Joint Test of Parametric and Nonparametric Effects in Partial Linear Models for Gene-Environment Interaction

Xu Liu[1], Ping-Shou Zhong[2], and Yuehua Cui[3]

[1]*School of Statistics and Management,*

Shanghai University of Finance and Economics,Shanghai 200433 China

[2]*Department of Mathematics, Statistics and Computer Science,*

University of Illinois at Chicago, Chicago, IL 60607 U.S.A.

[3]*Department of Statistics and Probability,*

Michigan State University, East Lansing, MI 48824 U.S.A.

Supplementary Material

Denote $q_k(\tilde{\eta}_i)$ by $q_k\{\tilde{\eta}(\mathbf{V}_i;\boldsymbol{\alpha}_0,\lambda)\}$, $k=1,2$, $i=1,\cdots,n$. 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

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

where $D_i = (\mathbf{B}_r(U_i)^T\tilde{X}_{i,l}, l = 1,\cdots,2p)^T$, $D_i(\mathbf{Z}) = (\mathbf{Z}_i^T, D_i^T)^T$, $\mathbf{D} =$

Corresponding author: cuiy@msu.edu

$(D_1, \cdots, D_n)^T$ which is an $n \times pJ_n$ matrix, and $\mathbf{D}(\mathbf{Z}) = (D_1(\mathbf{Z}), \cdots, D_n(\mathbf{Z}))^T$ which is an $n \times 2(q + pJ_n)$ matrix.

Lemma S.1 *Let assumptions* (A1) and (A4) *be satisfied. For any vector* $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1^T, \cdots, \boldsymbol{\zeta}_{2p}^T)^T$ *with* $\boldsymbol{\zeta}_l = (\boldsymbol{\zeta}_{s,l} : 1 \leq s \leq J_n)^T$ *and* $||\boldsymbol{\zeta}_l|| = 1$, $l = 1, \cdots, 2p$, *there exists constants* $0 < c_U < C_U < \infty$, *such that for any* $\boldsymbol{\alpha} \in \Theta$ *and for large enough* n,

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

$$\sup_{1 \leq s, s' \leq J_n, 1 \leq l \leq 2p} \left| n^{-1} \sum_{i=1}^{n} D_{i,sl} D_{i,s'l} - E[D_{i,sl} D_{i,s'l}] \right| = O\left(\sqrt{J_n^{-1} n^{-1} \log n}\right), \quad a.s., \tag{S.3}$$

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

and with probability approaching 1,

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

where $D_{i,sl} = B_{s,r}(u) X_{i,l}$, $i = 1, \cdots, n$, $s = 1, \cdots, J_n$, $l = 1, \cdots, 2p$.

Proof of Lemma S.1: By Theorem 5.4.2 of DeVore and Lorentz (1993) and assumption (A1), for any vector $\boldsymbol{\zeta}_l$ with $||\boldsymbol{\zeta}_l||_2 = 1$ and for large enough n, there exist constants $0 < c_l \leq C_l < \infty$, $l = 1, \cdots, 2p$, such that

$$c_l J_n^{-1} \leq \boldsymbol{\zeta}_l^T E\left[\mathbf{B}_r(U_i)\mathbf{B}_r(U_i)^T\right]\boldsymbol{\zeta}_l \leq C_l J_n^{-1}.$$

Let $\pi_{i,l} = \sum_{s=1}^{J_n} \zeta_{s,l} B_{s,q}(U_i)$ and $\pi_i = (\pi_{i,1}, \cdots, \pi_{i,2p})^T$. By assumptions (A1) and (A4), for large enough n, we have

$$\begin{aligned}
\boldsymbol{\zeta}^T E\left[\mathbf{U}\right]\boldsymbol{\zeta} =& E\left[q_2(\tilde{\eta}_i)\sum_{l=1}^{2p}\pi_{i,l}\tilde{X}_{i,l}\right]^2 \\
\leq& C_{q_2}^2 C_x E[\pi_i^T \pi_i] \\
=& C_{q_2}^2 C_x \sum_{l=1}^{2p} \boldsymbol{\zeta}_l^T E\left[\mathbf{B}_r(U_i)\mathbf{B}_r(U_i)^T\right]\boldsymbol{\zeta}_l \\
\leq& 2pC_{q_2}^2 C_x \text{min}_{1\leq l\leq 2p} C_l J_n^{-1}.
\end{aligned}$$

As the same way, we have $\boldsymbol{\zeta}^T E\left[\mathbf{U}\right]\boldsymbol{\zeta} \geq 2pc_{q_2}^2 c_x \text{max}_{1\leq l\leq 2p} c_l J_n^{-1}$. The second result in (S.2) can be shown directly from the first inequalities. Similarly it is easy to prove that (S.5) holds. (S.3) and (S.4) can be shown by Bernstein's inequality as Bosq (1998). □

Lemma S.2 *Let assumptions* (A1) (A3) and (A4) *be satisfied. For any* $\boldsymbol{\alpha} \in \boldsymbol{\Theta}_\alpha$, $||n^{-1}\mathbf{D}^T\mathbf{q}_1(\eta)||_2 = O_p(n^{-1/2})$.

Proof of Lemma S.2: By the law of large numbers, with probability approaching 1, we have

$$
\begin{aligned}
||n^{-1}\mathbf{D}^T\mathbf{q}_1(\eta)||_2^2 = & \sum_{1\le s\le J_n, 1\le l\le 2p}\left[n^{-1}\sum_{i=1}^{n} B_{s,q}(U_i)\tilde{X}_{il}q_1(\eta_i)\right]^2 \\
= & n^{-1}\sum_{1\le s\le J_n, 1\le l\le 2p} E\left[B_{s,q}(U_i)\tilde{X}_{il}q_1(\eta_i)\right]^2 + \boldsymbol{o}_{\text{a.s.}}(n^{-1}) \\
= & \boldsymbol{O}_{\text{a.s.}}(n^{-1}).
\end{aligned}
$$

□

Lemma S.3 *Let assumptions* (A1) and (A4) *be satisfied. There exists a constant* $0 \le c_D \le \infty$, *such that for large enough* n,

$$
\|n^{-1}\widehat{\mathbf{U}}(\mathbf{Z})^{-1}\mathbf{D}(\mathbf{Z})\mathbf{q}_2\|_\infty \le C_D.
$$

