $u_l\in[a_l,b_l]$, $B_{s,l}(u_l)$, $s=1,\cdots,J_n, l=0,1$, have the banded first derivatives, by (S.9) and Theorem 1, we have, for any $u_l\in[a_l,b_l]$,

$$\begin{aligned}|\tilde{m}_l(u_l,\widehat{\boldsymbol{\beta}})-\tilde{m}_l(u_l,\boldsymbol{\beta}^0)| =&|D(\widehat{\boldsymbol{\beta}})^T\widehat{\lambda}(\widehat{\boldsymbol{\beta}})-D(\boldsymbol{\beta}^0)^T\lambda(\boldsymbol{\beta}^0)|\\ \le&|D(\boldsymbol{\beta}^0)^T\{\widehat{\lambda}(\widehat{\boldsymbol{\beta}})-\lambda(\boldsymbol{\beta}^0)\}|+|\{D(\widehat{\boldsymbol{\beta}})-D(\boldsymbol{\beta}^0)\}^T\widehat{\lambda}(\widehat{\boldsymbol{\beta}})|\\ \le&|n^{-1}D(\boldsymbol{\beta}^0)^T\widehat{\mathbf{U}}(\boldsymbol{\beta}^0)^{-1}\mathbf{D}(\boldsymbol{\beta}^0)^T\mathbf{q_1}\{\tilde{\eta}\}|+O_p(n^{-1/2})\\ =&O_p(\sqrt{N/n}).\end{aligned}$$

Then, we have, together with Lemma S.4,

$$\begin{aligned}\sup_{u_l\in[a_l,b_l]}|\tilde{m}_l(u_l,\widehat{\boldsymbol{\beta}})-m_l(u_l)|\le&\sup_{u_l\in[a_l,b_l]}|\tilde{m}_l(u_l,\widehat{\boldsymbol{\beta}})-\tilde{m}_l(u_l,\boldsymbol{\beta}^0)|+\sup_{u_l\in[a_l,b_l]}|\tilde{m}_l(u_l,\boldsymbol{\beta}^0)-m_l(u_l)|\\ =&O_p(\sqrt{N/n}+N^{-r}).\end{aligned}$$

This completes the proof of Theorem 2. □

Lemma S.7 *If assumptions* (A1)-(A7) *hold, and* $nN^{-4}\to\infty$ *and* $nN^{-2r-2}\to 0$, *as* $n\to\infty$. *We have*

$$\sup_{u_l\in[a_l,b_l]}\left|\widehat{m}_l^O(u_l,\widehat{\boldsymbol{\beta}})-m_l(u_l)\right| = O_p(n^{-2/5}\sqrt{\log(n)}),$$

and for any $u_l\in[a_l,b_l]$,

$$\sqrt{nh_l}\left\{\widehat{m}_l^O(u_l,\widehat{\boldsymbol{\beta}})-m_l(u_l)-b_l(u_l)h_l^2\right\}\xrightarrow{\mathcal{L}}N(0,v_l(u_l)),$$

where $b_l(u_l) = \mu_1 m_l''(u_l)/2,\ v_l(u_l) = \{E[\tilde{G}_l^2|u_l]\}^{-2}\|K\|_2^2 E[\tilde{G}_l^2\sigma^2(\mathbf{v})|u_l]f_l(u_l)^{-1},\ \tilde{G}_0 = 1,$ *and* $\tilde{G}_1 = G.$

Proof of Lemma S.7: Let

$$\tilde{\mathbf{X}}(\boldsymbol{\beta}_1) \equiv \tilde{\mathbf{X}}(u_1, \boldsymbol{\beta}_1) = \begin{pmatrix} G_1 & \cdots & G_n \\ (\boldsymbol{\beta}_1^T \mathbf{X}_1 - u_1)G_i/h_1 & \cdots & (\boldsymbol{\beta}_1^T \mathbf{X}_n - u_1)G_n/h_1 \end{pmatrix}^T,$$

$$\tilde{\mathbf{W}}_{q_j}^O(\boldsymbol{\theta}) \equiv \tilde{\mathbf{W}}_{q_j}^O(u_1, \boldsymbol{\theta}) = \mathrm{diag}\left\{q_j\{\eta_{-1,1}^O\}K_{h_1}(\boldsymbol{\beta}_1^T \mathbf{X}_1 - u_1), \cdots, q_j\{\eta_{-1,n}^O\}K_{h_1}(\boldsymbol{\beta}_1^T \mathbf{X}_n - u_1)\right\},$$

where $q_j\{\eta_{-1,i}^O\} = q_j\{\eta_{-1}^O(\mathbf{V}_i; a, b, \boldsymbol{\theta})\}$ with $\eta_{-1}^O(\mathbf{V}_i; a, b, \boldsymbol{\theta}) = \boldsymbol{\alpha}_0^T \mathbf{Z}_i + m_0(\boldsymbol{\beta}_0^T \mathbf{X}_i) + \boldsymbol{\alpha}_1^T \mathbf{Z}_i + aG_i + b(\boldsymbol{\beta}_1^T \mathbf{X}_i - u_1)G_i$ and $j = 1, 2,\ i = 1, \cdots, n$. The estimate $(\widehat{a}^O, \widehat{b}^O)$ solves equation

$$0 = n^{-1}\sum_{i=1}^n q_1\{\widehat{\eta}_{-1}^O(\mathbf{V}_i; \widehat{a}^O, \widehat{b}^O, \widehat{\boldsymbol{\theta}})\} \begin{pmatrix} G_i \\ (\widehat{\boldsymbol{\beta}}_1^T \mathbf{X}_i - u_1)G_i/h_1 \end{pmatrix} K_{h_1}(\widehat{\boldsymbol{\beta}}_1^T \mathbf{X}_i - u_1).$$

