Supplemental Material for "Estimation for Extreme Conditional Quantiles of Functional Quantile Regression"

Hanbing Zhu[1], Riquan Zhang[1], Yehua Li[2] and Weixin Yao[2]

[1] *East China Normal University and* [2] *University of California, Riverside*

We begin by defining notations to be used in the proofs. Let $\xi_i = (1, \xi_{i1}, \ldots, \xi_{im_n})^\top$, $\widehat{\xi}_i = \left(1, \widehat{\xi}_{i1}, \ldots, \widehat{\xi}_{im_n}\right)^\top$, $\widetilde{\theta}_j = (\alpha(\tau_j), \beta_1, \ldots, \beta_{m_n})^\top$, $A_j = (a_j, b_1, \ldots, b_{m_n})^\top$, and $\zeta_i = \sum_{j=m_n+1}^{\infty} \beta_j \xi_{ij}$. Then the objective function in (2.1) can be rewritten as

$$\sum_{j=1}^{l} \sum_{i=1}^{n} \rho_{\tau_j} \left(\xi_i^\top \widetilde{\theta}_j - \widehat{\xi}_i^\top A_j + \zeta_i + \varepsilon_{i\tau_j} \right), \tag{1}$$

where $\varepsilon_{i\tau_j} = Y_i - \alpha(\tau_j) - \langle X_i, \beta_0 \rangle$. Let

$$W_i = \left(\xi_i - \widehat{\xi}_i \right)^\top \cdot \left(0, \beta^\top \right)^\top + \zeta_i, \tag{2}$$

$\Lambda = \mathrm{diag}(1, \lambda_1, \ldots, \lambda_{m_n})$, $\theta = \left(n^{1/2}(a_1 - \alpha(\tau_1)), n^{1/2}(a_2 - \alpha(\tau_2)), \ldots, n^{1/2}(a_l - \alpha(\tau_l)), U^\top \right)^\top$, $\widetilde{\xi}_i = n^{-1/2} \Lambda^{-1/2} \widehat{\xi}_i$, $\theta_j = \left(n^{1/2}(a_j - \alpha(\tau_j)), U^\top \right)^\top$, where $\beta = (\beta_1, \ldots, \beta_{m_n})^\top$,

$$\mathrm{diag}(1, \lambda_1, \ldots, \lambda_{m_n}) = \begin{pmatrix} 1 & & & \\ & \lambda_1 & & \\ & & \ddots & \\ & & & \lambda_{m_n} \end{pmatrix},$$

$U=(u_1,\ldots,u_{m_n})^\top = n^{1/2}\Lambda_1^{1/2}(b-\beta)$, $b=(b_1,\ldots,b_{m_n})^\top$, and $\Lambda_1=\operatorname{diag}(\lambda_1,\ldots,\lambda_{m_n})$. Let

$$S_n(\theta)=\sum_{j=1}^{l}\sum_{i=1}^{n}\left\{\rho_{\tau_j}\left(W_i+\varepsilon_{i\tau_j}-\widetilde{\xi}_i^\top\theta_j\right)-\rho_{\tau_j}\left(W_i+\varepsilon_{i\tau_j}\right)\right\}. \tag{3}$$

It's easy to see that minimizing (1) with respect to A_j is equivalent to minimizing (3) over θ. Let $\psi_\tau(u)=\tau-\mathbf{1}(u<0)$, $\Omega=\{X_1,\ldots,X_n\}$, $S_{nij}=\rho_{\tau_j}\left(W_i+\varepsilon_{i\tau_j}-\widetilde{\xi}_i^\top m_n^{1/2}\theta_j\right)-\rho_{\tau_j}\left(W_i+\varepsilon_{i\tau_j}\right)$, $\Gamma_{nij}=\mathbb{E}(S_{nij}|\Omega)$, $R_{nij}=S_{nij}-\Gamma_{nij}+m_n^{1/2}\widetilde{\xi}_i^\top\theta_j\psi_{\tau_j}\left(\varepsilon_{i\tau_j}\right)$, $S_{nj}=\sum_{i=1}^n S_{nij}$, $\Gamma_{nj}=\sum_{i=1}^n\Gamma_{nij}$, $R_{nj}=\sum_{i=1}^n R_{nij}$, $S_n=\sum_{j=1}^l S_{nj}$, $\Gamma_n=\sum_{j=1}^l\Gamma_{nj}$, and $R_n=\sum_{j=1}^l R_{nj}$. Then

$$S_n\left(m_n^{1/2}\theta\right)=\Gamma_n-m_n^{1/2}\sum_{j=1}^{l}\sum_{i=1}^{n}\widetilde{\xi}_i^\top\theta_j\psi_{\tau_j}\left(\varepsilon_{i\tau_j}\right)+R_n.$$

S1 Lemmas

Lemma 1. *Let $Z_1,\ldots,Z_n$ be arbitrary scalar random variables such that $\max_{1\le i\le n}\mathbb{E}\left(|Z_i|^{L_0}\right)<\infty$ for some $L_0\ge 1$. Then, we have*

$$\mathbb{E}\left(\max_{1\le i\le n}|Z_i|\right)\le n^{1/L_0}\max_{1\le i\le n}\left\{\mathbb{E}\left(|Z_i|^{L_0}\right)\right\}^{1/L_0}.$$