Proof of Lemma S.3: Let $S_n = \widehat{\mathbf{U}}(\mathbf{Z})$ with

$$
S_n = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix},
$$

where $S_{11} = n^{-1}\mathbf{Z}^T\mathbf{Z}$, $S_{12} = S_{21}^T = n^{-1}\mathbf{Z}^T\mathbf{D}$ and $S_{22} = \widehat{\mathbf{U}}$. Denote $S_{22.1} = S_{22} - S_{21}S_{11}^{-1}S_{12}$. For any $\boldsymbol{\zeta}$ as given in Lemma S.1, we have

$$
\begin{aligned}
\boldsymbol{\zeta}^T S_{22.1}\boldsymbol{\zeta} = & \boldsymbol{\zeta}^T\widehat{\mathbf{U}}\boldsymbol{\zeta} - n^{-2}\boldsymbol{\zeta}^T\mathbf{D}^T\mathbf{Z}S_{11}^{-1}\mathbf{Z}^T\mathbf{D}\boldsymbol{\zeta} \\
= & \boldsymbol{\zeta}^T\widehat{\mathbf{U}}\boldsymbol{\zeta} - c_z\boldsymbol{\zeta}^T\widehat{\mathbf{U}}\boldsymbol{\zeta} = (1-c_z)\boldsymbol{\zeta}^T\widehat{\mathbf{U}}\boldsymbol{\zeta},
\end{aligned}
$$

which is also followed by $\boldsymbol{\zeta}^T S_{22.1}^{-1}\boldsymbol{\zeta} = c_s\boldsymbol{\zeta}^T\widehat{\mathbf{U}}^{-1}\boldsymbol{\zeta}$, where c_z and c_s are constants. Thus, we have

$$
\begin{aligned}
&\|n^{-1}(\mathbf{0}_{pJ_n\times q}, I_{pJ_n})\widehat{\mathbf{U}}(\mathbf{Z})^{-1}\mathbf{D}(\mathbf{Z})^T\mathbf{q}_2\|_\infty \\
=&\|n^{-1}\begin{pmatrix}-S_{22.1}^{-1}S_{21}S_{11}^{-1} & S_{22.1}^{-1}\end{pmatrix}\mathbf{D}(\mathbf{Z})^T\mathbf{q}_2\|_\infty \\
=&\|n^{-1}(-S_{22.1}^{-1}S_{21}S_{11}^{-1}\mathbf{Z}^T + S_{22.1}^{-1}\mathbf{D}^T)\mathbf{q}_2\|_\infty \\
\le&\|n^{-1}S_{22.1}^{-1}\mathbf{D}^T\mathbf{Z}S_{11}^{-1}\mathbf{Z}^T\mathbf{q}_2\|_\infty + \|n^{-1}S_{22.1}^{-1}\mathbf{D}^T\mathbf{q}_2\|_\infty \\
\le&c_1\|n^{-1}\widehat{\mathbf{U}}^{-1}\mathbf{D}\mathbf{q}_2\|_\infty \\
\le&c_1\|\widehat{\mathbf{U}}^{-1}\|_\infty\|n^{-1}\mathbf{D}\mathbf{q}_2\|_\infty,
\end{aligned}
$$

where c_1 is a constant. By Bernstein's inequality in Bosq (1998), it can be shown that $||n^{-1}\sum_{i=1}^n \mathbf{D}_i\mathbf{q}_2||_\infty = O_p(J_n^{-1})$. We have $\|n^{-1}(\mathbf{0}_{2pJ_n\times q}, I_{2pJ_n})\widehat{\mathbf{U}}(\mathbf{Z})^{-1}\mathbf{D}(\mathbf{Z})^T\mathbf{q}_2\|_\infty \le c_2$, where c_2 is a constant. As the proof above, we have

$$
\begin{aligned}
&\|n^{-1}(I_q, \mathbf{0}_{q\times 2pJ_n})\widehat{\mathbf{U}}(\mathbf{Z})^{-1}\mathbf{D}(\mathbf{Z})^T\mathbf{q}_2\|_\infty \\
=&\|n^{-1}\begin{pmatrix}S_{11}^{-1} + S_{11}^{-1}S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1} & -S_{11}^{-1}S_{12}S_{22.1}^{-1}\end{pmatrix}\mathbf{D}(\mathbf{Z})^T\mathbf{q}_2\|_\infty \\
=&\|n^{-1}\left(S_{11}^{-1}\mathbf{Z}^T + S_{11}^{-1}S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1}\mathbf{Z}^T - S_{11}^{-1}S_{12}S_{22.1}^{-1}\mathbf{D}^T\right)\mathbf{q}_2\|_\infty \\
\le&\|n^{-1}S_{11}^{-1}\mathbf{Z}^T\mathbf{q}_2\|_\infty + \|n^{-3}S_{11}^{-1}\mathbf{Z}^T\mathbf{D}S_{22.1}^{-1}\mathbf{D}^T\mathbf{Z}S_{11}^{-1}\mathbf{Z}^T\mathbf{q}_2\|_\infty \\
&\quad + \|n^{-2}S_{11}^{-1}\mathbf{Z}^T\mathbf{D}S_{22.1}^{-1}\mathbf{D}^T\mathbf{q}_2\|_\infty \\
\le&c_3\|n^{-1}S_{11}^{-1}\mathbf{Z}^T\mathbf{q}_2\|_\infty \\
\le&c_4,
\end{aligned}
$$

combining with above proof, which arrives at the second part of Lemma

S.2, where c_3 and c_4 are constants.

□

The following Lemma states the convergence rate of the nonparametric estimators $\tilde{m}_l(u)$, $i = 1, \cdots, p$ and their first derivatives $\tilde{m}_l'(u)$.

Lemma S.4 *Let assumptions* (A1)-(A4) *be satisfied. We have*

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

$$|\tilde{\beta}_l(u) - \beta_l(u)| = \boldsymbol{O}_{a.s.}(\sqrt{N/n} + N^{-r})$$

uniformly for any $u \in [a_u, b_u]$;

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

$$|\tilde{\beta}_l'(u) - \beta_l'(u)| = \boldsymbol{O}_{a.s.}(\sqrt{N^3/n} + N^{1-r})$$

uniformly for any $u \in [a_u, b_u]$.

Proof of Lemma S.4: According to the result of de Boor (2001), for $\beta_l(\cdot)$ satisfying assumption (A2), there is a function $\bar{\beta}_l(u) = \mathbf{B}_r(u)^T \lambda_l$, such that

$$\sup_{u \in [a_u, b_u]} |\bar{\beta}_l(u) - \beta_l(u)| = O(J_n^{-r}). \tag{S.6}$$

Let $\mathbb{B}_r(\mathbf{u}) = (\mathbf{B}_r(u)^T, \cdots, \mathbf{B}_r(u)^T)^T$, and $\lambda = (\lambda_1, \cdots, \lambda_{2p})^T$. The estimate of λ solves equation

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

Recalling $\widehat{\mathbf{U}} = \boldsymbol{O}_{\text{a.s.}}(J_n^{-1})$ and $nN^{-2r-2} \to 0$, via Taylor series, we have

$$
\begin{aligned}
0 =& \frac{1}{n}\sum_{i=1}^{n} D_i q_1\{\tilde{\eta}_i)\} \\
&- \frac{1}{n}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\} D_i D_i^T \{(\widehat{\lambda} - \lambda)\}(1 + \boldsymbol{o}_{\text{a.s.}}(1)) \\
=& \frac{1}{n}\mathbf{D}^T \mathbf{q}_1\{\tilde{\eta}\} - \widehat{\mathbf{U}}(\widehat{\lambda} - \lambda)(1 + \boldsymbol{o}_{\text{a.s.}}(1)),
\end{aligned}
$$

which results in

$$
\widehat{\lambda} - \lambda = \frac{1}{n}\widehat{\mathbf{U}}^{-1}\mathbf{D}^T \mathbf{q}_1\{\tilde{\eta}\}(1 + \boldsymbol{o}_{\text{a.s.}}(1)).
$$