Thus,

$$\widehat{m}_1^O(u_1, \widehat{\boldsymbol{\beta}}) = (1, 0)\{\tilde{\mathbf{X}}(\widehat{\boldsymbol{\beta}}_1)^T \tilde{\mathbf{W}}_{q_2}^O(\widehat{\boldsymbol{\theta}})\tilde{\mathbf{X}}(\widehat{\boldsymbol{\beta}}_1)\}^{-1}\tilde{\mathbf{X}}(\widehat{\boldsymbol{\beta}}_1)^T \tilde{\mathbf{W}}_{q_1}^O(\widehat{\boldsymbol{\theta}})(1 + o_p(1)).$$

Because $\|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\| = O_p(n^{-1/2})$ by Theorem 1, we can prove that, for any $u_l \in [a_l, b_l]$, $l = 0, 1$, $\tilde{\mathbf{X}}(\widehat{\boldsymbol{\beta}}_1) = \tilde{\mathbf{X}}(\boldsymbol{\beta}_1^0) + O_p(n^{-1/2})$ and $\tilde{\mathbf{W}}_{q_j}^O(\widehat{\boldsymbol{\theta}}) = \tilde{\mathbf{W}}_{q_j}^O(\boldsymbol{\theta}^0) + O_p(n^{-1/2})$, which implies that

$$\sup_{u_l \in [a_l, b_l]} \left|\widehat{m}_l^O(u_l, \widehat{\boldsymbol{\beta}}) - \widehat{m}_l^O(u_l, \boldsymbol{\beta}^0)\right| = O_p(n^{-1/2}). \tag{S.9}$$

Along the lines of Theorem 2.5 and 2.6 in Li and Racine (2007), we have

$$\sup_{u_l \in [a_l, b_l]} \left|\widehat{m}_l^O(u_l, \boldsymbol{\beta}^0) - m_l(u_l)\right| = O_p(n^{-2/5}\sqrt{\log(n)}),$$

which implies that

$$\begin{aligned}\sup_{u_l\in[a_l,b_l]}\left|\widehat{m}_l^O(u_l,\widehat{\boldsymbol{\beta}})-m_l(u_l)\right|\leq&\sup_{u_l\in[a_l,b_l]}\left|\widehat{m}_l^O(u_l,\widehat{\boldsymbol{\beta}})-\widehat{m}_l^O(u_l,\boldsymbol{\beta}^0)\right|+\sup_{u_l\in[a_l,b_l]}\left|\widehat{m}_l^O(u_l,\boldsymbol{\beta}^0)-m_l(u_l)\right|\\ =&O_p(n^{-1/2})+O_p(n^{-2/5}\sqrt{\log(n)})\\ =&O_p(n^{-2/5}\sqrt{\log(n)}),\end{aligned}$$

Therefore, this arrives the first part of Lemma S.7, and

$$\sqrt{nh_l}\left\{\widehat{m}_l^O(u_l,\boldsymbol{\beta}^0)-m_l(u_l)-b_l(u_l)h_l^2\right\}\xrightarrow{\mathcal{L}}N(0,v_l(u_l)).$$

By combining (S.9) and the asymptotic normality of $\widehat{m}_l^O(u_1,\boldsymbol{\beta}^0)$, we can prove the second part.
□

Let $\widehat{\ell}(a,b)=\sum_{i=1}^n Q(g^{-1}\{\widehat{\eta}_{-1}(\mathbf{V}_i;a,b)\},Y_i)K_{h_1}(\widehat{\boldsymbol{\beta}}_1^T\mathbf{X}_i-u_1)$, where $\widehat{\eta}_{-1}(\mathbf{V}_i;a,b)=\widehat{\boldsymbol{\alpha}}_1^T\mathbf{Z}_i+\tilde{m}_0(\widehat{\boldsymbol{\beta}}_0^T\mathbf{X}_i,\widehat{\boldsymbol{\beta}})+aG_i+b(\widehat{\boldsymbol{\beta}}_1^T\mathbf{X}_i-u_1)G_i$. Let $\tilde{\ell}(a,b)=\sum_{i=1}^n Q(g^{-1}\{\widehat{\eta}_{-1}^O(\mathbf{V}_i;a,b)\},Y_i)K_{h_1}(\widehat{\boldsymbol{\beta}}_1^T\mathbf{X}_i-u_1)$, where $\widehat{\eta}_{-1}^O(\mathbf{V}_i;a,b)=\widehat{\boldsymbol{\alpha}}_1^T\mathbf{Z}_i+m_0(\widehat{\boldsymbol{\beta}}_0^T\mathbf{X}_i,\widehat{\boldsymbol{\beta}})+aG_i+b(\widehat{\boldsymbol{\beta}}_1^T\mathbf{X}_i-u_1)G_i$

Lemma S.8 *Suppose that assumptions* (A.1)-(A.7) *in the Appendix hold, and* $nN^{-4}\to\infty$ *and* $nN^{-\delta}\to 0$ *with* $\delta=\min(2r+2,5r/2)$, *then we have*

$$\sup_{u_1\in[a_1,b_1]}|\widehat{\ell}'(\widehat{a}^O,\widehat{b}^O)|=\boldsymbol{O}_{a.s.}(n^{-1/2}\log n).$$

Proof of Lemma S.8: We only prove the case of $l=1$, that is, consider $\widehat{a}^O=\widehat{\beta}_1(u)$ and $\widehat{b}^O=\widehat{\beta}_1'(u)$ in oracle case. Recall $\tilde{\ell}'(\widehat{a}^O,\widehat{b}^O)=0$. We have

$$\begin{aligned}\widehat{\ell}'(\widehat{a}^O,\widehat{b}^O)=&\widehat{\ell}'(\widehat{a}^O,\widehat{b}^O)-\tilde{\ell}'(\widehat{a}^O,\widehat{b}^O)\\ =&\frac{1}{n}\sum_{i=1}^n\left[q_1\{\widehat{\eta}_{-1}(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}-q_1\{\widehat{\eta}_{-1}^O(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}\right]X_{i,1}K_{h_1}(u_i-u)\\ =&\frac{1}{n}\sum_{i=1}^n q_2\{\widehat{\eta}_{-1}^O(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}\left[\tilde{\boldsymbol{\beta}}_{-1}(u_i)-\boldsymbol{\beta}_{-1}(u_i)\right]^T\mathbf{X}_{i,-1}X_{i,1}K_{h_1}(u_i-u)\\ &+\boldsymbol{O}_{\text{a.s.}}(\frac{1}{n}\sum_{i=1}^n\sum_{l=2}^p(\tilde{\beta}_1(u_i)-\beta_1(u_i))^2)\end{aligned}$$

