Biometrika, *pp*. 1–10
Printed in Great Britain
Advance Access publication

Supplementary material for "Functional L-Optimality Subsampling for Functional Generalized Linear Models with Massive Data "

BY H. LIU
School of Economics and Finance, Xi'an Jiaotong University, 74 West Yanta Road, Xi'an 710049, Shaanxi, China
liuhua_sufe@163.com

J. YOU
School of Statistics and Management, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai 200433, Shanghai, China
johnyou07@163.com

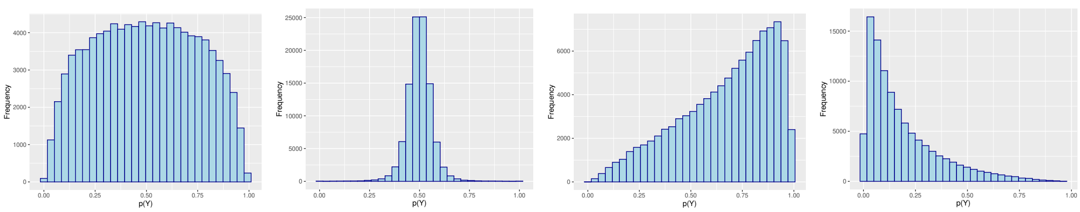
AND J. CAO
Department of Statistics and Actuarial Science, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia V5A 1S6, Canada
jiguo_cao@sfu.ca

SUMMARY

This supplementary file includes the additional figures of the simulation studies, the introduction of the Newton–Raphson algorithm for estimating the functional generalized linear model, and the detailed proofs of the theoretical results.

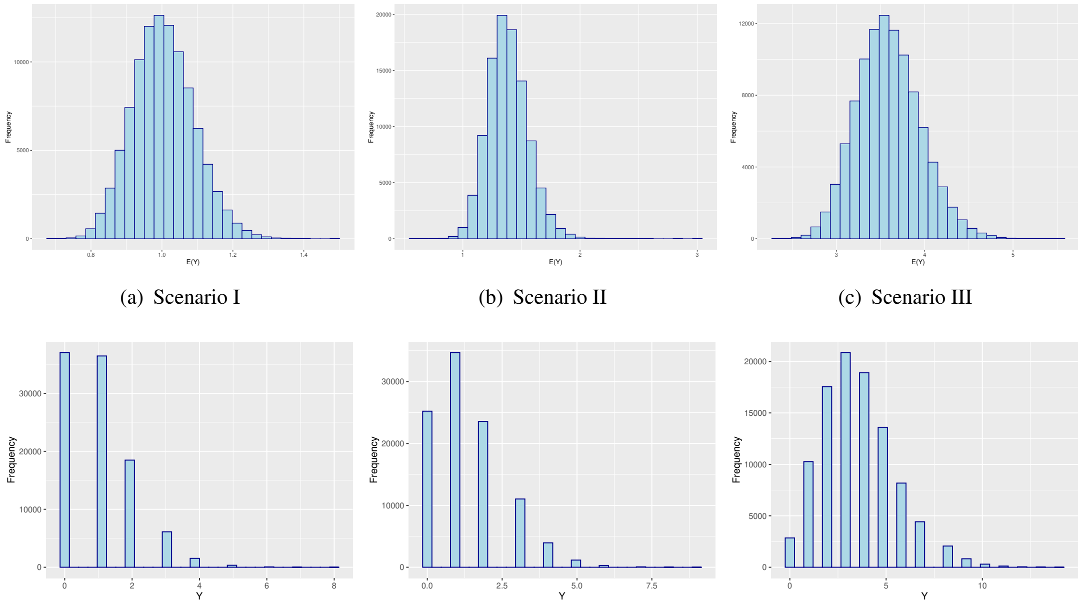
1. SIMULATION STUDIES (CONTINUED)

This section lists two additional plots (Fig S1 - S2) of the simulation I and II in the manuscript.



(a) Scenario I (b) Scenario II (c) Scenario III (d) Scenario IV

Fig. S1: The histogram of $p(x_i)$ under four scenarios in Simulation I when the full data size is $n = 10^5$.



(a) Scenario I (b) Scenario II (c) Scenario III

(d) Scenario I (e) Scenario II (f) Scenario III

Fig. S2: The histogram of $E(y_i) = \lambda(x_i)$ and y_i under three scenarios in Simulation II when the full data size is $n = 10^5$.

2. Estimation Steps for Functional Generalized Linear Model

In this section, we provide the estimation procedure for the functional generalized linear model to estimate the coefficient function $\boldsymbol{\beta}(\cdot)$ by solving the equation (1). Now we can apply the Newton–Raphson algorithm to iterativly solve it. The detailed estimation steps are given below:

Step 0. Obtain an initial estimate $\boldsymbol{c}_0$.

Step 1. At the $(k+1)$th iteration,

$$\boldsymbol{c}_{\mathrm{PQL}}^{(k+1)} = \boldsymbol{c}_{\mathrm{PQL}}^{(k)} + \dot{Q}_{\mathrm{PQL}}^{-1}(\boldsymbol{c}_{\mathrm{PQL}}^{(k)}) Q_{\mathrm{PQL}}(\boldsymbol{c}_{\mathrm{PQL}}^{(k)}). \tag{S1}$$

Step 2. Repeat Step 1 until convergence is reached.