For each $u \in [a_u, b_u]$, by Lemma S.3 and (S.5) and assumption (A3), with probability approaching 1, we have

$$
\begin{aligned}
& E\left[\mathbb{B}_r(u)^T(\widehat{\lambda} - \lambda)\Big| \mathbf{X}, \mathbf{Z}, U\right]^2 \\
=& E\left[\left|C_D n^{-1}\mathbb{B}_r(u)^T \widehat{\mathbf{U}}^{-1}\mathbf{D}^T \mathbf{q}_1\{\tilde{\eta}\}\right|\Big| \mathbf{X}, \mathbf{Z}, U\right]^2 (1 + o_p(1)) \\
=& n^{-2}C_D^2 \mathbb{B}_r(u)^T \widehat{\mathbf{U}}^{-1}\mathbf{D}^T E[\mathbf{q}_1\{\tilde{\eta}\}^{\otimes 2}|\mathbf{X}, \mathbf{Z}, U]\mathbf{D}\widehat{\mathbf{U}}^{-1}\mathbb{B}_r(\mathbf{u})(1 + o_p(1)) \qquad \text{(S.7)} \\
=& n^{-1}C_D^2 C_\sigma \mathbb{B}_r(u)^T \widehat{\mathbf{U}}^{-1}\mathbb{B}_r(u)(1 + \boldsymbol{o}_{\text{a.s.}}(1)) \\
=& \boldsymbol{O}_{\text{a.s.}}(J_n/n),
\end{aligned}
$$

which implies by the law of large numbers that for each $u \in [a_u, b_u]$, $\left\|\mathbb{B}_r(u)^T(\widehat{\lambda} - \lambda)\right\|_2 = \boldsymbol{O}_{\text{a.s.}}(\sqrt{J_n/n})$. Thus, (S.7) combining (S.6) results in Lemma S.4 (a). Noting that $\|\mathbf{W}_1\|_\infty = O(J_n)$ where $\mathbf{W}_1$ is defined in Section 2.2, one can show similarly that the second part of Lemma S.3 holds.

□

Let $\mathbf{P}_n(\tilde{\mathbf{Z}}_i) = D_i^T \widehat{\boldsymbol{\zeta}}^Z$, where $\widehat{\boldsymbol{\zeta}}^Z = (\widehat{\boldsymbol{\zeta}}_1^Z, \cdots, \widehat{\boldsymbol{\zeta}}_{2q}^Z)$ is a $2pJ_n \times 2q$ matrix and for $k = 1, \cdots, 2q$,

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

Lemma S.5 *Let assumptions* (A1)-(A5) *be satisfied, and* $nN^{-4} \to \infty$ *and* $nN^{-2r-2} \to 0$, *as* $n \to \infty$. *We have*

$$\begin{aligned} D_i^T\{\widehat{\lambda} - \lambda\} =& n^{-1} D_i^T \widehat{\mathbf{U}}^{-1} \mathbf{D}^T \mathbf{q}_1\{\tilde{\eta}\} - \mathbf{P}(\tilde{\mathbf{Z}}_i)^T(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) \\ &+ o_p(\|\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\|_2) + o_p(\sqrt{N/n} + N^{-r}), \end{aligned} \tag{S.8}$$

where λ *is defined in Lemma S.3.*

Proof of Lemma S.5: The estimate of $\lambda(\boldsymbol{\theta})$ solves equation

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

Recalling $\widehat{\mathbf{U}} = O_p(J_n^{-1})$ and $nN^{-2r-2} \to 0$, via Taylor series, we have

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

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

$$
\begin{aligned}
& D_i^T\{\widehat{\lambda}-\lambda\} \\
=& n^{-1} D_i^T \widehat{\mathbf{U}}^{-1} \mathbf{D}^T \mathbf{q}_1\{\tilde{\eta}\} \\
& -n^{-1} D_i^T \widehat{\mathbf{U}}^{-1} \mathbf{D}^T \mathbf{W}_{q2} \tilde{\mathbf{Z}}\{(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha})+o_p(\|\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}\|_2\}+o_p(\sqrt{N/n}+N^{-r}).
\end{aligned} \tag{S.9}
$$

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

$$
n^{-1} D_i^T \widehat{\mathbf{U}}^{-1} \mathbf{D}^T \mathbf{W}_{q2} \tilde{\mathbf{Z}} = \mathbf{P}_n(\tilde{\mathbf{Z}}_i) + O_p(J_n^{1-r}).
$$

As the same as Lemma S.3, we can show that $\|\mathbf{P}_n(\tilde{\mathbf{Z}}_i) - \mathbf{P}(\tilde{\mathbf{Z}}_i)\|_\infty = O_p(\sqrt{N/n}+N^{-r})$. Therefore, with (S.9), this arrives at the result of Lemma S.4. □

Proof of Theorem 1: The estimates of $\boldsymbol{\alpha}$ solve

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

Then by Taylor expansion, we obtain

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

which arrives

$$
\begin{aligned}
& n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\}\tilde{\mathbf{Z}}_i\left[\tilde{\mathbf{Z}}_i-\mathbf{P}(\tilde{\mathbf{Z}}_i)\right]^T\{(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}) + o_p(\|\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}\|_2)\} \\
& \qquad = n^{-1}\sum_{i=1}^{n}\left[\tilde{\mathbf{Z}}_i-\mathbf{P}(\tilde{\mathbf{Z}}_i)\right] q_1\{\tilde{\eta}_i\}\{1+O_p(J_n^{1-r})\} + o_p(n^{-1/2}),
\end{aligned}
$$

Thus, by the law of large numbers, we have

$$n^{-1}\sum_{i=1}^{n} q_2\{\tilde{\eta}_i\}\tilde{\mathbf{Z}}_i\left[\tilde{\mathbf{Z}}_i-\mathbf{P}(\tilde{\mathbf{Z}}_i)\right]^T=\Sigma+O_p(n^{-2}).$$

Because the observations $\mathbf{V}_1,\cdots,\mathbf{V}_n$ are independent, by Lindeberg-Feller central limit theorem, it is easy to prove that

$$n^{-1/2}\sum_{i=1}^{n}\left[\tilde{\mathbf{Z}}_i-\mathbf{P}(\tilde{\mathbf{Z}}_i)\right]q_1\{\tilde{\eta}_i\}\xrightarrow{\mathcal{L}}N(\mathbf{0},\Sigma).$$

Then the proof of Theorem 1 can be completed by Slusky's theorem.

□

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

$$\sup_{u\in[a_u,b_u]}\left|\widehat{\beta}_l^O(u)-\beta_l(u)\right|=O_p(n^{-2/5}\sqrt{\log(n)}),$$

and for any $u\in[a_u,b_u]$*,*

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

where

$$\begin{aligned}b_l(u)=&\mu_2\beta_l''(u)/2,\\ v_l(u)=&\|K\|_2^2\{E[\rho(\mathbf{V})\tilde{X}_l^2|U=u]f(u)\}^{-1}.\end{aligned}$$

Proof of Lemma S.6: To facilitate expression, we prove only

$$\sup_{u\in[a_u,b_u]}\left|\widehat{\beta}_1^O(u)-\beta_1(u)\right|=O_p(n^{-2/5}\sqrt{\log(n)}).$$