By Lemma S.5, we have $\frac{1}{n}\sum_{i=1}^{n}\sum_{l=2}^{p}(\tilde{\beta}_l(u_i)-\beta_l(u_i))^2 = \boldsymbol{O}_{\text{a.s.}}(N/n+N^{-2r})$. Let $\bar{\boldsymbol{\beta}}_{-1}(u) = (\bar{\beta}_2(u),\cdots,\bar{\beta}_p(u))^T$ with $\bar{\beta}_2(u)$ being defined in Lemma S.5. Thus, we can rewrite

$$\widehat{\ell}'(\widehat{a}^O,\widehat{b}^O) = A_n + B_n + \boldsymbol{O}_{\text{a.s.}}(N/n+N^{-2r}),$$

where

$$\begin{aligned}
A_n =& \frac{1}{n}\sum_{i=1}^{n} q_2\{\widehat{\eta}_{-1}^O(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}\left[\tilde{\boldsymbol{\beta}}_{-1}(u_i)-\bar{\boldsymbol{\beta}}_{-1}(u_i)\right]^T \mathbf{X}_{i,-1}X_{i,1}K_{h_1}(u_i-u),\\
B_n =& \frac{1}{n}\sum_{i=1}^{n} q_2\{\widehat{\eta}_{-1}^O(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}\left[\bar{\boldsymbol{\beta}}_{-1}(u_i)-\boldsymbol{\beta}_{-1}(u_i)\right]^T \mathbf{X}_{i,-1}X_{i,1}K_{h_1}(u_i-u).
\end{aligned}$$

Noting that $q_2(\cdot)$ is bounded, by Lemma S.5, we have

$$\begin{aligned}
B_n \leq& C\sum_{l=2}^{p}|\bar{\boldsymbol{\beta}}_{-1}(u_i)-\boldsymbol{\beta}_{-1}(u_i)|\|\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_{i,-1}X_{i,1}K_{h_1}(u_i-u)\|\\
=& \boldsymbol{O}_{\text{a.s.}}(\sqrt{N/n}+N^{-r}).
\end{aligned}$$

Define $\boldsymbol{\Phi}_b = (\phi_{b,1}^T,\cdots,\phi_{b,p}^T)^T$, $\boldsymbol{\Phi}_v = (\phi_{v,1}^T,\cdots,\phi_{v,p}^T)^T$ and $\boldsymbol{\Phi}_r = (\phi_{r,1}^T,\cdots,\phi_{r,p}^T)^T$, where

$$\begin{aligned}
\boldsymbol{\Phi}_b =& \mathbf{U}^{-1}\frac{1}{n}\sum_{i=1}^{n}\left[q_1\{\tilde{\eta}(\mathbf{V}_i;\widehat{\boldsymbol{\alpha}},\lambda)\}-q_1\{\eta(\mathbf{V}_i,\alpha)\}\right]D_i,\\
\boldsymbol{\Phi}_v =& \mathbf{U}^{-1}\frac{1}{n}\sum_{i=1}^{n}q_1\{\eta(\mathbf{V}_i,\alpha)\}D_i,\\
\boldsymbol{\Phi}_r =& \widehat{\lambda}-\lambda-\boldsymbol{\Phi}_b-\boldsymbol{\Phi}_v.
\end{aligned}$$

Along the line of proof in Liu eta al. (2013), and by Lemma S.1 and Lemma S.2, we can prove that $\|\boldsymbol{\Phi}_b\|_2 = \boldsymbol{O}_{\text{a.s.}}(N^{-r})$ and $\|\boldsymbol{\Phi}_r\|_2 = \boldsymbol{O}_{\text{a.s.}}(N^{3/2}n^{-1})$. We can decompose $A_n = A_{1n}+A_{2n}+A_{2n}$, where

$$
\begin{aligned}
A_{1n} =& \frac{1}{n}\sum_{i=1}^{n} q_2\{\widehat{\eta}_{-1}^{O}(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}(\phi_{b,l}^T\mathbf{B}_r(u_i), l=2,\cdots,p)^T\mathbf{X}_{i,-1}X_{i,1}K_{h_1}(u_i-u)\\
A_{2n} =& \frac{1}{n}\sum_{i=1}^{n} q_2\{\widehat{\eta}_{-1}^{O}(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}(\phi_{v,l}^T\mathbf{B}_r(u_i), l=2,\cdots,p)^T\mathbf{X}_{i,-1}X_{i,1}K_{h_1}(u_i-u)\\
A_{3n} =& \frac{1}{n}\sum_{i=1}^{n} q_2\{\widehat{\eta}_{-1}^{O}(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}(\phi_{r,l}^T\mathbf{B}_r(u_i), l=2,\cdots,p)^T\mathbf{X}_{i,-1}X_{i,1}K_{h_1}(u_i-u).
\end{aligned}
$$