3. Some Lemmas

To prove our theorems in the manuscripts, we start from proving the following lemmas.

Lemma S1. *Under Assumptions A1, $s_\beta(t) - \beta(t) = b_a(t) + o(K^{-d})$.*

Proof. The proof of this lemma can be found in Barrow & Smith (1978). □

Lemma S2. *Under Assumption A2, A4, A5 and A7, (i) there exists constants $C_G > c_G > 0$ such that*

$$c_G K^{-1} \le \rho_{min}(\boldsymbol{G}_{k,n}^{\psi}) \le \rho_{max}(\boldsymbol{G}_{k,n}^{\psi}) \le C_G K^{-1},$$

where ρ_{min} and ρ_{max} denote the smallest and largest eigenvalues of a matrix, respectively.

(ii) we can have $\|\boldsymbol{G}_{k,n}^{\psi}\|_\infty = O(K^{-1})$.

Proof. Using the result in (i), (ii) can be derived directly from Lemma 6.3 and Lemma 6.4 in (Shen et al., 1998). We only give the proof of (i). For any non-zero vector $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_{K+p+1})^T$ with $\|\boldsymbol{\mu}\| = 1$, note that

$$\rho_{min}(\boldsymbol{G}_{k,n}^{\psi}) = \min \boldsymbol{\mu}^T \boldsymbol{G}_{k,n}^{\psi} \boldsymbol{\mu}, \quad \rho_{max}(\boldsymbol{G}_{k,n}^{\psi}) = \max \boldsymbol{\mu}^T \boldsymbol{G}_{k,n}^{\psi} \boldsymbol{\mu}.$$

According to the definition of $\boldsymbol{G}_{k,n}^{\psi}$, we get

$$\begin{aligned} \boldsymbol{\mu}^T \boldsymbol{G}_{k,n}^{\psi} \boldsymbol{\mu} &= \frac{1}{n} \boldsymbol{\mu}^T \boldsymbol{N}^T \boldsymbol{\Psi} \boldsymbol{N} \boldsymbol{\mu} \\ &= \frac{1}{n} \sum_{i=1}^{n} \dot{\psi}(\boldsymbol{N}_i^T \boldsymbol{c}) \boldsymbol{\mu}^T \boldsymbol{N}_i \boldsymbol{N}_i^T \boldsymbol{\mu} \\ &= \frac{1}{n} \sum_{i=1}^{n} \dot{\psi}(\boldsymbol{N}_i^T \boldsymbol{c}) \left(\sum_{j=1}^{K+p+1} \mu_j N_{ij} \right)^2 . \end{aligned}$$

For $|\sum_{j=1}^{K+p+1} \mu_j N_{ij}|$, by the property $\sum_{j=1}^{K+p+1} N_{p+1,j}(t) = 1$, we have

$$\begin{aligned} \left| \sum_{j=1}^{K+p+1} \mu_j N_{ij} \right| &= \left| \int \sum_{j=1}^{K+p+1} \mu_j N_{p+1,j}(t) z_i(t) dt \right| \\ &\leq \int \left| \sum_{j=1}^{K+p+1} \mu_j N_{p+1,j}(t) z_i(t) \right| dt \\ &\leq \int \left\{ \left(\sum_{j=1}^{K+p+1} \mu_j^2 N_{p+1,j}(t) z_i^2(t) \right)^{1/2} \left(\sum_{j=1}^{K+p+1} N_{p+1,j}(t) \right)^{1/2} \right\} dt \\ &= \int \left\{ \left(\sum_{j=1}^{K+p+1} \mu_j^2 N_{p+1,j}(t) z_i^2(t) \right)^{1/2} \right\} dt \\ &\leq \int \left(\sum_{j=1}^{K+p+1} |\mu_j| |z_i(t)| N_{p+1,j}^{1/2}(t) \right) dt \\ &= \sum_{j=1}^{K+p+1} |\mu_j| \int |z_i(t)| N_{p+1,j}^{1/2}(t) dt \\ &\leq \sum_{j=1}^{K+p+1} |\mu_j| \left(\int z_i^2(t) dt \right)^{1/2} \left(\int N_{p+1,j}(t) dt \right)^{1/2} . \end{aligned}$$

By the property $\int N_{p+1,j}(t)dt = O(K^{-1})$ and Assumption A2, one has $\left| \sum_{j=1}^{K+p+1} \mu_j N_{ij} \right| = O(K^{-1/2})$. Next, by Assumption A2, $\boldsymbol{\mu}^T \boldsymbol{G}_{k,n}^{\psi} \boldsymbol{\mu} \leq C_G K^{-1}$. Applying the second condition in Assumption A4, $c_G K^{-1} \leq \rho_{min}(\boldsymbol{G}_{k,n}^{\psi})$ holds as well. This completes the proof of (i). □

LEMMA S3. *Under Assumptions A3 and A4, we can get (i)* $\|\boldsymbol{D}_q\|_\infty = O(K^{2q-1})$*; (ii) there are two positive constants* $c_D < C_D$ *such that* $c_D K^{2q-1}\|\boldsymbol{\mu}\|_2^2 \le \boldsymbol{\mu}^T \boldsymbol{D}_q \boldsymbol{\mu} \le C_D K^{2q-1}\|\boldsymbol{\mu}\|_2^2$.