Let

$$\tilde{\mathbf{X}} \equiv \tilde{\mathbf{X}}(u) = \begin{pmatrix} \tilde{X}_{1,1} & \cdots & \tilde{X}_{n,1} \\ (u_1-u)\tilde{X}_{1,1}/h_1 & \cdots & (u_n-u)\tilde{X}_{n,1}/h_1 \end{pmatrix}^T,$$

$$\tilde{\mathbf{W}}_{q_j}^O(\boldsymbol{\alpha}) \equiv \tilde{\mathbf{W}}_{q_j}^O(u,\boldsymbol{\alpha}) = \mathrm{diag}\left\{q_j\{\eta_{-1,1}^O\}K_{h_1}(u_1-u),\cdots,q_j\{\eta_{-1,n}^O\}K_{h_1}(u_n-u)\right\},$$

where $q_j\{\eta_{-1,i}^O\} = q_j\{\eta_{-1}^O(\mathbf{V}_i;a,b,\boldsymbol{\alpha})\}$ with $\eta_{-1}^O(\mathbf{V}_i;a,b,\boldsymbol{\alpha}) = \boldsymbol{\alpha}^T\tilde{\mathbf{Z}}_i+\sum_{l=2}^p\beta_l(u_i)^T\tilde{X}_{i,l}+a\tilde{X}_{i,1}+b(u_i-u)\tilde{X}_{i,1}$ and $j=1,2$, $i=1,\cdots,n$. The estimate $(\widehat{a}^O,\widehat{b}^O)$ solves equation

$$0 = n^{-1}\sum_{i=1}^n q_1\{\widehat{\eta}_{-1}^O(\mathbf{V}_i;\widehat{a}^O,\widehat{b}^O,\widehat{\boldsymbol{\alpha}})\}\begin{pmatrix} \tilde{X}_{i,1} \\ (u_i-u)\tilde{X}_{i,1}/h_1 \end{pmatrix}K_{h_1}(u_i-u).$$

Thus, we have

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

Noting that $\|\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}\| = O_p(n^{-1/2})$ by Property 1, it is not hard to see that, for any $u\in[a_u,b_u]$, $\tilde{\mathbf{W}}_{q_j}^O(\widehat{\boldsymbol{\alpha}}) = \tilde{\mathbf{W}}_{q_j}^O(\boldsymbol{\alpha})+O_p(n^{-1/2})$, which implies that

$$\sup_{u\in[a_u,b_u]}\left|\widehat{\beta}_1^O(u,\widehat{\boldsymbol{\alpha}})-\widehat{\beta}_1^O(u,\boldsymbol{\alpha})\right| = O_p(n^{-1/2}). \tag{S.10}$$

Along the lines of Theorem 2.5 and 2.6 in Li and Racine (2007), we have, under Assumption A.?

$$\sup_{u\in[a_u,b_u]}\left|\widehat{\beta}_1^O(u,\boldsymbol{\alpha})-\beta_l(u)\right| = O_p(n^{-2/5}\sqrt{\log(n)}),$$

which implies that

$$\sup_{u\in[a_u,b_u]}\left|\widehat{\beta}_1^O(u,\widehat{\boldsymbol{\alpha}})-\beta_1(u)\right| \leq \sup_{u\in[a_u,b_u]}\left|\widehat{\beta}_1^O(u,\widehat{\boldsymbol{\alpha}})-\widehat{\beta}_1^O(u,\boldsymbol{\alpha})\right| + \sup_{u\in[a_u,b_u]}\left|\widehat{\beta}_1^O(u,\boldsymbol{\alpha})-\beta_1(u)\right|$$
$$=O_p(n^{-1/2})+O_p(n^{-2/5}\sqrt{\log(n)})$$
$$=O_p(n^{-2/5}\sqrt{\log(n)}),$$

Therefore, this arrives the first part of Lemma S.6, and

$$\sqrt{nh_l}\left\{\widehat{\beta}_1^O(u,\boldsymbol{\alpha})-\beta_1(u)-b_1(u)h_1^2\right\}\xrightarrow{\mathcal{L}} N(0,v_1(u)).$$

The second part can be show by combining (S.16) and the asymptotic normality of $\widehat{\beta}_1^O(u,\boldsymbol{\alpha})$.

□

Lemma S.7 *Suppose that assumptions* (A.1)-(A.6) *in the Appendix hold, and* $nN^{-4}\to\infty$ *and* $nN^{-\delta}\to 0$ *with* $\delta=\min(2r+2,5r/2)$*, then we have*

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

Proof of Lemma S.7: 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

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

By Lemma S.4, 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.4.

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

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

Noting that $q_2(\cdot)$ is bounded, by Lemma S.4, we have

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

Define $\boldsymbol{\Phi}_b = (\phi_{b,1}^T, \cdots, \phi_{b,2p}^T)^T$, $\boldsymbol{\Phi}_v = (\phi_{v,1}^T, \cdots, \phi_{v,2p}^T)^T$ and $\boldsymbol{\Phi}_r = (\phi_{r,1}^T, \cdots, \phi_{r,2p}^T)^T$, 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 et 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

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

Applying Cauchy-Schwarz inequality, we have

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

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

with

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

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

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

By de Boor (2001), we have, for any $u\in[a_u,b_u]$,

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

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

which implies that

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

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

□

Lemma S.8 *Suppose that assumptions* (A.1)-(A.6) *in the Appendix hold, and* $nN^4\to\infty$ *and* $nN^{-\delta}\to 0$ *with* $\delta = \min(2r+2, 5r/2)$,

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

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

Proof of Lemma S.8: Let $\check{\mathbf{X}}_{i1}(u) = (X_{i,1},(U_i-u)X_{i,1}/h_1)^T$. According to the boundedness of $q_2(\cdot)$ and the law of large numbers, we have

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

□

Lemma S.9 *Suppose that assumptions* (A.1)-(A.6) *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* $l = 1, \cdots, p$*,*

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

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.9: 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.7 and Lemma S.8, 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.9 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 2: Due to $nh^5 = O(1)$, we have $\sqrt{nh_l}n^{-1/2}\log n = o_p(1)$. By Lemma S.9, we have

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

Lemma S.6 has proved that

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

where $b_l(u) = \mu_2\beta_l''(u)/2$, and $v_l(u) = \|K\|_2^2\{E[\rho(\mathbf{V})\tilde{X}_l^2|U = u]f(u)\}^{-1}$. Thus Theorem **??** can be shown straightforwardly by Lemma S.6 and Lemma S.9 in the Supplementary Materials.

□

$\ell_n(H_0) = \sum_{i=1}^n Q(g^{-1}\{\widehat{\eta}_{H_0}(\mathbf{V}_i; \widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\theta}})\}, Y_i)$ and $\ell_n(H_1) = \sum_{i=1}^n Q(g^{-1}\{\widehat{\eta}_{H_1}(\mathbf{V}_i; \widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\theta}})\}, Y_i)$, where $\widehat{\eta}_{H_0}(\mathbf{V}_i; \widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\theta}}) = \mathbf{Z}_i^T\widehat{\boldsymbol{\alpha}}_{0,H_0} + \mathbf{X}_i^T\widehat{\boldsymbol{\theta}}_{0,H_0}(u) + \mathbf{Z}_i^T\widehat{\boldsymbol{\alpha}}_{1,H_0}G_i$ and $\widehat{\eta}_{H_1}(\mathbf{V}_i; \widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\theta}}) = \mathbf{Z}_i^T\widehat{\boldsymbol{\alpha}}_{0,H_1} + \mathbf{X}_i^T\widehat{\boldsymbol{\theta}}_{0,H_1}(u) + \{\mathbf{Z}_i^T\widehat{\boldsymbol{\alpha}}_{1,H_1} + \mathbf{X}_i^T\widehat{\boldsymbol{\theta}}_{1,H_1}(u)\}G_i$. We define the following GLR test statistic,

$$T_{NP} = 2(\ell_n(H_1) - \ell_n(H_0)). \tag{S.11}$$

To prove Theorem 3, we need following Lemma. Similarly to Liu et. al. (2016), assuming that nonparametric functions $\boldsymbol{\beta}(u)$ are known, we can construct a GLR statistic based on the "Oracle" estimator $\widehat{\boldsymbol{\beta}}^O(u)$. Consider hypothesis test (3.14). Let $\widehat{\beta}^O_{l,H_0}(u)$ and $\widehat{\beta}^O_{l,H_1}(u)$ be the "Oracle" estimates under H_0 and H_1 as the same as Section 2.1, respectively. The resulting likelihoods under H_0 and H_1 in hypothesis test (3.14) are