Proof of Lemma 1. This inequality follows from the observation that

$$\begin{aligned}\mathbb{E}\left(\max_{1\le i\le n}|Z_i|\right)&\le\left\{\mathbb{E}\left(\max_{1\le i\le n}|Z_i|^{L_0}\right)\right\}^{1/L_0}\\&\le\left\{\sum_{i=1}^{n}\mathbb{E}\left(|Z_i|^{L_0}\right)\right\}^{1/L_0}\le n^{1/L_0}\max_{1\le i\le n}\left\{\mathbb{E}\left(|Z_i|^{L_0}\right)\right\}^{1/L_0}.\end{aligned}$$

Lemma 2. *Under conditions (C1)–(C3) and (C5), we have*

$$\max_{1\le i\le n}\left\|\widetilde{\xi}_i\right\|_2=o_p\left\{m_n^{-1/2}(\log n)^{-1}\right\},$$

where $\left\|\widetilde{\xi}_i\right\|_2=\left(\widetilde{\xi}_i^\top\widetilde{\xi}_i\right)^{1/2}$.

Proof of Lemma 2. It is easy to show

$$\begin{aligned}\left\|\widetilde{\xi}_i\right\|_2^2 \leq & \frac{2}{n}\sum_{j=1}^{m_n}\lambda_j^{-1}\left\{\left(\widehat{\xi}_{ij}-\xi_{ij}\right)^2+\xi_{ij}^2\right\}+\frac{1}{n}\\ \leq & \frac{2}{n}||X_i||^2\sum_{j=1}^{m_n}\lambda_j^{-1}\left\|\widehat{\phi}_j-\phi_j\right\|^2+\frac{2}{n}\sum_{j=1}^{m_n}\lambda_j^{-1}\xi_{ij}^2+\frac{1}{n}.\end{aligned}$$

By (5.21) and (5.22) in Hall and Horowitz (2007), we have $\left\|\widehat{\phi}_j-\phi_j\right\|^2 = O_p(j^2/n)$, uniformly in $j \in \{1,\ldots,m_n\}$. Combining $\mathbb{E}\left(||X||^4\right) < \infty$ and Lemma 1, we have $\max_{1\leq i\leq n}||X_i||^2 = O_p\left(n^{1/2}\right)$. Hence, $\max_{1\leq i\leq n} 2||X_i||^2\sum_{j=1}^{m_n}\lambda_j^{-1}\left\|\widehat{\phi}_j-\phi_j\right\|^2 \leq O_p\left(n^{-1/2}\right)\sum_{j=1}^{m_n}j^2/\lambda_j \leq O_p\left(m_n^{\nu+3}n^{-1/2}\right)$. By assumption (C2) and Lemma 1, we have $\max_{1\leq i\leq n}\sum_{j=1}^{m_n}2\lambda_j^{-1}\xi_{ij}^2 = O_p\left(m_n n^{1/2}\right)$. Since $\nu > 1$ and $\zeta > \nu/2+1$, there exists a small constant $c > 0$ (depending on ν and ζ) such that $\max_{1\leq i\leq n} 2||X_i||^2\sum_{j=1}^{m_n}\lambda_j^{-1}\left\|\widehat{\phi}_j-\phi_j\right\|^2 = O_p\left\{n^{-c}\left(n/m_n\right)\right\}$ and $\max_{1\leq i\leq n}\sum_{j=1}^{m_n}2\lambda_j^{-1}\xi_{ij}^2 = O_p\left\{n^{-c}\left(n/m_n\right)\right\}$, and complete the proof of Lemma 2.

Lemma 3. *Under assumptions (C1)–(C5), we have*

$$\max_{1\leq i\leq n}|W_i| = o_p(1),$$

where W_i are defined in (2).

Proof of Lemma 3. Following assumptions (C1), (C4) and $\left\|\widehat{\phi}_j-\phi_j\right\|^2 = O_p\left(j^2/n\right)$, uniformly in $j \in \{1,\ldots,m_n\}$, we have

$$\max_{1\leq i\leq n}\left|\sum_{j=1}^{m_n}\left(\xi_{ij}-\widehat{\xi}_{ij}\right)\beta_j\right|$$

$$
\begin{aligned}
&\le \max_{1\le i\le n}\|X_i\|\sum_{j=1}^{m_n}\left\|\widehat{\phi}_j-\phi_j\right\||\beta_j| = O_p\left(n^{-1/4}\sum_{j=1}^{m_n} j^{1-\zeta}\right)\\
&=\begin{cases} O_p\left(n^{-1/4}m_n^{2-\zeta}\right) & \text{if } \ 3/2<\zeta<2,\\ O_p\left(n^{-1/4}\log m_n\right) & \text{if } \ \zeta=2,\\ O_p\left(n^{-1/4}\right) & \text{if } \ \zeta>2.\end{cases}
\end{aligned}
$$