Proof. (i) Let $\boldsymbol{N}_{p+1-q}(t) = (N_{j,p+1-q}(t) : -p+q \le j \le K)^T$, from the derivative formula for B-spline functions (de Boor, 1978), we can get

$$s_\beta^{(q)}(t) = \boldsymbol{N}_{p+1-q}^T(t)\boldsymbol{c}^{(q)},$$

where $\boldsymbol{c}^{(q)}$ are defined recursively via

$$\boldsymbol{c}^{(1)} = \boldsymbol{\nabla}_1 \boldsymbol{c},$$
$$\boldsymbol{c}^{(q)} = \boldsymbol{\nabla}_q \boldsymbol{c}^{(q-1)},$$

where

$$\boldsymbol{\nabla}_1 = (p+1-1) \times \begin{pmatrix} \frac{-1}{k_1 - k_{-p+1}} & \frac{1}{k_1 - k_{-p+1}} & 0 & \cdots & 0 \\ 0 & \frac{-1}{k_{-p+2} - k_{-p+3}} & \frac{1}{k_{-p+2} - k_{-p+3}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{-1}{k_{K+p} - k_K} & \frac{1}{k_{K+p} - k_K} \end{pmatrix}_{(K+p+1-1)\times(K+p+1)},$$

$$\boldsymbol{\nabla}_q = (p+1-q) \times \begin{pmatrix} \frac{-1}{k_1 - k_{-p+q}} & \frac{1}{k_1 - k_{-p+q}} & 0 & \cdots & 0 \\ 0 & \frac{-1}{k_{-q+1} - k_{-p+2}} & \frac{1}{k_{-q+1} - k_{-p+2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{-1}{k_{K+p+1-q} - k_K} & \frac{1}{k_{K+p+1-q} - k_K} \end{pmatrix}_{(K+p+1-q)\times(K+p+1)}.$$

So, we can rewrite the second penalty term in (1) as $\lambda \boldsymbol{c}^T \boldsymbol{\Delta}_q^T \boldsymbol{R} \boldsymbol{\Delta}_q \boldsymbol{c}$, where the matrix $\boldsymbol{R} = \int_a^b \boldsymbol{N}_{p+1-q}(t)\boldsymbol{N}_{p+1-q}^T(t)dt$ and $\boldsymbol{\Delta}_q = \boldsymbol{\nabla}_1 \ldots \boldsymbol{\nabla}_q$. Note that (A1) implies that $\delta \sim K^{-1}$, i.e., δ and K^{-1} are rate-wise equivalent. Moreover, by definition $\|\Delta_l\|_\infty = O(K), l = 1, \ldots, q$, thus, $\|\boldsymbol{\Delta}_q\|_\infty = O(\delta^{-q}) = O(K^q)$. And, from the definition of B-spline, we can get $\|\boldsymbol{R}\|_\infty = O(K^{-1})$. Because $\boldsymbol{D}_q = \boldsymbol{\Delta}_q^T \boldsymbol{R} \boldsymbol{\Delta}_q$, the desired result holds. For (ii), it can be derived from Lemma 6.1 in Cardot et al. (2003) and the proof of it is omitted. □

LEMMA S4. *Under Assumption A2-A5 and A7, (i) there are some positive constants* $c_H < C_H$ *such that*

$$c_H K^{-1} \le \rho_{min}(\boldsymbol{H}_{k,n}^\psi) \le \rho_{max}(\boldsymbol{H}_{k,n}^\psi) \le C_H K^{-1}.$$

(ii) $\|\boldsymbol{H}_{k,n}^\psi\|_\infty = O(K^{-1})$.

Proof. Under Assumption A3, it can be directly derived from Lemma S2 and S3. □

LEMMA S5. *Under Assumptions A2, A4, A6 and A7, as* $n, L \to \infty$*, in probability,*

$$\frac{1}{n}\sum_{i=1}^n \frac{R_i}{Lp_i}\dot{\psi}(\boldsymbol{N}_i^T\boldsymbol{c})\boldsymbol{N}_i\boldsymbol{N}_i^T - \frac{1}{n}\sum_{i=1}^n \dot{\psi}(\boldsymbol{N}_i^T\boldsymbol{c})\boldsymbol{N}_i\boldsymbol{N}_i^T = o_{p|\mathcal{F}_n}(1). \quad \text{(S2)}$$

Proof. Direct calculation shows that conditionally on $\mathcal{F}_n$,

$$\mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{R_i}{Lp_i}\dot{\psi}(\boldsymbol{N}_i^T\boldsymbol{c})\boldsymbol{N}_i\boldsymbol{N}_i^T|\mathcal{F}_n\right\}=\frac{1}{n}\sum_{i=1}^{n}\dot{\psi}(\boldsymbol{N}_i^T\boldsymbol{c})\boldsymbol{N}_i\boldsymbol{N}_i^T.$$