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

where $\widehat{\eta}^O_{H_0}(\mathbf{V}_i;\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{\theta}}) = \mathbf{Z}_i^T\widehat{\boldsymbol{\alpha}}_{0,H_0}+\mathbf{X}_i^T\widehat{\boldsymbol{\theta}}^O_{0,H_0}(U_i)+\mathbf{Z}_i^T\widehat{\boldsymbol{\alpha}}_{1,H_0}G_i$ and $\widehat{\eta}^O_{H_1}(\mathbf{V}_i;\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{\theta}}) = \mathbf{Z}_i^T\widehat{\boldsymbol{\alpha}}_{0,H_1}+\mathbf{X}_i^T\widehat{\boldsymbol{\theta}}^O_{0,H_1}(U_i)+\{\mathbf{Z}_i^T\widehat{\boldsymbol{\alpha}}_{1,H_1}+\mathbf{X}_i^T\widehat{\boldsymbol{\theta}}^O_{1,H_1}(U_i)\}G_i$. We define the following Oracle-version GLR test statistic,

$$T^O_{NP} = 2(\ell_n^O(H_1) - \ell_n^O(H_0)). \tag{S.12}$$

Let $a_K = \{K(0) - 1/2\int K^2(u)du\}[\int\{K(u) - 1/2K*K(u)\}du]^{-1}$, where $K*K(u)$ denotes the convolution of K.

The following Lemma states the asymptotic distribution of T^O_{NP} under H_0^{NP}.

Lemma S.10 *Suppose that assumptions* (A.1)-(A.6) *in the Appendix hold, and* $nN^4 \to \infty$ *and* $nN^{-2r-2} \to 0$, *then under* H_0^{NP} *in* (3.8), *when* $nh^{9/2} \to$

0,

$$\sigma_n^{-1}(T_{NP}^O - \mu_n) \xrightarrow{\mathcal{L}} N(0,1),$$

where $\sigma_n^2 = 2h^{-1}|\Omega| \int \{K(u) - 1/2K*K(u)\}^2\, du$ *and* $\mu_n = h^{-1}|\Omega| \left\{K(0) - 1/2\int K^2(u)du\right\}$.

Proof of Lemma S.10: Define

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

where $\eta_{H_0}^*(\mathbf{V}_i;\boldsymbol{\alpha},\widehat{\boldsymbol{\theta}}^O) = \mathbf{Z}_i^T\boldsymbol{\alpha}_0^0 + \mathbf{X}_i^T\widehat{\boldsymbol{\theta}}_{0,H_0}^O(U_i) + \mathbf{Z}_i^T\boldsymbol{\alpha}_1^0$ and $\widehat{\eta}_{H_1}^*(\mathbf{V}_i,\boldsymbol{\alpha},\widehat{\boldsymbol{\theta}}^O) = \mathbf{Z}_i^T\boldsymbol{\alpha}_0^0 + \mathbf{X}_i^T\widehat{\boldsymbol{\theta}}_{0,H_1}^O(U_i) + \{\mathbf{Z}_i^T\boldsymbol{\alpha}_1^0 + \mathbf{X}_i^T\widehat{\boldsymbol{\theta}}_{1,H_1}^O(U_i)\}G_i$. It is easy to see that

$$\begin{aligned}
\ell_n^O(H_0) - \ell_n^*(H_0) =& O_p(1),\\
\ell_n^O(H_1) - \ell_n^*(H_1) =& O_p(1),
\end{aligned}$$

which implies that $T_{NP}^O = T_{NP}^* + O_p(1)$. It remains to show that

$$\sigma_n^{-1}(T_{NP}^* - \mu_n) \xrightarrow{\mathcal{L}} N(0,1).$$

Denote $\eta_{H_0}^*(\mathbf{V}_i;\boldsymbol{\alpha},\widehat{\boldsymbol{\theta}}^O)$ by η_{i,H_0}^* and $q_1\{\eta_{i,H_0}^*\}$ by ε_i. Let $\check{\mathbf{X}}_i = \mathbf{X}_i G_i$. By the mean value theory, there exists a $\bar{\eta}^*(\mathbf{V}_i;\boldsymbol{\alpha},\widehat{\boldsymbol{\theta}}^O)$ between η_{i,H_0}^* and $\widehat{\eta}_{H_1}^*(\mathbf{V}_i;\boldsymbol{\alpha},\widehat{\boldsymbol{\theta}}^O)$ such that

$$T_{NP}^* = 2\sum_{i=1}^n \varepsilon_i \check{\mathbf{X}}_i^T \widehat{\boldsymbol{\theta}}_{1,H_1}^O(U_i) + \sum_{i=1}^n \bar{q}_{i,2}\widehat{\boldsymbol{\theta}}_{1,H_1}^O(U_i)^T\check{\mathbf{X}}_i\check{\mathbf{X}}_i^T\widehat{\boldsymbol{\theta}}_{H_1}^O(U_i),$$

where $\bar{q}_{i,2}$ denotes $q_2\{\bar{\eta}^*(\mathbf{V}_i;\boldsymbol{\alpha})\}$. Invoking Lemma S.6 and $\beta_{l0}(u)=0$, we have

$$\widehat{\boldsymbol{\theta}}^{O}_{H_1}(u)=\{\mathbf{a}_n(u)+R_n(u)\}(1+o_p(1)),$$

where

$$\begin{aligned}
\mathbf{a}_n(u)=&\frac{1}{nh}\Gamma(u)\sum_{i=1}^n\varepsilon_i\check{\mathbf{X}}_iK((U_i-u)/h),\\
R_n(u)=&\frac{1}{nh}\Gamma(u)\sum_{i=1}^n\{\boldsymbol{\theta}(U_i)^T\check{\mathbf{X}}_i-\tau(u)^T\tilde{\mathbf{X}}_i\}\mathbf{X}_iK((U_i-u)/h),
\end{aligned}$$

and $\Gamma(u)=\{E[\rho(\mathbf{V})\check{\mathbf{X}}\check{\mathbf{X}}^T|U=u]f(u)\}^{-1}$, $\tau(u)=(\boldsymbol{\theta}(u)^T,h\boldsymbol{\theta}'(u)^T)^T$ and $\tilde{\mathbf{X}}_i=\tilde{\mathbf{X}}_i(u)=(\mathbf{X}_i^TG_i,\mathbf{X}_i^TG_i(U_i-u)/h)^T$. Define