Therefore, by condition (C5), we obtain $\max_{1\le i\le n}\left|\sum_{j=1}^{m_n}\left(\xi_{ij}-\widehat{\xi}_{ij}\right)\beta_j\right| = o_p(1)$. Following conditions (C2)–(C5) and Lemma 1, we have

$$
\begin{aligned}
\mathbb{E}\left(\max_{1\le i\le n}|\zeta_i|\right) &\le \mathbb{E}\left(\sum_{j=m_n+1}^{\infty}|\beta_j|\max_{1\le i\le n}|\xi_{ij}|\right)\le \sum_{j=m_n+1}^{\infty} Cn^{1/4}|\beta_j|\lambda_j^{1/2}\\
&\le Cn^{1/4}\sum_{j=m_n+1}^{\infty} j^{-(\zeta+v/2)} = Cn^{1/4}\sum_{j=m_n+1}^{\infty}\int_{j-1}^{j} j^{-(\zeta+v/2)}dx\\
&\le Cn^{1/4}\sum_{j=m_n+1}^{\infty}\int_{j-1}^{j} x^{-(\zeta+v/2)}dx = Cn^{1/4}\int_{m_n}^{\infty} x^{-(\zeta+v/2)}dx\\
&= Cn^{1/4}\frac{1}{\zeta+v/2-1}m_n^{-(\zeta+v/2)+1} = \frac{C}{\zeta+v/2-1}n^{1/4}n^{\frac{-(\zeta+v/2)+1}{v+2\zeta}}\\
&= \frac{C}{\zeta+v/2-1}n^{\frac{1}{v+2\zeta}-\frac{1}{4}} = o(1),
\end{aligned}
$$

where $v>1$ and $\zeta>v/2+1$. Thus, we have

$$
\max_{1\le i\le n}|W_i| \le \max_{1\le i\le n}\left|\sum_{j=1}^{m_n}\left(\xi_{ij}-\widehat{\xi}_{ij}\right)\beta_j\right| + \max_{1\le i\le n}|\zeta_i| = o_p(1).
$$

Lemma 4. *For $k\in\{1,\ldots,m_n\}$, define the following expressions:*

$$
\vartheta^{(1)} = \sum_{i=1}^{n}\sum_{s=m_n+1}^{\infty}\xi_{is}\beta_s\Lambda^{-1/2}\xi_i,\ \ \vartheta_k^{(2)} = \sum_{i=1}^{n}\left(\widehat{\xi}_{ik}-\xi_{ik}\right)/\sqrt{\lambda_k}\sum_{s=m_n+1}^{\infty}\xi_{is}\beta_s,
$$

$$
\vartheta^{(3)} = \sum_{i=1}^{n}\sum_{s=1}^{m_n}\left(\widehat{\xi}_{is}-\xi_{is}\right)\beta_s\Lambda^{-1/2}\xi_i,\ \vartheta_k^{(4)} = \sum_{i=1}^{n}\sum_{s=1}^{m_n}\beta_s\left(\widehat{\xi}_{is}-\xi_{is}\right)\left(\widehat{\xi}_{ik}-\xi_{ik}\right)/\sqrt{\lambda_k}.
$$

Under conditions (C1)–(C5), we have

$$\left\|\vartheta^{(1)}\right\|_2 = o_p\left\{(m_n n)^{1/2}\right\},\ \ \vartheta_k^{(2)} = o_p\left(n^{1/2}\right),$$

$$\left\|\vartheta^{(3)}\right\|_2 = O_p\left\{(m_n n)^{1/2}\right\},\ \ \vartheta_k^{(4)} = O_p\left(k^{\nu/2+1}\right),$$

where o_p and O_p are uniform for $k \in \{1, \ldots, m_n\}$.

Proof of Lemma 4. Please refer to the proof of Lemma 2 in Kong et al. (2016).

S2 Proof of Theorem 1

By simple calculation, we have

$$\left|S_{nij} + m_n^{1/2}\widetilde{\xi}_i^\top \theta_j \psi_{\tau_j}\left(\varepsilon_{i\tau_j}\right)\right| \le 2\left|m_n^{1/2}\widetilde{\xi}_i^\top \theta_j\right| \mathbf{1}\left\{\left|\varepsilon_{i\tau_j}\right| \le \left|m_n^{1/2}\widetilde{\xi}_i^\top \theta_j - W_i\right|\right\}.$$

Therefore,

$$\begin{aligned}
&\mathbb{E}\left(R_{nj}\right)^2 \\
&= \mathbb{E}\left[\mathbb{E}\left\{\left(\sum_{i=1}^n R_{nij}^2 + 2\sum_{i=1}^n \sum_{k:i<k} R_{nij}R_{nkj}\right)\middle|\Omega\right\}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left\{\left(\sum_{i=1}^n R_{nij}^2\right)\middle|\Omega\right\}\right] \\
&= \sum_{i=1}^n \mathbb{E}\left(\mathbb{E}\left[\left\{S_{nij} - \Gamma_{nij} + m_n^{1/2}\widetilde{\xi}_i^\top \theta_j \psi_{\tau_j}\left(\varepsilon_{i\tau_j}\right)\right\}^2 \middle|\Omega\right]\right) \\
&\le \sum_{i=1}^n \mathbb{E}\left(\mathbb{E}\left[\left\{S_{nij} + m_n^{1/2}\widetilde{\xi}_i^\top \theta_j \psi_{\tau_j}\left(\varepsilon_{i\tau_j}\right)\right\}^2 \middle|\Omega\right]\right)
\end{aligned}$$

$$\leq 2m_n \mathbb{E}\left[\sum_{i=1}^{n}\left(\widetilde{\xi}_i^\top \theta_j\right)^2 \mathbb{E}\left\{\mathbf{1}\left(\left|\varepsilon_{i\tau_j}\right| \leq \max_{1\leq i\leq n}\left|m_n^{1/2}\widetilde{\xi}_i^\top \theta_j - W_i\right|\right)\Big|\Omega\right\}\right]$$

$$\leq 2m_n \left[\mathbb{E}\left\{\sum_{i=1}^{n}\left(\widetilde{\xi}_i^\top \theta_j\right)^2\right\}^2\right]^{1/2} \mathbb{E}\left\{\mathbf{1}\left(\left|\varepsilon_{i\tau_j}\right| \leq \max_{1\leq i\leq n}\left|m_n^{1/2}\widetilde{\xi}_i^\top \theta_j - W_i\right|\right)\right\}.$$