Let $\boldsymbol{G}_{k,n}^{\psi,*}=n^{-1}\sum_{i=1}^{n}R_i\dot{\psi}(\boldsymbol{N}_i^T\boldsymbol{c})\boldsymbol{N}_i\boldsymbol{N}_i^T/(Lp_i)$, then $\boldsymbol{G}_{k,n}^{\psi,*}-\boldsymbol{G}_{k,n}^{\psi}=n^{-1}\sum_{i=1}^{n}(R_i-Lp_i)\dot{\psi}(\boldsymbol{N}_i^T\boldsymbol{c})\boldsymbol{N}_i\boldsymbol{N}_i^T/(Lp_i)$. For any component of $\boldsymbol{G}_{k,n}^{\psi,*}-\boldsymbol{G}_{k,n}^{\psi}$, we have

$$\begin{aligned}
&\mathrm{E}\left\{\boldsymbol{G}_{k,n}^{\psi,*j_1j_2}-\boldsymbol{G}_{k,n}^{\psi,j_1j_2}|\mathcal{F}_n\right\}^2\\
=&\frac{1}{n^2L}\sum_{i=1}^{n}\frac{1-p_i}{p_i}\left\{\dot{\psi}(\boldsymbol{N}_i^T\boldsymbol{c})\left(\boldsymbol{N}_i\boldsymbol{N}_i^T\right)^{j_1j_2}\right\}^2\\
\leq&\frac{1}{n^2L}\sum_{i=1}^{n}\frac{1}{p_i}\left\{\dot{\psi}(\boldsymbol{N}_i^T\boldsymbol{c})\left(\boldsymbol{N}_i\boldsymbol{N}_i^T\right)^{j_1j_2}\right\}^2\\
\leq&\max(\frac{1}{nLp_i})\frac{1}{n}\sum_{i=1}^{n}\left\{\dot{\psi}(\boldsymbol{N}_i^T\boldsymbol{c})\left(\boldsymbol{N}_i\boldsymbol{N}_i^T\right)^{j_1j_2}\right\}^2\\
=&o_p(\frac{1}{\sqrt{L}})O(K^{-2})=o_p(1).
\end{aligned}$$

Thus, through Chebyshev's inequality, the desired conclusion follows. □

LEMMA S6. *Under Assumptions A1 - A7, conditional on $\mathcal{F}_n$, as $n, L\to\infty$,*

$$\frac{\sqrt{L}}{n}\boldsymbol{V}_p^{\psi\,-1/2}Q^*_{\mathrm{PQL}}(\widehat{\boldsymbol{c}}_{\mathrm{PQL}})\to N(0,1),$$

in distribution.

Proof. Direct calculation shows that

$$\mathrm{E}\left\{\frac{1}{n}Q^*_{\mathrm{PQL}}(\widehat{\boldsymbol{c}}_{\mathrm{PQL}})|\mathcal{F}_n\right\}=\frac{1}{n}Q_{\mathrm{PQL}}(\widehat{\boldsymbol{c}}_{\mathrm{PQL}})=0,$$

and

$$\begin{aligned}
&\mathrm{Var}\left\{\frac{1}{n}Q^*_{\mathrm{PQL}}(\widehat{\boldsymbol{c}}_{\mathrm{PQL}})|\mathcal{F}_n\right\}\\
=&\frac{1}{n^2L}\sum_{i=1}^{n}\frac{1}{p_i}\left\{y_i-\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})\right\}^2\boldsymbol{N}_i\boldsymbol{N}_i^T-\frac{1}{n^2L}\sum_{i=1}^{n}\left\{y_i-\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})\right\}^2\boldsymbol{N}_i\boldsymbol{N}_i^T\\
=&\frac{1}{L}\boldsymbol{V}_p^{\psi}-o_p(1).
\end{aligned}$$

Next, we check the Lindeberg-Feller condition under the conditional distribution. Denote $\boldsymbol{\varpi}_i = \frac{1}{n}\left\{\frac{R_i}{Lp_i}\left\{y_i - \psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})\right\}\boldsymbol{N}_i - \lambda\boldsymbol{D}\widehat{\boldsymbol{c}}_{\mathrm{PQL}}\right\}$. For any $\epsilon > 0$,

$$
\begin{aligned}
&\sum_{i=1}^{n}\mathrm{E}\left\{\|\boldsymbol{\varpi}_i\|^2\boldsymbol{I}(\|\boldsymbol{\varpi}_i\| > \epsilon)|\mathcal{F}_n\right\}\\
\leq&\frac{1}{\epsilon}\sum_{i=1}^{n}\mathrm{E}\left\{\|\boldsymbol{\varpi}_i\|^3|\mathcal{F}_n\right\}\\
\leq&\frac{1}{n^3}\sum_{i=1}^{n}\left\{\mathrm{E}(\frac{R_i^3|y_i-\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})|^3\|\boldsymbol{N}_i\|^3}{L^3p_i^3}|\mathcal{F}_n) + \|\lambda\boldsymbol{D}\widehat{\boldsymbol{c}}_{\mathrm{PQL}}\|^3\right\}\\
=&\frac{1}{n^3}\sum_{i=1}^{n}\frac{(L(L-1)(L-2)p_i^3 + 3L(L-1)p_i^2 + Lp_i)|y_i-\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})|^3\|\boldsymbol{N}_i\|^3}{L^3p_i^3} + o(K^{-3})\\
=&o_p(1),
\end{aligned}
$$

where the last equality holds by Assumptions A1 - A7. By Lindeberg-Feller central limit theorem, the desired result follows. □