$$\begin{aligned}
R_n^{(1)}=&2\sum_{i=1}^n\varepsilon_iR_n(U_i)^T\check{\mathbf{X}}_i,\\
R_n^{(2)}=&2\sum_{i=1}^n\bar{q}_{i,2}\mathbf{a}_n(U_i)^T\check{\mathbf{X}}_i\check{\mathbf{X}}_i^TR_n(U_i),\\
R_n^{(3)}=&\sum_{i=1}^n\bar{q}_{i,2}R_n(U_i)^T\check{\mathbf{X}}_i\check{\mathbf{X}}_i^TR_n(U_i).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
T^*_{NP}=&\frac{2}{n}\sum_{i=1}^n\varepsilon_i\left\{\sum_{k=1}^n\varepsilon_k\check{\mathbf{X}}_k^T\Gamma(U_i)\check{\mathbf{X}}_iK_h(U_k-U_i)\right\}\\
&+\frac{1}{n^2}\sum_{k=1}^n\sum_{i=1}^n\sum_{j=1}^n\bar{q}_{k,2}\varepsilon_i\varepsilon_j\check{\mathbf{X}}_i^T\Gamma(U_k)\check{\mathbf{X}}_k\check{\mathbf{X}}_k^T\Gamma(U_k)\check{\mathbf{X}}_jK_h(U_i-U_k)K_h(U_j-U_k)\\
&+R_n^{(1)}+R_n^{(2)}+R_n^{(3)}+O_p(n^{-1}h^{-2})\\
\equiv&\bar{T}_n+\bar{S}_n+R_n^{(1)}+R_n^{(2)}+R_n^{(3)}+O_p(n^{-1}h^{-2}).
\end{aligned}$$

By the condition $nh^{9/2} \to 0$ and some direct but tedious calculations, it can be shown

$$\bar{R}_n \equiv R_n^{(1)} + R_n^{(2)} + R_n^{(3)} = O_p(nh^4 + n^{1/2}h^2) = o_p(h^{-1/2}). \tag{S.13}$$

It is easy to prove that

$$\bar{T}_n = 2h^{-1}K(0)|\Omega| + \frac{2}{n}\sum_{i\neq k}^{n} \varepsilon_i\varepsilon_k \check{\mathbf{X}}_k^T \Gamma(U_i)\check{\mathbf{X}}_i K_h(U_k - U_i) + o_p(h^{-1/2}). \tag{S.14}$$

Let $\bar{S}_n = S_{n1} + S_{n2}$, where

$$\begin{aligned} S_{n1} =& \frac{1}{n^2}\sum_{i=1}^{n}\varepsilon_i^2 \sum_{k=1}^{n} \bar{q}_{k,2}\left\{\check{\mathbf{X}}_k^T\Gamma(U_k)\check{\mathbf{X}}_i K_h(U_k - U_i)\right\}^2, \\ S_{n2} =& \frac{1}{n^2}\sum_{i\neq j}^{n}\varepsilon_i\varepsilon_j \sum_{k=1}^{n} \bar{q}_{k,2}\check{\mathbf{X}}_i^T\Gamma(U_k)\check{\mathbf{X}}_k\check{\mathbf{X}}_k^T\Gamma(U_k)\check{\mathbf{X}}_j K_h(U_i - U_k)K_h(U_j - U_k). \end{aligned}$$

It is easy to see that

$$S_{n1} = \varrho_n(1 + o(1)) + O_p(n^{-3/2}h^{-2}) + O_p((nh^2)^{-1}) + o_p(h^{-1/2}),$$

where

$$\varrho_n = \frac{1}{n(n-1)}\sum_{i\neq k}^{n} \bar{q}_{k,2}\varepsilon_i^2\left\{\check{\mathbf{X}}_k^T\Gamma(U_k)\check{\mathbf{X}}_i K_h(U_k - U_i)\right\}^2.$$

According to Hoeffding's decomposition (Serfling 1980) for the variance of U-statistics, by tedious calculations we can show that the variance of ϱ_n is

$$\varrho_n = O_p(n^{-1}h^{-2}).$$

It is straightforward to calculate the expectation of ϱ_n

$$E\varrho_n = -h^{-1}|\Omega| \int K^2(t)dt + o_p(h^{-1}),$$

which implies that

$$S_{n1} = -h^{-1}|\Omega| \int K^2(t)dt + o_p(h^{-1/2}). \tag{S.15}$$

S_{n2} can be further decomposed as $S_{n2} = S_{n21} + S_{n22}$, where

$$S_{n21} = \frac{1}{n^2} \sum_{i \neq j}^{n} \varepsilon_i \varepsilon_j \sum_{k \neq i,j}^{n} \bar{q}_{k,2} \check{\mathbf{X}}_i^T \Gamma(U_k) \check{\mathbf{X}}_k \check{\mathbf{X}}_k^T \Gamma(U_k) \check{\mathbf{X}}_j K_h(U_i - U_k) K_h(U_j - U_k),$$

$$S_{n22} = \frac{K(0)}{n^2 h} \sum_{i \neq j}^{n} \varepsilon_i \varepsilon_j \left[\bar{q}_{i,2} \check{\mathbf{X}}_i^T \Gamma(U_i) \check{\mathbf{X}}_i \check{\mathbf{X}}_i^T \Gamma(U_i) \check{\mathbf{X}}_j + \bar{q}_{j,2} \check{\mathbf{X}}_i^T \Gamma(U_j) \check{\mathbf{X}}_j \check{\mathbf{X}}_j^T \Gamma(U_j) \check{\mathbf{X}}_j \right] K_h(U_i - U_j).$$

It can be show by tedious calculation that

$$\mathrm{var}(S_{n22}) = O(n^{-2}h^{-3}),$$

which implies that

$$S_{n22} = o(h^{-1/2}). \tag{S.16}$$

Let

$$Q_{ijk} = \bar{q}_{k,2} \Gamma(U_k) \check{\mathbf{X}}_k \check{\mathbf{X}}_k^T \Gamma(U_k) K_h(U_i - U_k) K_h(U_j - U_k).$$

We can prove that

$$E\left\{ n^{-1} \sum_{k \neq i,j}^{n} Q_{ijk} - E\left[Q_{ijk} | (U_i, U_j) \right] \right\}^2 \leq n^{-2} \sum_{k \neq i,j}^{n} EQ_{ijk}^2 = O((nh^2)^{-1}),$$

which implies that

$$
\begin{aligned}
S_{n21} =& \frac{n-2}{n^2} \sum_{i \neq j}^{n} \varepsilon_i \varepsilon_j \check{\mathbf{X}}_i^T E\left[Q_{ijk} | (U_i, U_j)\right] \check{\mathbf{X}}_j + o_p(h^{-1/2}) \\
=& - \frac{1}{nh} \sum_{i \neq j}^{n} \varepsilon_i \varepsilon_j \check{\mathbf{X}}_i^T \Gamma(U_i) \check{\mathbf{X}}_j \tilde{K}((U_i - U_j)/h) + o_p(h^{-1/2}),
\end{aligned} \tag{S.17}
$$

where $\tilde{K}(t)$ denotes the convolution of $K(\cdot)$. By (S.13)-(S.17), we have

$$
T_{NP}^{*} = \mu_n + \Upsilon(n) h^{-1/2} + o_p(h^{-1/2}),
$$

where

$$
\begin{aligned}
\mu_n =& h^{-1} |\Omega| \left\{ 2K(0) - \int K^2(t) dt \right\}, \\
\Upsilon(n) =& \frac{1}{n} h^{-1/2} \sum_{i \neq j}^{n} \varepsilon_i \varepsilon_j \check{\mathbf{X}}_i^T \Gamma(U_i) \check{\mathbf{X}}_j \{ 2K((U_i - U_j)/h) - \tilde{K}((U_i - U_j)/h) \}.
\end{aligned}
$$