Following

$$\frac{1}{n}\sum_{i=1}^{n}\widehat{\xi}_{ik}\widehat{\xi}_{ij} = \iint_{\mathcal{I}^2}\widehat{K}(s,t)\widehat{\phi}_k(s)\widehat{\phi}_j(t)dsdt = \begin{cases}\widehat{\lambda}_k & \text{if } k=j,\\ 0 & \text{if } k\neq j\end{cases}$$

and (5.2) in Hall and Horowitz (2007), we get

$$\begin{aligned}
&\sum_{i=1}^{n}\left(\widetilde{\xi}_i^\top \theta_j\right)^2\\
&= \sum_{i=1}^{n}\left(\theta_j^\top\left(n^{-1/2}, 0_{m_n}^\top\right)^\top + \theta_j^\top\left(0, \left\{n^{-1/2}\Lambda_1^{-1/2}\left(\widehat{\xi}_{i1},\ldots,\widehat{\xi}_{im_n}\right)^\top\right\}^\top\right)^\top\right)^2\\
&= \theta_{j1}^2 + \theta_j^\top \mathrm{diag}\left(0, \widehat{\lambda}_1/\lambda_1, \ldots, \widehat{\lambda}_{m_n}/\lambda_{m_n}\right)\theta_j\\
&= \|\theta_j\|_2^2 + \theta_j^\top \mathrm{diag}\left(0, \left(\widehat{\lambda}_1-\lambda_1\right)/\lambda_1, \ldots, \left(\widehat{\lambda}_{m_n}-\lambda_{m_n}\right)/\lambda_{m_n}\right)\theta_j \qquad \text{(S2.1)}\\
&\leq \|\theta_j\|_2^2 + \|\theta_j\|_2^2\widehat{\Delta}/\lambda_{m_n},
\end{aligned}$$

where θ_{j1} is the first component of θ_j, 0_{m_n} represents an m_n-dimensional vector whose components are all zero, and $\widehat{\Delta} = \left\{\iint_{\mathcal{I}^2}\left(\widehat{K}-K\right)^2\right\}^{1/2}$. By (5.9) in Hall and Horowitz (2007) and conditions (C3)–(C5), we get $\mathbb{E}\left(\widehat{\Delta}^2/\lambda_{m_n}^2\right) = o(1)$. Therefore,

$$\mathbb{E}\left\{\sum_{i=1}^{n}\left(\widetilde{\xi}_i^\top \theta_j\right)^2\right\}^2 \leq 2\|\theta_j\|_2^4\left\{1+\mathbb{E}\left(\widehat{\Delta}^2/\lambda_{m_n}^2\right)\right\} = C\|\theta_j\|_2^4. \qquad \text{(S2.2)}$$

Suppose that L_1 is a positive constant. For $\left\|\theta_j\right\|_2 \leq \|\theta\|_2 \leq L_1$, following Lemmas 2 and 3, and $\varepsilon_{i\tau_j} = O_p(1)$, we have

$$\begin{aligned}
&\mathbb{E}\left\{\mathbf{1}\left(\left|\varepsilon_{i\tau_j}\right| \leq \max_{1\leq i\leq n}\left|m_n^{1/2}\widetilde{\xi}_i^\top\theta_j - W_i\right|\right)\right\} \\
&= \mathbb{P}\left(\left|\varepsilon_{i\tau_j}\right| \leq \max_{1\leq i\leq n}\left|m_n^{1/2}\widetilde{\xi}_i^\top\theta_j - W_i\right|\right) = o(1). \qquad \text{(S2.3)}
\end{aligned}$$

Combining (S2.2) and (S2.3), we obtain

$$\mathbb{E}\left(R_{nj}\right)^2 = o\left(m_n\left\|\theta_j\right\|_2^2\right). \qquad \text{(S2.4)}$$

It is easy to show that

$$\begin{aligned}
\mathbb{E}\left\{m_n^{1/2}\sum_{i=1}^n\widetilde{\xi}_i^\top\theta_j\psi_{\tau_j}\left(\varepsilon_{i\tau_j}\right)\right\}^2 &= \mathbb{E}\left(\mathbb{E}\left[\left\{m_n^{1/2}\sum_{i=1}^n\widetilde{\xi}_i^\top\theta_j\psi_{\tau_j}\left(\varepsilon_{i\tau_j}\right)\right\}^2\middle|\Omega\right]\right) \\
&= m_n\sum_{i=1}^n\mathbb{E}\left(\mathbb{E}\left[\left\{\widetilde{\xi}_i^\top\theta_j\psi_{\tau_j}\left(\varepsilon_{i\tau_j}\right)\right\}^2\middle|\Omega\right]\right) \\
&\leq m_n\mathbb{E}\left\{\sum_{i=1}^n\left(\widetilde{\xi}_i^\top\theta_j\right)^2\right\} = O\left(m_n\|\theta_j\|_2^2\right).
\end{aligned} \qquad \text{(S2.5)}$$