LEMMA S7. *Under Assumptions A1 - A7, as $n, L \to \infty$,*

$$
\frac{\sqrt{L}}{n}\boldsymbol{W}_p^{\psi\,-1/2}Q^*_{\mathrm{PQL}}(\boldsymbol{c}) \to N(0,1),
$$

in distribution, where

$$
\boldsymbol{W}_p^{\psi} = \frac{1}{n^2}\sum_{i=1}^{n}\frac{E(y_i-\psi(\boldsymbol{N}_i^T\boldsymbol{c}))^2\boldsymbol{N}_i\boldsymbol{N}_i^T}{p_i}.
$$

Proof. Note that

$$
\begin{aligned}
&\mathrm{E}\left\{\frac{1}{n}Q^*_{\mathrm{PQL}}(\boldsymbol{c})\right\} = \mathrm{E}\left\{\mathrm{E}\left\{\frac{1}{n}Q^*_{\mathrm{PQL}}(\boldsymbol{c})|\mathcal{F}_n\right\}\right\}\\
=&\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{N}_i(\psi(<\boldsymbol{z}_i,\boldsymbol{\beta}>) - \psi(\boldsymbol{N}_i^T\boldsymbol{c})) - \frac{1}{n}\lambda\boldsymbol{D}\boldsymbol{c}\\
=&o((nK)^{-1/2}),
\end{aligned}
$$

and

$$
\mathrm{Var}\left\{\frac{1}{n}Q^*(\boldsymbol{c})\right\} = \mathrm{E}\left\{\mathrm{Var}\left\{\frac{1}{n}Q^*(\boldsymbol{c})|\mathcal{F}_n\right\}\right\} + \mathrm{Var}\left\{\mathrm{E}\left\{\frac{1}{n}Q^*(\boldsymbol{c})|\mathcal{F}_n\right\}\right\}. \tag{S3}
$$

For the first term in (S3), we can get

$$
\begin{aligned}
\mathrm{E}\left\{\mathrm{Var}\left\{\frac{1}{n}Q^*_{\mathrm{PQL}}(\boldsymbol{c})|\mathcal{F}_n\right\}\right\} &= \mathrm{E}\left\{\frac{1}{n^2L}\sum_{i=1}^{n}\frac{(y_i-\psi(\boldsymbol{N}_i^T\boldsymbol{c}))^2\boldsymbol{N}_i\boldsymbol{N}_i^T}{p_i} - \frac{1}{n^2L}\sum_{i=1}^{n}(y_i-\psi(\boldsymbol{N}_i^T\boldsymbol{c}))^2\boldsymbol{N}_i\boldsymbol{N}_i^T\right\}\\
&= \frac{1}{n^2L}\sum_{i=1}^{n}\frac{\mathrm{E}(y_i-\psi(\boldsymbol{N}_i^T\boldsymbol{c}))^2\boldsymbol{N}_i\boldsymbol{N}_i^T}{p_i} - \frac{1}{n^2L}\sum_{i=1}^{n}\mathrm{E}(y_i-\psi(\boldsymbol{N}_i^T\boldsymbol{c}))^2\boldsymbol{N}_i\boldsymbol{N}_i^T\\
&= \frac{1}{L}\boldsymbol{W}_p^{\psi} - O(\frac{1}{nLK}).
\end{aligned}
$$

Similarly, we deal with the second term in (S3),

$$
\begin{aligned}
\mathrm{Var}\left\{\mathrm{E}\left\{\frac{1}{n}Q^*_{\mathrm{PQL}}(\boldsymbol{c})|\mathcal{F}_n\right\}\right\} &= \mathrm{Var}\left\{\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{N}_i(y_i-\psi(\boldsymbol{N}_i^T\boldsymbol{c}))-\lambda\boldsymbol{D}\boldsymbol{c}\right\}\\
&= \frac{1}{n^2}\sum_{i=1}\sigma^2\boldsymbol{N}_i\boldsymbol{N}_i^{`}\\
&= O(\frac{1}{nK}).
\end{aligned}
$$

Under Assumption A1 - A7, $\frac{1}{L}\boldsymbol{W}_p^{\psi} = o(K^{-1}L^{-1/2})$. Consequently,

$$
\mathrm{Var}\left\{\frac{1}{n}Q^*_{\mathrm{PQL}}(\boldsymbol{c})\right\} = \frac{1}{L}\boldsymbol{W}_p^{\psi} + O(\frac{1}{nK}).
$$

Next, we check the Lindeberg-Feller condition, denote

$$
\boldsymbol{\zeta}_i = \frac{1}{n}\left\{\frac{R_i}{Lp_i}\left\{y_i-\psi(\boldsymbol{N}_i^T\boldsymbol{c})\right\}\boldsymbol{N}_i-\lambda\boldsymbol{D}\boldsymbol{c}\right\}
$$

. For any $\epsilon > 0$,

$$
\begin{aligned}
&\sum_{i=1}^{n}\mathrm{E}\left\{\|\boldsymbol{\zeta}_i\|^2\boldsymbol{I}(\|\boldsymbol{\zeta}_i\|>\epsilon)\right\}\\
\le&\frac{1}{\epsilon}\sum_{i=1}^{n}\mathrm{E}\left\{\|\boldsymbol{\zeta}_i\|^3\right\}\\
\le&\frac{1}{n^3}\sum_{i=1}^{n}\left\{\mathrm{E}(\frac{R_i^3|y_i-\psi(\boldsymbol{N}_i^T\boldsymbol{c})|^3\|\boldsymbol{N}_i\|^3}{L^3p_i^3})+\|\lambda\boldsymbol{D}\boldsymbol{c}\|^3\right\}\\
=&\frac{1}{n^3}\sum_{i=1}^{n}\frac{(L(L-1)(L-2)p_i^3+3L(L-1)p_i^2+Lp_i)\mathrm{E}|y_i-\psi(\boldsymbol{N}_i^T\boldsymbol{c})|^3\|\boldsymbol{N}_i\|^3}{L^3p_i^3}+o(K^{-3})\\
=&o_p(1),
\end{aligned}
$$

where the last equality holds by Assumptions A1 - A7. By Lindeberg-Feller central limit theorem, the desired result follows. □

4. Proofs of Main Results

In this section, we give the proofs of the theoretical results.