Applying Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
\sup_{u_1\in[a_1,b_1]}|A_{1n}| \leq& C\sup_{u_1\in[a_1,b_1]}\frac{1}{n}\sum_{i=1}^{n}|(\phi_{b,l}^T\mathbf{B}_r(u_i), l=2,\cdots,p)^T\mathbf{X}_{i,-1}X_{i,1}K_{h_1}(u_i-u)|\\
\leq& C\sup_{u_1\in[a_1,b_1]}\|\boldsymbol{\Phi}_b\|_2\left[\sum_{2\leq l\leq p,1\leq s\leq J_n}[\frac{1}{n}\sum_{i=1}^{n}|B_{s,r}(u_i)X_{i,l}X_{i,1}|K_{h_1}(u_i-u)]^2\right]^{1/2}\\
=& C\sup_{u_1\in[a_1,b_1]}\|\boldsymbol{\Phi}_b\|_2 J_n\boldsymbol{O}\text{a.s.}(J_n^{-1})\\
=& \boldsymbol{O}\text{a.s.}(N^{-r}),
\end{aligned}
$$

and similarly, $\sup_{u_1\in[a_1,b_1]}|A_{3n}| = \boldsymbol{O}\text{a.s.}(N^{3/2}n^{-1})$. Let $\boldsymbol{\zeta}_{l,1}(u) = (\boldsymbol{\zeta}_{1,l,1},\cdots,\boldsymbol{\zeta}_{J_n,l,1})^T$ with

$$
\boldsymbol{\zeta}_{s,l,1}(u) = n^{-1}\sum_{i=1}^{n} q_2\{\widehat{\eta}_{-1}^{O}(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}X_{i,1}K_{h_1}(u_i-u)B_{s,r}(u_i)^TX_{i,l},\quad l=2,\cdots,p.
$$

By Lemma A.2 in Xia and Li (1999), we have

$$
\sup_{u_1\in[a_1,b_1]}|\boldsymbol{\zeta}_{s,l,1}(u) - E\{\boldsymbol{\zeta}_{s,l,1}(u)\}| = \boldsymbol{O}\text{a.s.}(\sqrt{\log n/(nh)}).
$$

By de Boor (2001), we have, for any $u_1\in[a_1,b_1]$,

$$
\begin{aligned}
|E\{\boldsymbol{\zeta}_{s,l,1}(u)\}| =& |E\{K_{h_1}(u_i-u)E[q_2\{\widehat{\eta}_{-1}^{O}(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}X_{i,1}B_{s,r}(u_i)^TX_{i,l}|u_i]\}|\\
\leq& \mu_0 f(u)|E\{X_{i,1}X_{i,2}|u_i=u\}||E[B_{s,r}(u_i)]| + O(h_2^2)O(|E[B_r(u_i)]|)\\
=& O(J_n^{-1}),
\end{aligned}
$$

which implies that $|\boldsymbol{\zeta}_{s,l,1}(u)| = O(J_n^{-1})$. It is easy can be shown further that $\sup_{u_1\in[a_1,b_1]}\max_{1\le s\le J_n,1\le l\le p}|\boldsymbol{\zeta}_{s,l,1}(u)| = \boldsymbol{O}_{\text{a.s.}}(J_n^{-1/2})$. Define $\boldsymbol{\Phi}_{v,-1} = (\phi_{v,2}^T,\cdots,\phi_{v,p}^T)^T$. By the definition of A_{2n}, we have $A_{2n} = \boldsymbol{\zeta}_1(u)^T\boldsymbol{\Phi}_{v,-1}$, which implies that

$$\begin{aligned}
&\sup_{u_1\in[a_1,b_1]} E\left(A_{2n}|\mathbf{X},\mathbf{Z},U\right)^2\\
&\qquad= \sup_{u_1\in[a_1,b_1]} E\left(\boldsymbol{\zeta}_1(u)^T\boldsymbol{\Phi}_{v,-1}\boldsymbol{\Phi}_{v,-1}^T\boldsymbol{\zeta}_1(u)|\mathbf{X},\mathbf{Z},U\right)\\
&\qquad= \sup_{u_1\in[a_1,b_1]} \boldsymbol{\zeta}_1(u)^T E\left(\boldsymbol{\Phi}_{v,-1}\boldsymbol{\Phi}_{v,-1}^T|\mathbf{X},\mathbf{Z},U\right)\boldsymbol{\zeta}_1(u)\\
&\qquad\le \sup_{u_1\in[a_1,b_1]} \|\boldsymbol{\zeta}_1(u)\|_2^2\|E\left(\boldsymbol{\Phi}_v\boldsymbol{\Phi}_v^T|\mathbf{X},\mathbf{Z},U\right)\|_\infty)\\
&\qquad= \sup_{u_1\in[a_1,b_1]} n^{-2}\|\boldsymbol{\zeta}_1(u)\|_2^2\|\widehat{\mathbf{U}}^{-1}\mathbf{D}^T E\left(\mathbf{q}_2\{\widehat{\eta}_{-1}^O(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}^{\otimes 2}|\mathbf{X},\mathbf{Z},U\right)\mathbf{D}\widehat{\mathbf{U}}^{-1}\|_2\\
&\qquad\le C_{q_2}\sup_{u_1\in[a_1,b_1]} n^{-1}\|\boldsymbol{\zeta}_1(u)\|_2^2\|\widehat{\mathbf{U}}^{-1}\|_2\\
&\qquad= \boldsymbol{O}_{\text{a.s.}}(n^{-1}).
\end{aligned}$$

Therefore, by the law of large numbers, we have $\sup_{u_1\in[a_1,b_1]}|A_{2n}| = \boldsymbol{O}_{\text{a.s.}}(n^{-1/2}\log n)$.

□

Lemma S.9 *If assumptions* (A.1)-(A.7) *in the Appendix hold, and* $nN^4\to\infty$ *and* $nN^{-\delta}\to 0$ *with* $\delta = \min(2r+2, 5r/2)$,

$$\sup_{u_1\in[a_1,b_1]}\|\widehat{\ell}''(a,b)\|_2 = \boldsymbol{O}_{a.s.}(1),$$

for any $(a,b)\in\mathcal{A}$.