Since $\mathbb{E}\left(\widehat{\Delta}^2/\lambda_{m_n}^2\right) = o(1)$, we have

$$\begin{aligned}
&\left|\theta_j^\top\operatorname{diag}\left(0, \left(\widehat{\lambda}_1 - \lambda_1\right)/\lambda_1, \ldots, \left(\widehat{\lambda}_{m_n} - \lambda_{m_n}\right)/\lambda_{m_n}\right)\theta_j\right| \\
&\leq \|\theta_j\|_2^2\widehat{\Delta}/\lambda_{m_n} = o_p\left(\|\theta_j\|_2^2\right).
\end{aligned}$$

Hence, by (S2.1), we get

$$\sum_{i=1}^n\left(m_n^{1/2}\widetilde{\xi}_i^\top\theta_j\right)^2$$

$$= m_n\|\theta_j\|_2^2 + m_n\theta_j^\top \mathrm{diag}\left(0, \left(\widehat{\lambda}_1 - \lambda_1\right)/\lambda_1, \ldots, \left(\widehat{\lambda}_{m_n} - \lambda_{m_n}\right)/\lambda_{m_n}\right)\theta_j$$

$$= m_n\|\theta_j\|_2^2\{1 + o_p(1)\}. \tag{S2.6}$$

It is easy to prove that

$$\begin{aligned}
&\sum_{i=1}^{n} m_n^{1/2} W_i \widetilde{\xi}_i^\top \theta_j \\
&= (m_n/n)^{\frac{1}{2}} \left(\theta_j^\top \vartheta^{(3)} + \sum_{s=1}^{m_n} \vartheta_s^{(4)} u_s + \theta_j^\top \vartheta^{(1)} + \sum_{s=1}^{m_n} \vartheta_s^{(2)} u_s\right).
\end{aligned}$$

Following Lemma 4 and condition (C5), we conclude that

$$(m_n/n)^{\frac{1}{2}}\theta_j^\top \vartheta^{(3)} = O_p(m_n)\left\|\theta_j\right\|_2, \quad (m_n/n)^{\frac{1}{2}}\theta_j^\top \vartheta^{(1)} = o_p(m_n)\left\|\theta_j\right\|_2,$$

$$(m_n/n)^{\frac{1}{2}} \sum_{s=1}^{m_n} \vartheta_s^{(2)} u_s = o_p(m_n)\left\|\theta_j\right\|_2,$$

$$\begin{aligned}
(m_n/n)^{\frac{1}{2}} \sum_{s=1}^{m_n} \vartheta_s^{(4)} u_s &= O_p\left\{(m_n/n)^{\frac{1}{2}} \left(\sum_{s=1}^{m_n} s^{\nu+2}\right)^{1/2}\right\}\left\|\theta_j\right\|_2 \\
&= O_p\left\{(m_n/n)^{\frac{1}{2}}\, m_n^{(\nu+3)/2}\right\}\left\|\theta_j\right\|_2 = o_p\left(m_n\right)\left\|\theta_j\right\|_2.
\end{aligned}$$

Thus,

$$\sum_{i=1}^{n} m_n^{1/2} W_i \widetilde{\xi}_i^\top \theta_j = O_p\left(m_n\right)\|\theta_j\|_2. \tag{S2.7}$$

By the proof of Lemma 2 in Yao et al. (2017) and condition (C6),

$$\begin{aligned}
\Gamma_{nj} &= \sum_{i=1}^{n} \mathbb{E}\left\{\rho_{\tau_j}\left(W_i + \varepsilon_{i\tau_j} - \widetilde{\xi}_i^\top m_n^{1/2}\theta_j\right)\Big|\Omega\right\} - \mathbb{E}\left\{\rho_{\tau_j}\left(W_i + \varepsilon_{i\tau_j}\right)\big|\Omega\right\} \\
&= \frac{1}{2}\sum_{i=1}^{n} f_{\tau_j}(0|X_i)\left\{\left(m_n^{1/2}\widetilde{\xi}_i^\top \theta_j\right)^2 - 2m_n^{1/2} W_i \widetilde{\xi}_i^\top \theta_j\right\}\{1 + o_p(1)\}
\end{aligned}$$

Following condition (C6), (S2.6) and (S2.7), we get

$$\Gamma_{nj} = m_n \|\theta_j\|_2^2 \{1 + o_p(1)\} + O_p\left(m_n\right) \|\theta_j\|_2. \tag{S2.8}$$