Proof of Theorem 1. By Taylor expansion, we can get

$$
0 = \frac{1}{n}Q^*_{\mathrm{PQL}}(\widetilde{\boldsymbol{c}}_{\mathrm{PQL}}) = \frac{1}{n}Q^*_{\mathrm{PQL}}(\boldsymbol{c}) + \frac{1}{n}\dot{Q}^*_{\mathrm{PQL}}(\boldsymbol{c})(\widetilde{\boldsymbol{c}}_{\mathrm{PQL}}-\boldsymbol{c}) + o_p(1),
$$

where $\dot{Q}^*_{\mathrm{PQL}}(\boldsymbol{c})$ is the first derivations of $Q^*_{\mathrm{PQL}}(\boldsymbol{c})$ about $\boldsymbol{c}$, and $\dot{Q}^*_{\mathrm{PQL}}(\boldsymbol{c}) = -\sum_{i=1}^{n}\frac{R_i}{Lp_i}\dot{\psi}(\boldsymbol{N}^T\boldsymbol{c})\boldsymbol{N}_i\boldsymbol{N}_i^T-\lambda\boldsymbol{D} = -n\boldsymbol{G}_{k,n}^{\psi,*}-\lambda\boldsymbol{D} \triangleq -n\boldsymbol{H}_{k,n}^{\psi,*}$.

It is obviously that

$$
\widetilde{\boldsymbol{c}}_{\mathrm{PQL}}-\boldsymbol{c} = \left\{-\frac{1}{n}\dot{Q}^*_{\mathrm{PQL}}(\boldsymbol{c})\right\}^{-1}\frac{1}{n}Q^*_{\mathrm{PQL}}(\boldsymbol{c}) + \left\{-\frac{1}{n}\dot{Q}^*_{\mathrm{PQL}}(\boldsymbol{c})\right\}^{-1}\cdot o_p(1),
$$

and given t, we have

$$\begin{aligned}\widetilde{\boldsymbol{\beta}}_{\mathrm{PQL}}(t)-\boldsymbol{\beta}(t)&=\boldsymbol{N}^T(t)\left\{-\frac{1}{n}\dot{Q}^*_{\mathrm{PQL}}(\boldsymbol{c})\right\}^{-1}\frac{1}{n}Q^*_{\mathrm{PQL}}(\boldsymbol{c})+\boldsymbol{N}^T(t)\left\{-\frac{1}{n}\dot{Q}^*_{\mathrm{PQL}}(\boldsymbol{c})\right\}^{-1}\cdot o_p(1)+s_\beta(t)-\beta(t)\\&=\boldsymbol{N}^T(t)\left\{-\frac{1}{n}\dot{Q}^*_{\mathrm{PQL}}(\boldsymbol{c})\right\}^{-1}\frac{\boldsymbol{W}_p^{\psi\,1/2}}{\sqrt{L}}\sqrt{L}\boldsymbol{W}_p^{\psi\,-1/2}\frac{1}{n}Q^*_{\mathrm{PQL}}(\boldsymbol{c})\\&\quad+\boldsymbol{N}^T(t)\left\{-\frac{1}{n}\dot{Q}^*_{\mathrm{PQL}}(\boldsymbol{c})\right\}^{-1}\cdot o_p(1)+s_\beta(t)-\beta(t).\end{aligned}\tag{S4}$$

From Lemma S5, we can get that $\boldsymbol{G}^{\psi,*}_{k,n}-\boldsymbol{G}^{\psi}_{k,n}\to 0$ and $\boldsymbol{H}^{\psi,*}_{k,n}-\boldsymbol{H}^{\psi}_{k,n}\to 0$ as $n,L\to\infty$. Thus, (S4) can be written as

$$\widetilde{\boldsymbol{\beta}}_{\mathrm{PQL}}(t)-\boldsymbol{\beta}(t)=\boldsymbol{N}^T(t)\boldsymbol{H}^{\psi}_{k,n}\frac{\boldsymbol{W}_p^{\psi\,1/2}}{\sqrt{L}}\sqrt{L}\boldsymbol{W}_p^{\psi\,-1/2}\frac{1}{n}Q^*_{\mathrm{PQL}}(\boldsymbol{c})+\boldsymbol{N}^T(t)\boldsymbol{H}^{\psi}_{k,n}\cdot o_p(1)+s_\beta(t)-\beta(t).\tag{S5}$$