It remains to prove that

$$
\Upsilon(n) \xrightarrow{\mathcal{L}} N(0, v^2)
$$

with $v^2 = 2|\Omega| \int \{K(t) - 1/2\tilde{K}(t)\}^2 dt$. Define

$$
\Phi_{ij} = h^{-1/2} \varepsilon_i \varepsilon_j \check{\mathbf{X}}_i^T \Gamma(U_i) \check{\mathbf{X}}_j \{ 2K((U_i - U_j)/h) - \tilde{K}((U_i - U_j)/h) \}
$$

and

$$
\Upsilon(n) = \sum_{i<j} \Upsilon_{ij}.
$$

where $\Upsilon_{ij} = \frac{1}{n}(\Phi_{ij} + \Phi_{ji})$. Define $v_n^2 = \mathrm{Var}(\Upsilon(n))$. According to Proposition 3.2 in de Jong (1987), it suffices to check following conditions:

(a) $\Upsilon(n)$ is clean [see de Jong (1987) for the definition],

(b) $v_n^2 \to v^2$,

(c) ζ_1 is of lower order than v_n^4,

(d) ζ_2 is of lower order than v_n^4,

(e) ζ_3 is of lower order than v_n^4,

where

$$\begin{aligned}
\zeta_1 =& E \sum_{1 \le i < j \le n} \Upsilon_{ij}^4, \\
\zeta_2 =& E \sum_{1 \le i < j < k \le n} \{\Upsilon_{ij}^2 \Upsilon_{ik}^2 + \Upsilon_{ji}^2 \Upsilon_{jk}^2 + \Upsilon_{ki}^2 \Upsilon_{kj}^2\}, \\
\zeta_3 =& E \sum_{1 \le i < j < k < l \le n} \{\Upsilon_{ij} \Upsilon_{ik} \Upsilon_{lj} \Upsilon_{lk} + \Upsilon_{ij} \Upsilon_{il} \Upsilon_{kj} \Upsilon_{kl} + \Upsilon_{ik} \Upsilon_{il} \Upsilon_{jk} \Upsilon_{jl}\}.
\end{aligned}$$

Below we check each of conditions as follows. Condition (a) holds obviously. Then we calculate the variance v_n^2 as follows. (a) implies that $v_n^2 = E\Upsilon(n)^2$. By de Jong (1987), we have

$$v_n^2 = E \sum_{1 \le i < j \le n} \Upsilon_{ij}^2 = \frac{4}{n^2} E \sum_{1 \le i < j \le n} \Phi_{ij}^2.$$

We can show that

$$E \sum_{1 \le i < j \le n} \Phi_{ij}^2 = \frac{n(n-1)}{2} p|\Omega| \int \{K(t) - \frac{1}{2}\tilde{K}(t)\}^2 dt,$$

which implies that (b) holds. Noting that

$$E\sum_{1\le i<j\le n}\{\Phi_{ij}^4\}=O(n^2h^{-3}),E\sum_{1\le i<j\le n}\{\Phi_{ij}^3\Phi_{ji}\}=O(n^2h^{-2}),E\sum_{1\le i<j\le n}\{\Phi_{ij}^2\Phi_{ji}^2\}=O(n^2h^{-2}),$$

we have $\zeta_1=O(h^2n^{-4})O(n^2h^{-3})=O(n^{-2}h^{-1})$. Similarly conditions (d) is shown by noting that

$$E\sum_{1\le i<j<k\le n}\{\Upsilon_{ij}^2\Upsilon_{ik}^2\}=h^2n^{-4}O(n^3h^{-2})=O(n^{-1}),$$

which implies that $\zeta_2=O(n^{-1})$. It is obvious by straightforward calculations that,

$$\begin{aligned}E\sum_{1\le i<j<k<l\le n}\{\Phi_{ij}\Phi_{ik}\Phi_{lj}\Phi_{lk}\}=&O(n^4h^{-1}),\\ E\sum_{1\le i<j<k<l\le n}\{\Phi_{ji}\Phi_{ik}\Phi_{lj}\Phi_{lk}\}=&O(n^4h^{-1}),\\ E\sum_{1\le i<j<k<l\le n}\{\Phi_{ji}\Phi_{ki}\Phi_{lj}\Phi_{lk}\}=&O(n^4h^{-1}),\end{aligned}$$

which result in

$$E\sum_{1\le i<j<k<l\le n}\{\Upsilon_{ij}\Upsilon_{ik}\Upsilon_{lj}\Upsilon_{lk}\}=O(h^2n^{-4})O(n^4h^{-1})=O(h).$$

Therefore, we have $\zeta_3=O(h)$, which implies that condition (e) holds. This completes the proof of Lemma S.10.

□

Proof of Theorem 3: According to Lemma S.9, we have

$$\ell_n^O(H_0) - \ell_n(H_0) = O_p(\log n),$$

$$\ell_n^O(H_1) - \ell_n(H_1) = O_p(\log n),$$

which implies that

$$T_{NP} = T_{NP}^O + O_p(\log n).$$

Therefore, applying Lemma S.10, Theorem 3 can be proved directly.

□

Proof of Lemma 1: Invoking the proof of Theorem 1 and 3, we have

$$n^{1/2}\Sigma_\alpha^{-1/2}(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) = n^{-1/2}\Sigma_\alpha^{-1/2}\sum_{i=1}^n (\mathbf{Z}_i - \mathbf{P}(\mathbf{Z}_i))\varepsilon_i + o_p(1),$$

$$\sigma_n^{-1}(T_{NP} - \mu_n) = v^{-1}\Upsilon(n) + o_p(1),$$

where $\varepsilon_i = q_1(\eta_{i,H_0}^*)$, $\Upsilon(n) = \frac{1}{n}h^{-1/2}\sum_{i\neq j}^n \varepsilon_i\varepsilon_j \check{\mathbf{X}}_i^T \Gamma(U_i)\check{\mathbf{X}}_j\{2K((U_i - U_j)/h) - \tilde{K}((U_i - U_j)/h)\}$ and $v^2 = 2|\Omega| \int \{K(t) - 1/2\tilde{K}(t)\}^2 dt$ are defined in the proof of Lemma S.10. Let

$$\mathbb{I}_{1n} = \sum_{k\neq i,j}^n (\mathbf{Z}_k - \mathbf{P}(\mathbf{Z}_k))\varepsilon_k \sum_{i\neq j}^n \varepsilon_i\varepsilon_j \check{\mathbf{X}}_i^T \Gamma(U_i)\check{\mathbf{X}}_j\{2K((U_i - U_j)/h) - \tilde{K}((U_i - U_j)/h)\},$$

$$\mathbb{I}_{2n} = \sum_{i\neq j}^n \varepsilon_i^2\varepsilon_j(\mathbf{Z}_i - \mathbf{P}(\mathbf{Z}_i))\check{\mathbf{X}}_i^T \Gamma(U_i)\check{\mathbf{X}}_j\{2K((U_i - U_j)/h) - \tilde{K}((U_i - U_j)/h)\}.$$

It is easy to see that $E[\mathbb{I}_{1n}] = 0$ and $E[\mathbb{I}_{2n}] = 0$. Therefore, we have

$$\mathrm{COV}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2) = n^{-1/2}v^{-1}\Sigma_\alpha^{-1/2}(\mathbb{I}_{1n} + \mathbb{I}_{2n}) + o_p(1),$$

which results directly in $\mathrm{COV}(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2) \stackrel{P}{\to} 0$.