Combining (S2.4), (S2.5) and (S2.8), for sufficiently large L_1, we have

$$\inf_{\|\theta\|_2 = L_1} S_n\left(m_n^{1/2}\theta\right) \geqslant c_0 L_1^2 m_n \{1 + o_p(1)\},$$

where c_0 is a positive constant. This implies that

$$\mathbb{P}\left(\inf_{\|\theta\|_2 \geqslant L_1}\left[\sum_{j=1}^{l}\sum_{i=1}^{n}\left\{\rho_{\tau_j}\left(W_i + \varepsilon_{i\tau_j} - m_n^{1/2}\widetilde{\xi}_i^\top \theta_j\right) - \rho_{\tau_j}\left(W_i + \varepsilon_{i\tau_j}\right)\right\}\right] > 0\right) \to 1$$

as $n \to \infty$. Hence, $\mathbb{P}\left(\left\|\widehat{\theta}\right\|_2 \leq L_1 m_n^{1/2}\right) \to 1$ as $n \to \infty$, where $\widehat{\theta}$ is the minimizer of (3). Hence, $\left\|\widehat{\theta}\right\|_2 = O_p\left(m_n^{1/2}\right)$. Therefore, we have

$$\left\|n^{1/2}\Lambda_1^{1/2}\left(\widehat{\beta} - \beta\right)\right\|_2 = O_p\left(m_n^{1/2}\right), \tag{S2.9}$$

where $\widehat{\beta} = \left(\widehat{\beta}_1, \ldots, \widehat{\beta}_{m_n}\right)^\top$. By some straightforward calculations, we get

$$\begin{aligned}
&\left\|\widehat{\beta}_0 - \beta_0\right\|^2 \\
&= \left\|\sum_{j=1}^{m_n}\widehat{\beta}_j\widehat{\phi}_j - \sum_{j=1}^{\infty}\beta_j\phi_j\right\|^2 \\
&\leq 2\left\|\sum_{j=1}^{m_n}\widehat{\beta}_j\widehat{\phi}_j - \sum_{j=1}^{m_n}\beta_j\phi_j\right\|^2 + 2\left\|\sum_{j=m_n+1}^{\infty}\beta_j\phi_j\right\|^2 \\
&\leq 4\left\|\sum_{j=1}^{m_n}\left(\widehat{\beta}_j - \beta_j\right)\widehat{\phi}_j\right\|^2 + 4\left\|\sum_{j=1}^{m_n}\beta_j\left(\widehat{\phi}_j - \phi_j\right)\right\|^2 + 2\sum_{j=m_n+1}^{\infty}\beta_j^2 \\
&\leq 4\sum_{j=1}^{m_n}\left(\widehat{\beta}_j - \beta_j\right)^2 + 4m_n\sum_{j=1}^{m_n}\beta_j^2\left\|\widehat{\phi}_j - \phi_j\right\|^2 + 2\sum_{j=m_n+1}^{\infty}\beta_j^2
\end{aligned}$$

$$\leq 4(n\lambda_{m_n})^{-1}\left\|n^{1/2}\Lambda_1^{1/2}\left(\widehat{\beta}-\beta\right)\right\|_2^2 + 4m_n\sum_{j=1}^{m_n}\beta_j^2\left\|\widehat{\phi}_j-\phi_j\right\|^2 + 2\sum_{j=m_n+1}^{\infty}\beta_j^2.$$

By (S2.9), conditions (C3) and (C5), we obtain

$$(n\lambda_{m_n})^{-1}\left\|n^{1/2}\Lambda_1^{1/2}\left(\widehat{\beta}-\beta\right)\right\|_2^2 \leq O_p\left(n^{-1}m_n^{v+1}\right) = O_p\left\{n^{-(2\zeta-1)/(\nu+2\zeta)}\right\}. \tag{S2.10}$$

Since $\left\|\widehat{\phi}_j-\phi_j\right\|^2 = O_p\left(j^2/n\right)$, uniformly in $j\in\{1,\ldots,m_n\}$, we have

$$m_n\sum_{j=1}^{m_n}\beta_j^2\left\|\widehat{\phi}_j-\phi_j\right\|^2 = m_n\sum_{j=1}^{m_n}j^{-2\zeta}O_p\left(n^{-1}j^2\right) = \frac{m_n}{n}\,O_p\left(\sum_{j=1}^{m_n}j^{2-2\zeta}\right)$$

$$= \begin{cases} O_p\left(\frac{m_n}{n}\right) & \text{if} \quad 2-2\zeta<-1, \\ O_p\left(\frac{m_n\log m_n}{n}\right) & \text{if} \quad 2-2\zeta=-1, \\ O_p\left(\frac{m_n^{4-2\zeta}}{n}\right) & \text{if} \quad 2-2\zeta>-1. \end{cases}$$

Since $\zeta>\nu/2+1$ and $\nu>1$, we conclude that

$$m_n\sum_{j=1}^{m_n}\beta_j^2\left\|\widehat{\phi}_j-\phi_j\right\|^2 = O_p(m_n/n) = o_p\left\{n^{-(2\zeta-1)/(\nu+2\zeta)}\right\}. \tag{S2.11}$$

By conditions (C4) and (C5), we get

$$\sum_{j=m_n+1}^{\infty}\beta_j^2 \leq C\sum_{j=m_n+1}^{\infty}j^{-2\zeta} = O\left\{m_n^{-(2\zeta-1)}\right\} = O\left\{n^{-(2\zeta-1)/(\nu+2\zeta)}\right\}. \tag{S2.12}$$

Combining (S2.10)–(S2.12), we complete the proof of Theorem 1.