For Theorem 1, applying (S5), as $n,L\to\infty$, it holds that

$$\begin{aligned}&(\boldsymbol{N}^T(t)(\boldsymbol{H}^{\psi}_{k,n})^{-1}\boldsymbol{W}_p^{\psi}(\boldsymbol{H}^{\psi}_{k,n})^{-1}\boldsymbol{N}(t))^{-1/2}\sqrt{L}(\widetilde{\boldsymbol{\beta}}_{\mathrm{PQL}}(t)-\boldsymbol{\beta}(t))\\=&(\boldsymbol{N}^T(t)(\boldsymbol{H}^{\psi}_{k,n})^{-1}\boldsymbol{W}_p^{\psi}(\boldsymbol{H}^{\psi}_{k,n})^{-1}\boldsymbol{N}(t))^{-1/2}\sqrt{L}\boldsymbol{N}^T(t)\boldsymbol{H}^{\psi}_{k,n}\frac{\boldsymbol{W}_p^{\psi\,1/2}}{\sqrt{L}}\sqrt{L}\boldsymbol{W}_p^{\psi\,-1/2}\frac{1}{n}Q^*_{\mathrm{PQL}}(\boldsymbol{c})+o_p(1).\end{aligned}$$

Therefore, from Assumption A6, Lemma S7 and Slutsky's theorem, we conclude that as $n,L\to\infty$, with probability approaching one,

$$(\boldsymbol{N}^T(t)(\boldsymbol{H}^{\psi}_{k,n})^{-1}\boldsymbol{W}_p^{\psi}(\boldsymbol{H}^{\psi}_{k,n})^{-1}\boldsymbol{N}(t))^{-1/2}\sqrt{L}(\widetilde{\boldsymbol{\beta}}_{\mathrm{PQL}}(t)-\boldsymbol{\beta}(t))\to\mathbb{N}(\boldsymbol{0}_2,\boldsymbol{I}_2).$$

Proof Proof of Theorem 2. With Lemma S5 and S6, and the similar reasoning as the proof for Theorem 1, we can prove Theorem 2. So, the details are omitted. □

Proof Proof of Theorem 3. Note that

$$\begin{aligned}\mathrm{tr}(\boldsymbol{V}_p^{\psi})&=\mathrm{tr}\left[\frac{1}{n^2}\sum_{i=1}^n\frac{\{y_i-\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})\}^2\boldsymbol{N}_i\boldsymbol{N}_i^T}{p_i}\right]\\&=\frac{1}{n^2}\sum_{i=1}^n\mathrm{tr}\left[\frac{\{y_i-\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})\}^2\boldsymbol{N}_i\boldsymbol{N}_i^T}{p_i}\right]\\&=\frac{1}{n^2}\sum_{i=1}^n\frac{\{y_i-\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})\}^2\|\boldsymbol{N}_i\|_2^2}{p_i}\\&=\frac{1}{n^2}(\sum_{i=1}^n p_i)\sum_{i=1}^n\frac{\{y_i-\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})\}^2\|\boldsymbol{N}_i\|_2^2}{p_i}\\&\geq\frac{1}{n^2}\left\{\sum_{i=1}^n|y_i-\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})|\|\boldsymbol{N}_i\|_2\right\}^2,\end{aligned}$$

where the last step is from the Cauchy-Schwarz inequality and the equality holds if and only if when $p_i \propto |y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})| \|\boldsymbol{N}_i\|_2$. And, to satisfy $\sum_{i=1}^n p_i = 1$, we let $p_i = \frac{|y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})| \|\boldsymbol{N}_i\|_2}{\sum_{i=1}^n |y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})| \|\boldsymbol{N}_i\|_2}$. □

Proof Proof of Theorem 4. Following the proof of Lemma S7, when $p_i = p_{\mathrm{PQL},i}^{\mathrm{FLoS},\widehat{\boldsymbol{c}}^0_{\mathrm{PQL}}}$, we have

$$
\begin{aligned}
\boldsymbol{V}_p^{\psi} &= \frac{1}{n^2} \sum_{i=1}^n \frac{(y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}}))^2 \boldsymbol{N}_i \boldsymbol{N}_i^T}{p_i} \\
&= \frac{1}{n} \sum_{i=1}^n \frac{\{y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})\}^2 \boldsymbol{N}_i \boldsymbol{N}_i^T}{|y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}})| \|\boldsymbol{N}_i\|_2} \times \frac{1}{n} \sum_{i=1}^n |y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}})| \|\boldsymbol{N}_i\|_2 \\
&\equiv \widetilde{\boldsymbol{V}}_1^{\psi} \times \widetilde{\boldsymbol{V}}_2^{\psi}.
\end{aligned}
$$

Now we want to prove that $\widetilde{\boldsymbol{V}}_1^{\psi} - \boldsymbol{V}_1^{\psi} = o_p(1)$ and $\widetilde{\boldsymbol{V}}_2^{\psi} - \boldsymbol{V}_2^{\psi} = o_p(1)$, where $\boldsymbol{V}_1^{\psi}$ and $\boldsymbol{V}_2^{\psi}$ have the same expression of $\widetilde{\boldsymbol{V}}_1^{\psi}$ and $\widetilde{\boldsymbol{V}}_2^{\psi}$, respectively, except that $y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}})$ is replaced by $y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})$. Combining $\widehat{\boldsymbol{c}}_{\mathrm{PQL}} - \boldsymbol{c} = o_p(1)$ with $\widehat{\boldsymbol{c}}^0_{\mathrm{PQL}} - \boldsymbol{c} = o_p(1)$, we can get $\widehat{\boldsymbol{c}}^0_{\mathrm{PQL}} - \widehat{\boldsymbol{c}}_{\mathrm{PQL}} = o_p(1)$, thus, $|\psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}}) - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})| = o_p(1)$ for each i. Therefore, $\mathrm{E}\left\{|\psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}}) - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})|\right\} \to 0$ and $\frac{|y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})|}{|y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}})|} \leq C_1$, where C_1 is a positive constant. Note that