□

Proof of Theorem 4: This Theorem follows by Theorem 1 and Lemma S.10.

□

The following Lemma states the asymptotic distribution of T_{NP}^O under H_1^{NP}.

Lemma S.11 *Suppose that assumptions in Theorem 5 hold, then under H_1^{NP} in* (3.8),

$$\sigma_n^{-1}(T_{NP}^O - \mu_n - d_n) \stackrel{\mathcal{L}}{\to} N(0,1),$$

where $d_{2n} = nE[\rho(V)\boldsymbol{\theta}_n(U)^T \check{\mathbf{X}}\check{\mathbf{X}}^T \boldsymbol{\theta}_n(U)]$.

Proof of Lemma S.11: Along the lines of Lemma S.10, we can prove that

$$\begin{aligned}\ell_n^O(H_0) - \ell_n^*(H_0) =& O_p(1),\\ \ell_n^O(H_1) - \ell_n^*(H_1) =& O_p(1),\end{aligned}$$

which implies that $T_{NP}^O = T_{NP}^* + O_p(1)$. It remains to show that

$$\sigma_n^{-1}(T_{NP}^* - \mu_n) \stackrel{\mathcal{L}}{\to} N(0,1).$$

Denote $\tilde{\eta}_{H_1}^*(\mathbf{V}_i;\boldsymbol{\alpha}) = \theta_0(U_i) + \boldsymbol{\alpha}^T\mathbf{Z}_i + \check{\mathbf{X}}_i^T\boldsymbol{\theta}_n(U_i)$ by $\tilde{\eta}_{i,H_1}^*$, and $q_1\{\tilde{\eta}_{i,H_1}^*\}$ by ε_i. Let $\tilde{\ell}_n^*(H_1) = \sum_{i=1}^n Q(g^{-1}\{\tilde{\eta}_{H_1}^*(\mathbf{V}_i;\boldsymbol{\alpha})\}, Y_i)$. By the mean value theory, there exists a $\bar{\eta}^*(\mathbf{V}_i;\boldsymbol{\alpha})$ between $\tilde{\eta}^*_{i,H_1}$ and $\widehat{\eta}^*_{H_1}(\mathbf{V}_i;\boldsymbol{\alpha})$ such that

$$\begin{aligned}
T^*_{NP} =& 2(\ell^*_n(H_1) - \ell^*_n(H_0)) \\
=& 2(\ell^*_n(H_1) - \tilde{\ell}^*_n(H_1)) - 2(\ell^*_n(H_0) - \tilde{\ell}^*_n(H_1)) \\
=& 2\sum_{i=1}^n \varepsilon_i \check{\mathbf{X}}_i^T(\widehat{\boldsymbol{\theta}}^O_{H_1}(U_i) - \boldsymbol{\theta}_n(U_i)) + \sum_{i=1}^n \bar{q}_{i,2}(\widehat{\boldsymbol{\theta}}^O_{H_1}(U_i) - \boldsymbol{\theta}_n(U_i))^T \check{\mathbf{X}}_i \check{\mathbf{X}}_i^T(\widehat{\boldsymbol{\theta}}^O_{H_1}(U_i) - \boldsymbol{\theta}_n(U_i)) \\
& + 2\sum_{i=1}^n \varepsilon_i \check{\mathbf{X}}_i^T \boldsymbol{\theta}_n(U_i) - \sum_{i=1}^n \bar{q}_{i,2}\boldsymbol{\theta}_n(U_i))^T \check{\mathbf{X}}_i \check{\mathbf{X}}_i^T \boldsymbol{\theta}_n(U_i) \\
=& 2\sum_{i=1}^n \varepsilon_i \check{\mathbf{X}}_i^T(\mathbf{a}_n(U_i) + R_n(U_i))(1 + o_p(1)) \\
& + \sum_{i=1}^n \bar{q}_{i,2}(\mathbf{a}_n(U_i) + R_n(U_i))^T \check{\mathbf{X}}_i \check{\mathbf{X}}_i^T(\mathbf{a}_n(U_i) + R_n(U_i))(1 + o_p(1)) \\
& + 2\sum_{i=1}^n \varepsilon_i \check{\mathbf{X}}_i^T \boldsymbol{\theta}_n(U_i) - \sum_{i=1}^n \bar{q}_{i,2}\boldsymbol{\theta}_n(U_i))^T \check{\mathbf{X}}_i \check{\mathbf{X}}_i^T \boldsymbol{\theta}_n(U_i) \\
=& 2\sum_{i=1}^n \varepsilon_i \check{\mathbf{X}}_i^T \mathbf{a}_n(U_i) + \sum_{i=1}^n \bar{q}_{i,2}\mathbf{a}_n(U_i)^T \check{\mathbf{X}}_i \check{\mathbf{X}}_i^T \mathbf{a}_n(U_i) \\
& + 2\sum_{i=1}^n \varepsilon_i \check{\mathbf{X}}_i^T \boldsymbol{\theta}_n(U_i) - d_{2n} + R_n^{(1)} + R_n^{(2)} + R_n^{(3)} + o_p(h^{-1/2}) \\
=& \bar{T}_n + \bar{S}_n + 2\sum_{i=1}^n \varepsilon_i \check{\mathbf{X}}_i^T \boldsymbol{\theta}_n(U_i) - d_{2n} + R_n^{(1)} + R_n^{(2)} + R_n^{(3)} + o_p(h^{-1/2})
\end{aligned}$$

where $\bar{q}_{i,2}$ denotes $q_2\{\bar{\eta}^*_{H_0}(\mathbf{V}_i;\boldsymbol{\alpha})\}$, d_{2n} is defined in Theorem 5, and $\bar{T}_n$ and $\bar{S}_n$ are defined in the proof of Lemma S.10. As the same argument of Lemma S.10, we have

$$T^*_{NP} - \mu_n + d_{2n} = \Upsilon(n)h^{-1/2} + 2\sum_{i=1}^n \varepsilon_i \check{\mathbf{X}}_i^T \boldsymbol{\theta}_n(U_i) + o_p(h^{-1/2}),$$

where $\Upsilon(n)$ is defined in the proof of Lemma S.10. It is easy to see that the variance of the second term on the right hand is $\tilde{v}^2 = nE[\rho(\mathbf{V})\boldsymbol{\theta}_n(U)^T\check{\mathbf{X}}\check{\mathbf{X}}^T\boldsymbol{\theta}_n(U)] = O_p(h^{-1/4})$ by noting the assumption $nh^4 \to 0$ and $\boldsymbol{\theta}_n(u) = O_p(n^{-1/2}h^{-1/4})$.

$$(T^*_{NP} - \mu_n + d_{2n})/\sigma_n \xrightarrow{\mathcal{L}} N(0,1),$$

which completes the proof of Lemma S.11.

□

Proof of Theorem 5: According to Lemma S.11, this Theorem can be proved similarly to Theorem 3.

□

Bibliography

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

Cui, X., Hardle, 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.

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

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

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.