S3 Proof of Theorem 2

Recall that $v_i \in \{\tau : Q_Y(\tau|X_i) = Y_i\}$, $Q_Y(v_i|X_i) = Y_i$, $v_i \sim$ Uniform(0,1), and $\widehat{e}_i = Y_i - \langle X_i, \widehat{\beta}_0\rangle = Q_Y(v_i|X_i) - \langle X_i, \widehat{\beta}_0\rangle$, for $i = 1, \dots, n$. Hence, the ordering of $\{\widehat{e}_1, \dots, \widehat{e}_n\}$ is not necessarily the same as the ordering of $\{v_1, \dots, v_n\}$. The main task of this proof is to illustrate that the $k_n + 1$ largest $\widehat{e}_i$'s correspond to the $k_n + 1$ largest v_i's. To this end, we first show that with probability tending to one, $\widehat{e}_{(n-j)}$ for $j = 0, 1, \dots, k_n$ can be decomposed as follows:

$$\widehat{e}_{(n-j)} = \alpha(v_{i(j)}) + \langle X_{i(j)}, \beta_0 - \widehat{\beta}_0\rangle, \tag{S3.13}$$

where $i(j)$ is the index function such that $\widehat{e}_{(n-j)} = \widehat{e}_{i(j)}$ for $j = 0, 1, \dots, n-1$. It can be seen from (2.2), in order to get (S3.13), we only need to prove that with probability tending to one, $v_{i(j)} > \tau_0$ jointly for all $j = 0, 1, \dots, k_n$. Let $v_{(1)} \le \cdots \le v_{(n)}$ be the order statistics of $\{v_1, \dots, v_n\}$. Define $\widetilde{i}(j)$ by $v_{\widetilde{i}(j)} = v_{(n-j)}$ for $j = 0, 1, \dots, n-1$. Then

$$\begin{aligned}
&\mathbb{P}\left(\cup_{j=0}^{k_n}\left\{v_{i(j)} < \tau_0\right\}\right)\\
&= \mathbb{P}\left(\cup_{j=0}^{k_n}\left\{Y_{i(j)} < Q_Y\left(\tau_0|X_{i(j)}\right)\right\}\right)\\
&= \mathbb{P}\left(\min_{0\le j\le k_n}\left(Y_{i(j)} - \alpha(\tau_0) - \langle X_{i(j)}, \beta_0\rangle\right) < 0\right)\\
&= \mathbb{P}\left(\min_{0\le j\le k_n}\left(Y_{i(j)} - \langle X_{i(j)}, \widehat{\beta}_0\rangle - \langle X_{i(j)}, \beta_0 - \widehat{\beta}_0\rangle - \alpha(\tau_0)\right) < 0\right)
\end{aligned}$$

$$
\begin{aligned}
&\le \mathbb{P}\left(\min_{0\le j\le k_n} \widehat{e}_{(n-j)} - \max_{1\le i\le n}\left|\langle X_i, \beta_0 - \widehat{\beta}_0\rangle\right| - |\alpha(\tau_0)| < 0\right)\\
&= \mathbb{P}\left(\widehat{e}_{(n-k_n)} < \max_{1\le i\le n}\left|\langle X_i, \beta_0 - \widehat{\beta}_0\rangle\right| + |\alpha(\tau_0)|\right)\\
&= 1 - \mathbb{P}\left(\widehat{e}_{(n-k_n)} \ge \max_{1\le i\le n}\left|\langle X_i, \beta_0 - \widehat{\beta}_0\rangle\right| + |\alpha(\tau_0)|\right)\\
&\le 1 - \mathbb{P}\left(\bigcap_{j=0}^{k_n}\left\{\widehat{e}_{\tilde{i}(j)} \ge \max_{1\le i\le n}\left|\langle X_i, \beta_0 - \widehat{\beta}_0\rangle\right| + |\alpha(\tau_0)|\right\}\right)\\
&= 1 - \mathbb{P}\left(\bigcap_{j=0}^{k_n}\left\{\alpha\left(v_{(n-j)}\right) + \langle X_{\tilde{i}(j)}, \beta_0 - \widehat{\beta}_0\rangle \ge \max_{1\le i\le n}\left|\langle X_i, \beta_0 - \widehat{\beta}_0\rangle\right| + |\alpha(\tau_0)|\right\}\right)\\
&\le 1 - \mathbb{P}\left(\alpha\left(v_{(n-k_n)}\right) \ge 2\max_{1\le i\le n}\left|\langle X_i, \beta_0 - \widehat{\beta}_0\rangle\right| + |\alpha(\tau_0)|\right),
\end{aligned}
$$

where following (2.2) and $\mathbb{P}\left(v_{(n-k_n)} > \tau_0\right) \to 1$ as $n \to \infty$, we get the last equality holds with probability tending to one. By Theorem 1, Lemma 1 and conditions (C1) and (C4), we have

$$
\max_{1\le i\le n}\left|\langle X_i, \beta_0 - \widehat{\beta}_0\rangle\right| = O_p\left(n^{1/4-\frac{2\zeta-1}{2(\nu+2\zeta)}}\right) = O_p\left(n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\right) = o_p(1). \tag{S3.14}
$$