Proof of Lemma S.9: Let $\tilde{\mathbf{X}}_{i1}(u) = (X_{i,1},(U_i-u)X_{i,1}/h_1)^T$. According to the boundedness of

$q_2(\cdot)$ and the law of large numbers, we have

$$\begin{aligned}\widehat{\ell}''(a,b) =& \frac{1}{n}\sum_{i=1}^{n} q_2\{\widehat{\eta}_{-1}(\mathbf{V}_i;a,b,\widehat{\boldsymbol{\alpha}})\}\tilde{\mathbf{X}}_{i1}(u)^{\otimes 2}K_{h_1}(u_i-u)\\ =& E\left[q_2\{\widehat{\eta}_{-1}(\mathbf{V}_i;a,b,\widehat{\boldsymbol{\alpha}})\}\tilde{\mathbf{X}}_{i1}(u)^{\otimes 2}K_{h_1}(u_i-u)\right]+\boldsymbol{O}_{\text{a.s.}}(1)\\ =& \boldsymbol{O}_{\text{a.s.}}(1).\end{aligned}$$

□

Lemma S.10 *Suppose that assumptions* (A.1)-(A.7) *in the Appendix hold, and* $nN^4 \to \infty$ *and* $nN^{-\delta} \to 0$ *with* $\delta = \min(2r+2, 5r/2)$*, then we have, for* $l = 1, \cdots, p$*,*

$$\sup_{u_1\in[a_1,b_1]} |\widehat{\beta}_l(u,\widehat{\boldsymbol{\alpha}}) - \widehat{\beta}_l^O(u,\widehat{\boldsymbol{\alpha}})| = \boldsymbol{O}_{a.s.}(n^{-1/2}\log n),$$

where $\widehat{\beta}_l(u,\widehat{\boldsymbol{\alpha}})$ *and* $\widehat{\beta}_l^O(u,\widehat{\boldsymbol{\alpha}})$ *denotes* $\widehat{\beta}_l(u)$ *and* $\widehat{\beta}_l^O(u)$*, respectively.*

Proof of Lemma S.10: Noting that $\widehat{\ell}'(\widehat{a},\widehat{b}) = 0$, by the mean value theorem, there exists a $(\bar{a},\bar{b}) \in \mathcal{A}$ between $(\widehat{a},\widehat{b})$ and $(\widehat{a}^O,\widehat{b}^O)$ such that

$$-\widehat{\ell}'(\widehat{a}^O,\widehat{b}^O) = \widehat{\ell}'(\widehat{a},\widehat{b}) - \widehat{\ell}'(\widehat{a}^O,\widehat{b}^O) = \widehat{\ell}''(\bar{a},\bar{b})\left[(\widehat{a},\widehat{b})^T - (\widehat{a}^O,\widehat{b}^O)^T\right],$$

which follows by

$$(\widehat{a},\widehat{b})^T - (\widehat{a}^O,\widehat{b}^O)^T = -\widehat{\ell}''(\bar{a},\bar{b})^{-1}\widehat{\ell}'(\widehat{a}^O,\widehat{b}^O).$$

Combining Lemma S.8 and Lemma S.9, we have $(\widehat{a},\widehat{b})^T - (\widehat{a}^O,\widehat{b}^O)^T = \boldsymbol{O}_{\text{a.s.}}(n^{-1}\log n)$, which proves Lemma S.10 by noting $\widehat{\beta}_l(u,\widehat{\boldsymbol{\alpha}}) - \widehat{\beta}_l^O(u,\widehat{\boldsymbol{\alpha}}) = (1,0)\left[(\widehat{a},\widehat{b})^T - (\widehat{a}^O,\widehat{b}^O)^T\right]$.

□

Proof of Theorem 4: Due to $nh_l^5 = O(1)$, we have $\sqrt{nh_l}n^{-2/5} = o_o(1)$. By Theorem 3, we have

$$\sqrt{nh_l}\left\{\widehat{m}_l(u_l,\widehat{\boldsymbol{\beta}}) - m_l(u_l) - b_l(u_l)h_l^2\right\} = \sqrt{nh_l}\left\{\widehat{m}_l^O(u_l,\widehat{\boldsymbol{\beta}}) - m_l(u_l) - b_l(u_l)h_l^2\right\} + o_p(1).$$

Thus Theorem 4 can be shown straightforwardly by Lemma S.7. □

Assume $m_0(u_0)$ be known, similar to Liu et al. (2016). Let $\widehat{m}^O_{1,H_0}(u_1)$ and $\widehat{m}^O_{1,H_1}(u_1)$ be the estimates under H_0 and H_1 in hypothesis test (9), respectively. The log-likelihoods under H_0 and H_1 are

$$\ell^O_{n1}(H_0) = \sum_{i=1}^{n} Q(g^{-1}\{\widehat{\eta}^O_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\}, Y_i) \text{ and } \ell^O_{n1}(H_1) = \sum_{i=1}^{n} Q(g^{-1}\{\widehat{\eta}^O_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\}, Y_i),$$

where

$$\begin{aligned}\widehat{\eta}^O_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) =& \widehat{\boldsymbol{\alpha}}^T\tilde{\mathbf{Z}}_i + m_0(\widehat{\boldsymbol{\beta}}_0^T\mathbf{X}_i) + \widehat{m}^O_{1,H_0}(\widehat{\boldsymbol{\beta}}_1^T\mathbf{X}_i)G_i,\\ \widehat{\eta}^O_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) =& \widehat{\boldsymbol{\alpha}}^T\tilde{\mathbf{Z}}_i + m_0(\widehat{\boldsymbol{\beta}}_0^T\mathbf{X}_i) + \widehat{m}^O_{1,H_1}(\widehat{\boldsymbol{\beta}}_1^T\mathbf{X}_i)G_i.\end{aligned}$$