$$
\begin{aligned}
|\widetilde{\boldsymbol{V}}_1^{\psi} - \boldsymbol{V}_1^{\psi}| &= \frac{1}{n} \left| \sum_{i=1}^n \frac{\{y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})\}^2 \boldsymbol{N}_i \boldsymbol{N}_i^T}{|y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}})| \|\boldsymbol{N}_i\|_2} - \sum_{i=1}^n \frac{\{y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})\} \boldsymbol{N}_i \boldsymbol{N}_i^T}{\|\boldsymbol{N}_i\|_2} \right| \\
&\leq \frac{1}{n} \sum_{i=1}^n \left| \frac{(y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}}))(\psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}}) - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})) \boldsymbol{N}_i \boldsymbol{N}_i^T}{|y_i - \boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0| \|\boldsymbol{N}_i\|_2} \right| \\
&\leq \frac{C_1}{n} \sum_{i=1}^n \left| \frac{(\psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}}) - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})) \boldsymbol{N}_i \boldsymbol{N}_i^T}{\|\boldsymbol{N}_i\|_2} \right| \\
&\leq \frac{C_1}{n} \sum_{i=1}^n |(\psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}}) - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}}))| \left| \frac{\boldsymbol{N}_i \boldsymbol{N}_i^T}{\|\boldsymbol{N}_i\|_2} \right|.
\end{aligned}
$$

And

$$
\begin{aligned}
|\widetilde{\boldsymbol{V}}_2^{\psi} - \boldsymbol{V}_2^{\psi}| &= \frac{1}{n} \left| \sum_{i=1}^n |y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}})| \|\boldsymbol{N}_i\|_2 - \sum_{i=1}^n |y_i - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})| \|\boldsymbol{N}_i\|_2 \right| \\
&\leq \frac{1}{n} \sum_{i=1}^n \left| (\psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}}) - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})) \|\boldsymbol{N}_i\|_2 \right|.
\end{aligned}
$$

For any $\epsilon > 0$, by Chebyshev's inquality

$$
\begin{aligned}
& P\left\{ \frac{1}{n} \sum_{i=1}^n |\psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}}) - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})| \left| \frac{\boldsymbol{N}_i \boldsymbol{N}_i^T}{\|\boldsymbol{N}_i\|_2} \right| > \epsilon \right\} \\
\leq & \frac{1}{\epsilon n} \sum_{i=1}^n \mathrm{E}\left\{|\psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}^0_{\mathrm{PQL}}) - \psi(\boldsymbol{N}_i^T \widehat{\boldsymbol{c}}_{\mathrm{PQL}})|\right\} \left| \frac{\boldsymbol{N}_i \boldsymbol{N}_i^T}{\|\boldsymbol{N}_i\|_2} \right| \\
\to & 0.
\end{aligned}
$$

and

$$
\begin{aligned}
&P\left\{\frac{1}{n}\sum_{i=1}^{n}|\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}}^0)-\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})|\|\boldsymbol{N}_i\|_2>\epsilon\right\}\\
\leq&\frac{1}{\epsilon n}\sum_{i=1}^{n}\mathrm{E}\left\{|\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}}^0)-\psi(\boldsymbol{N}_i^T\widehat{\boldsymbol{c}}_{\mathrm{PQL}})|\right\}\|\boldsymbol{N}_i\|_2\\
\to&0.
\end{aligned}
$$

Thus, $\widetilde{\boldsymbol{V}}_1^{\psi}-\boldsymbol{V}_1=o_p(1)$, $\widetilde{\boldsymbol{V}}_2^{\psi}-\boldsymbol{V}_2^{\psi}=o_p(1)$ and $\widetilde{\boldsymbol{V}}_1^{\psi}\times\widetilde{\boldsymbol{V}}_2-\boldsymbol{V}_1^{\psi}\times\boldsymbol{V}_2^{\psi}=\widetilde{\boldsymbol{V}}_1^{\psi}\times\widetilde{\boldsymbol{V}}_2^{\psi}-\boldsymbol{V}_{\mathrm{FLoS}}^{\psi}=o_p(1)$.

Next, using the similar proof with the proof for Theorem 2, we can finish the proof of Theorem 4. □

REFERENCES

BARROW, D. & SMITH, P. (1978). Asymptotic properties of best $l_2[0,1]$ approximation by splines with variable knots. *Quarterly of applied mathematics* **36**, 293–304.

CARDOT, H., FERRATY, F. & SARDA, P. (2003). Spline estimators for the functional linear model. *Statistica Sinica* **13**, 571–591.

DE BOOR, C. (1978). *A Practical Guide to Splines*. New York: Springer.

SHEN, X., WOLFE, D. & ZHOU, S. (1998). Local asymptotics for regression splines and confidence regions. *The Annals of Statistics* **26**, 1760–1782.

[*Received on. Editorial decision on*]