When $m_0(u_0)$ is known, we can construct GLR statistic for hypothesis in (9) ,

$$T_1^O = 2(\ell^O_{n1}(H_1) - \ell^O_{n1}(H_0)). \tag{S.10}$$

Lemma S.11 *If assumptions* (A.1)-(A.7) *hold, and* $nN^4 \to \infty$ *and* $nN^{-2r-2} \to 0$*, then under* H_0 *in* (9)*, when* $m_0(u_0)$ *is known, and* $m_{01}(u_1)$ *is a linear function,*

$$\sigma_{2n}^{-1}(T_1^O - \mu_{2n}) \xrightarrow{\mathcal{L}} N(0,1),$$

where $\sigma^2_{2n} = 2|\Omega_2| \int \left\{K(u) - 1/2K * K^2(u)\right\}^2 du/h_1$ *and* $\mu_{2n} = |\Omega_2| \left\{K(0) - \frac{1}{2}\int K^2(u)du\right\}/h_1$*. Furthermore,* $a_K T_1^O \overset{a}{\sim} \chi^2_{d_2}$*, where* $d_2 = a_K\mu_{2n}$*.*

Proof of Lemma S.11: Let

$$\begin{aligned}\ell^*_{n1}(H_0) =& \sum_{i=1}^{n} Q(g^{-1}\{\widehat{\eta}^*_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\}, Y_i)\\ \ell^*_{n1}(H_1) =& \sum_{i=1}^{n} Q(g^{-1}\{\widehat{\eta}^*_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\}, Y_i),\\ T_1^* =& 2(\ell^*_{n1}(H_1) - \ell^*_{n1}(H_0)),\end{aligned}$$

where

$$\begin{aligned}\widehat{\eta}^*_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) =& \tilde{\mathbf{Z}}_i^T\boldsymbol{\alpha}^0 + m_0(\mathbf{X}_i^T\boldsymbol{\beta}_0^0) + \widehat{m}^O_{1,H_0}(\mathbf{X}_i^T\boldsymbol{\beta}_1^0)G_i,\\ \widehat{\eta}^*_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) =& \tilde{\mathbf{Z}}_i\boldsymbol{\alpha}^0 + m_0(\mathbf{X}_i^T\boldsymbol{\beta}_0^0) + \widehat{m}^O_{1,H_1}(\mathbf{X}_i^T\boldsymbol{\beta}_1^0)G_i,\end{aligned}$$

Then we have

$$\begin{aligned}\ell^O_{n1}(H_0) - \ell^*_{n1}(H_0) =& \sum_{i=1}^n q_1(\widehat{\eta}^*_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}^O_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}^*_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\}\\ &- 1/2\sum_{i=1}^n q_2(\widehat{\eta}^*_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}^O_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}^*_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\}^2\\ =& O_p(1),\end{aligned}$$

and

$$\begin{aligned}\ell^O_{n1}(H_1) - \ell^*_{n1}(H_1) =& \sum_{i=1}^n q_1(\widehat{\eta}^*_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}^O_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}^*_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\}\\ &- 1/2\sum_{i=1}^n q_2(\widehat{\eta}^*_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}^O_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}^*_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\}^2\\ =& O_p(1),\end{aligned}$$

This results in $\ell^O_{n1}(H_1) - \ell^O_{n1}(H_0) = \ell^*_{n1}(H_1) - \ell^*_{n1}(H_0) + O_p(1)$. So we have $T_1^O = T_1^* + O_p(1)$. By the similar arguments of Fan et al. (2001), we can prove $\sigma_{2n}^{-1}(T_1^* - \mu_{2n}) \xrightarrow{\mathcal{L}} N(0,1)$, which implies the results of Lemma S.11. □

Proof of Theorem 5: As the same proof of Lemma S.11, we have

$$\begin{aligned}\ell_{n1}(H_0) - \ell^O_{n1}(H_0) =& \sum_{i=1}^n q_1(\widehat{\eta}^O_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}^O_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\}\\ &- 1/2\sum_{i=1}^n q_2(\widehat{\eta}^O_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}^O_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\}^2\\ =& O_p(1),\end{aligned}$$

and

$$
\begin{aligned}
&\ell_{n1}(H_1) - \ell_{n1}^O(H_1) \\
=&\sum_{i=1}^n q_1(\widehat{\eta}_{H_1}^O(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}_{H_1}^O(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\} \\
&- 1/2\sum_{i=1}^n q_2(\widehat{\eta}_{H_1}^O(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}_{H_1}^O(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\}^2 \\
&= O_p(1).
\end{aligned}
$$

Therefore we can show $\ell_{n1}(H_1) - \ell_{n1}(H_0) = \ell_{n1}^O(H_1) - \ell_{n1}^O(H_0) + O_p(1)$, and furthermore $T_1 = T_1^O + o(1)$, which implies that T_1 and T_1^O have the same distribution. This completes the proof of Theorem 5 by Lemma S.11. □

Proof of Theorem 6: Let

$$
\begin{aligned}
\ell_{n2}^O(H_0) =& \sum_{i=1}^n Q(g^{-1}\{\widehat{\eta}_{2,H_0}^O(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\}, Y_i), \\
\ell_{n2}^O(H_1) =& \sum_{i=1}^n Q(g^{-1}\{\widehat{\eta}_{2,H_1}^O(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\}, Y_i), \\
\ell_{n2}^*(H_0) =& \sum_{i=1}^n Q(g^{-1}\{\widehat{\eta}_{2,H_0}^*(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\}, Y_i) \\
\ell_{n2}^*(H_1) =& \sum_{i=1}^n Q(g^{-1}\{\widehat{\eta}_{2,H_1}^*(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\}, Y_i), \\
T_2^O =& 2(\ell_{n2}^O(H_1) - \ell_{n2}^O(H_0)), \\
T_2^* =& 2(\ell_{n2}^*(H_1) - \ell_{n2}^*(H_0)),
\end{aligned}
$$

where $\tilde{\mathbf{Z}}_i = (\mathbf{Z}_i^T, \mathbf{Z}_i^T G_{i1}, \mathbf{Z}_i^T G_{i2})^T$ and

$$
\begin{aligned}
\widehat{\eta}_{2,H_0}^O(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) =& \tilde{\mathbf{Z}}_i^T\widehat{\boldsymbol{\alpha}} + m_0(\mathbf{X}_i^T\widehat{\boldsymbol{\beta}}_0) + \widehat{m}_{1,H_0}^O(\mathbf{X}_i^T\widehat{\boldsymbol{\beta}}_1)G_{i1} + \widehat{m}_{2,H_0}^O(\mathbf{X}_i^T\widehat{\boldsymbol{\beta}}_2)G_{i2}, \\
\widehat{\eta}_{2,H_1}^O(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) =& \tilde{\mathbf{Z}}_i^T\widehat{\boldsymbol{\alpha}} + m_0(\mathbf{X}_i^T\widehat{\boldsymbol{\beta}}_0) + \widehat{m}_{1,H_1}^O(\mathbf{X}_i^T\widehat{\boldsymbol{\beta}}_1)G_{i1} + \widehat{m}_{2,H_1}^O(\mathbf{X}_i^T\widehat{\boldsymbol{\beta}}_2)G_{i2}, \\
\widehat{\eta}_{2,H_0}^*(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) =& \tilde{\mathbf{Z}}_i^T\boldsymbol{\alpha}^0 + m_0(\mathbf{X}_i^T\boldsymbol{\beta}_0^0) + \widehat{m}_{1,H_0}^O(\mathbf{X}_i^T\boldsymbol{\beta}_1^0)G_{i1} + \widehat{m}_{2,H_0}^O(\mathbf{X}_i^T\boldsymbol{\beta}_2^0)G_{i2}, \\
\widehat{\eta}_{2,H_1}^*(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) =& \tilde{\mathbf{Z}}_i\boldsymbol{\alpha}^0 + m_0(\mathbf{X}_i^T\boldsymbol{\beta}_0^0) + \widehat{m}_{1,H_1}^O(\mathbf{X}_i^T\boldsymbol{\beta}_1^0)G_{i1} + \widehat{m}_{2,H_1}^O(\mathbf{X}_i^T\boldsymbol{\beta}_2^0)G_{i2}.
\end{aligned}
$$

By the same arguments of the proof of Lemma S.11, we can prove that $\ell^O_{n2}(H_1) - \ell^O_{n2}(H_0) = \ell^*_{n1}(H_2) - \ell^*_{n2}(H_0) + O_p(1)$, and furthermore, $T^O_2 = T^*_2 + o_p(1)$. By the same arguments of the proof of Theorem 5, we have $\ell_{n2}(H_1) - \ell_{n2}(H_0) = \ell^O_{n2}(H_1) - \ell^O_{n2}(H_0) + O_p(1)$, and $T_2 = T^O_2 + o(1)$. It remains to show that $\sigma_n^{-1}(T^*_2 - \mu_n) \xrightarrow{\mathcal{L}} N(0,1)$, which can be proved by the similar arguments of Liu et al. (2014) and Liu et al. (2016). □

Proof of Theorem 7: Let $B_n = \mathbf{A}\Gamma_n^{-1}\mathbf{A}^T$, and

$$\Gamma_n = \sum_{i=1}^n q_2(\eta(\mathbf{V}_i;\boldsymbol{\theta}^0))\widehat{\boldsymbol{\Psi}}_i(\mathbf{X},\mathbf{Z})\widehat{\boldsymbol{\Psi}}_i(\mathbf{X},\mathbf{Z})^T, \text{ with } \widehat{\boldsymbol{\Psi}}_i(\mathbf{X},\mathbf{Z}) = \begin{pmatrix} X_{\widehat{m},i} - \tilde{\mathbf{P}}_n(\mathbf{X}_i) \\ \tilde{Z}_i - \tilde{\mathbf{P}}_n(\mathbf{Z}_i) \end{pmatrix}.$$

By Lemma S.5, T_3 can be decomposed as

$$\begin{aligned}
2(\ell_{n3}(H_1) - \ell_{n3}(H_0)) =& 2\sum_{i=1}^n q_1(\widehat{\eta}_{3,H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}_{3,H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}_{3,H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\} \\
&- \sum_{i=1}^n q_2(\widehat{\eta}_{3,H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}_{3,H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}_{3,H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\}^2 + o_p(1) \\
=& 2\sum_{i=1}^n q_1(\widehat{\eta}^O_{3,H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}^O_{3,H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}^O_{3,H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\} \\
&- \sum_{i=1}^n q_2(\widehat{\eta}^O_{3,H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}))\left\{\widehat{\eta}^O_{3,H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) - \widehat{\eta}^O_{3,H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}})\right\}^2 + o_p(1) \\
=& (\widehat{\boldsymbol{\theta}}_{H_0} - \widehat{\boldsymbol{\theta}}_{H_1})^T\Gamma_n(\widehat{\boldsymbol{\theta}}_{H_0} - \widehat{\boldsymbol{\theta}}_{H_1}) + o_p(1),
\end{aligned}$$

where

$$\begin{aligned}
\widehat{\eta}_{3,H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) &= \widehat{m}_0(\widehat{\boldsymbol{\beta}}^T_{0,H_0}\mathbf{X}_i, \widehat{\boldsymbol{\theta}}_{H_0}) + \widehat{\boldsymbol{\alpha}}^T_{0,H_0}\mathbf{Z}_i + (\widehat{m}_1(\widehat{\boldsymbol{\beta}}^T_{1,H_0}\mathbf{X}_i, \widehat{\boldsymbol{\theta}}_{H_0}) + \widehat{\boldsymbol{\alpha}}^T_{1,H_0}\mathbf{Z}_i)G_i, \\
\widehat{\eta}_{3,H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) &= \widehat{m}_0(\widehat{\boldsymbol{\beta}}^T_{0,H_1}\mathbf{X}_i, \widehat{\boldsymbol{\theta}}_{H_1}) + \widehat{\boldsymbol{\alpha}}^T_{0,H_1}\mathbf{Z}_i + (\widehat{m}_1(\widehat{\boldsymbol{\beta}}^T_{1,H_1}\mathbf{X}_i, \widehat{\boldsymbol{\theta}}_{H_1}) + \widehat{\boldsymbol{\alpha}}^T_{1,H_1}\mathbf{Z}_i)G_i, \\
\widehat{\eta}^O_{3,H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) &= \widehat{m}^O_0(\widehat{\boldsymbol{\beta}}^T_{0,H_0}\mathbf{X}_i, \widehat{\boldsymbol{\theta}}_{H_0}) + \widehat{\boldsymbol{\alpha}}^T_{0,H_0}\mathbf{Z}_i + (\widehat{m}^O_1(\widehat{\boldsymbol{\beta}}^T_{1,H_0}\mathbf{X}_i, \widehat{\boldsymbol{\theta}}_{H_0}) + \widehat{\boldsymbol{\alpha}}^T_{1,H_0}\mathbf{Z}_i)G_i, \\
\widehat{\eta}^O_{3,H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\theta}}) &= \widehat{m}^O_0(\widehat{\boldsymbol{\beta}}^T_{0,H_1}\mathbf{X}_i, \widehat{\boldsymbol{\theta}}_{H_1}) + \widehat{\boldsymbol{\alpha}}^T_{0,H_1}\mathbf{Z}_i + (\widehat{m}^O_1(\widehat{\boldsymbol{\beta}}^T_{1,H_1}\mathbf{X}_i, \widehat{\boldsymbol{\theta}}_{H_1}) + \widehat{\boldsymbol{\alpha}}^T_{1,H_1}\mathbf{Z}_i)G_i.
\end{aligned}$$

It is not difficult to see that the relation $\widehat{\boldsymbol{\theta}}_{H_0} = \widehat{\boldsymbol{\theta}}_{H_1} + \Gamma_n^{-1}\mathbf{A}^T B_n^{-1}(\gamma - \mathbf{A}\widehat{\boldsymbol{\theta}}_{H_1})$, which results in

$$\begin{aligned} T_3 &= \left(\gamma - \mathbf{A}\widehat{\boldsymbol{\theta}}_{H_1}\right)^T B_n^{-1}\mathbf{A}\Gamma_n^{-1}\sum_{i=1}^{n}\widehat{\boldsymbol{\Psi}}_i(\mathbf{X},\mathbf{Z})^{\otimes 2}\Gamma_n^{-1}\mathbf{A}^T B_n^{-1}\left(\gamma - \mathbf{A}\widehat{\boldsymbol{\theta}}_{H_1}\right) + o_p(1) \\ &= \left(\gamma - \mathbf{A}\widehat{\boldsymbol{\theta}}_{H_1}\right)^T B_n^{-1}\left(\gamma - \mathbf{A}\widehat{\boldsymbol{\theta}}_{H_1}\right) + o_p(1) \end{aligned}$$

As the same arguments of Liu et al. (2016), we have $T_3 \xrightarrow{\mathcal{L}} \chi_k^2$ under null hypothesis, and T_3 asymptotically follows a noncentral Chi-squared distribution with k degrees of freedom and noncentrality parameter ϕ under alternative. This completes the proof of Theorem 7. $\square$

References

Bosq, D. (1998), *Nonparametric statistics for stochastic processes.* Springer-Verlag, New York.

Cai, Z., Fan, J. and Li, R. (2000), "Efficient estimation and inferences for varying-coefficient models", *Journal of the American statistical Association*, **95**, 888-902.

Carroll, R. J., Fan, J., Gijbels, I. and Wand, M. P. (1997), "Generalized partially linear single-index models", *Journal of the American statistical Association*, **92**, 477-489.

Carroll, R. J., Ruppert, D. and Welsh, A. H. (1998), "Local estimating equations," *Journal of the American statistical Association*, **93**, 214-227.

Cui, X., Härdle, W. and Zhu, L. (2011), "The EFM approach for single-index models", *Annals of Statistics*, **39**, 1658-1688.

de Boor, C. *A practical guide to splines*, Springer, New York.

de Jong, P. (1987). A central limit theorem for generalized quadratic forms. *Probability Theory Related Fields*, **75**, 262-277.

DeVore, R. A. and Lorentz, G. G. (1993), *Constructive Approximation.* Springer, New York.

Fan and Huang, T. (2005), "Profile likelihood inferences on semiparametric varying-coefficient partially linear models", *Bernoulli*, **29**, 1031-1057.

Fan, J., Zhang, C. and Zhang, J. (2001), "Generalized likelihood ratio statistics and Wilks phenomenon", *Annals of statistics*, **29**, 153-193.

Li, Q. and Racine, R. S. (2007), *Nonparametric Econometrics: Theory and practice*. Princeton University Press, Princeton, N. J.

Liang, H., Liu, X., Li, R. and Tsai, C. L. (2010), "Estimation and testing for partially linear singleindex models", *Annals of Statistics*, **38**, 3811-3836.

Liu, X., Jiang, H. and Zhou, Y. (2013), "Local empirical likelihood inference for varying-coefficient density-ratio models based on case-control data", *Journal of the American statistical Association*, to appear.

Liu, X., Cui, Y. and Li, R. (2016), "Partial linear varying multi-index coefficient model for integrative gene-environment interactions". *Statistica Sinica*, **26**, 1037-1060.

Liu, R., Yang, N. and Härdle, W. (2013), "Oracally efficient two-step estimation of generalized additive model", *Journal of the American Statistical Association*, **108**, 619-631.

Ma, S. and Song, P. X. (2013), "Varying index coefficient models"

Ruppert, D., Sheathers, S. J. and Wand, M. P. (1995), "An effective bandwidth selector for local least squares regression," *Journal of the American statistical Association*, **90**, 1257-1270.

Sepanski, J. H., Knickerbocker, R. and Carroll, R. J. (1994), "A semiparametric correction for attenuation," *Journal of the American statistical Association*, **89**, 1366-1373.

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

Xia, Y. C. and Li, W. K. (1999), "On single-index coefficient regression models," *Journal of the American statistical Association*, **94**, 1275-1285.