It is easy to show $\lim_{n\to\infty}\mathbb{P}\left(\alpha\left(v_{(n-k_n)}\right) > C\right) = 1$ for arbitrary large positive constant C and $\alpha(\tau_0) = O(1)$. Thus,

$$
\lim_{n\to\infty}\mathbb{P}\left(\cup_{j=0}^{k_n}\left\{v_{i(j)} < \tau_0\right\}\right) = 0,
$$

and (S3.13) is proved. Next, we show that

$$
\lim_{n\to\infty}\mathbb{P}\left(\cap_{j=0}^{k_n}\left\{\widehat{e}_{(n-j)} = \alpha\left(v_{(n-j)}\right) + \langle X_{i(j)}, \beta_0 - \widehat{\beta}_0\rangle\right\}\right) = 1. \tag{S3.15}
$$

To this aim, we know from (S3.13) that it is sufficient to prove

$$\max_{1\le i\neq j\le n}\left|\langle X_i - X_j, \beta_0 - \widehat{\beta}_0\rangle\right| = o_p\left(\min_{1\le j\le k_n}\left(\alpha\left(v_{(n-j+1)}\right) - \alpha\left(v_{(n-j)}\right)\right)\right). \tag{S3.16}$$

We now prove that (S3.16) holds under the conditions outlined in the paper. Under condition $C_n := \sqrt{k_n^{\gamma+1}(k_n+1)n^{-\gamma-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}} \to 0$ as $n\to\infty$, for $\forall$ $\varepsilon > 0$, we have

$$\begin{aligned}
&\mathbb{P}\left(\frac{n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}}{(n/k_n)^\gamma \frac{2E_1}{k_n(1+k_n)}} > \varepsilon\right)\\
&= \mathbb{P}\left(\frac{C_n^2}{2E_1} > \varepsilon, E_1 > C_n\right) + \mathbb{P}\left(\frac{C_n^2}{2E_1} > \varepsilon, E_1 \le C_n\right)\\
&\le \mathbb{P}\left(\frac{C_n}{2} > \varepsilon, E_1 > C_n\right) + \mathbb{P}\left(E_1 \le C_n\right) \to 0
\end{aligned}$$

as $n\to\infty$, where E_1 is a standard exponential random variable. Hence, we have

$$\max_{1\le i\neq j\le n}\left|\langle X_i - X_j, \beta_0 - \widehat{\beta}_0\rangle\right| = O_p\left(n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\right) = o_p\left((n/k_n)^\gamma\frac{E_1}{k_n(1+k_n)}\right). \tag{S3.17}$$

In what follows, we prove

$$\begin{aligned}
&\min_{1\le j\le k_n}\left(\alpha\left(v_{(n-j+1)}\right) - \alpha\left(v_{(n-j)}\right)\right)\\
&\ge C(n/k_n)^\gamma\left(\sqrt{\frac{\gamma}{2}\left(1 - \frac{1}{\min_{1\le i\le k_n} z_i}\right) + \frac{1}{4}} - \frac{1}{2}\right)\\
&\overset{d}{=} (n/k_n)^\gamma \frac{2E_1}{k_n(1+k_n)}(C + o_p(1))
\end{aligned} \tag{S3.18}$$

in probability, where $z_j = \frac{1-v_{(n-j)}}{1-v_{(n-j+1)}} > 1$, for $j = 1, \ldots, k_n$, $L_n \stackrel{d}{=} S_n$ means that L_n and S_n have the same distribution, and C denotes a generic fixed positive constant, the value of which varies in different places. Following (S3.17) and (S3.18), we conclude that (S3.16) holds.

It is easy to show that

$$\begin{aligned}\min_{1\le j\le k_n}\left[\alpha\left(v_{(n-j+1)}\right)-\alpha\left(v_{(n-j)}\right)\right] &= \min_{1\le j\le k_n}\left[\frac{\alpha\left(v_{(n-j+1)}\right)}{\alpha\left(v_{(n-j)}\right)}-1\right]\alpha\left(v_{(n-j)}\right)\\ &\ge \alpha\left(v_{(n-k_n)}\right)\min_{1\le j\le k_n}\left[\frac{\alpha\left(v_{(n-j+1)}\right)}{\alpha\left(v_{(n-j)}\right)}-1\right].\end{aligned}$$

Let $U(t) := H_Y(t|x=0) = \inf\{y : F_Y(y|x=0) \ge 1-1/t\}$ for $t > 1$ and $t_j = \frac{1}{1-v_{(n-j)}}$, for $j = 1, \ldots, k_n$. Then

$$\frac{\alpha\left(v_{(n-j+1)}\right)}{\alpha\left(v_{(n-j)}\right)} = \frac{\alpha\left(1-\frac{1}{t_j z_j}\right)}{\alpha\left(1-\frac{1}{t_j}\right)} = \frac{U(t_j z_j)}{U(t_j)}.$$

It is easy to show that

$$\log\left(U(t_j z_j)\right) - \log\left(U(t_j)\right) \ge \frac{U(t_j z_j)-U(t_j)}{U(t_j z_j)}.$$

Therefore, we have

$$\begin{aligned}\frac{U(t_j z_j)}{U(t_j)} &\ge \exp\left(\frac{U(t_j z_j)-U(t_j)}{U(t_j z_j)}\right)\\ &= \exp\left(\frac{(t_j z_j - t_j)U'(\pi_j)}{U(t_j z_j)}\right)\\ &= \exp\left(\frac{U'\left(\pi_j\right)\pi_j}{U(\pi_j)}\cdot\frac{U\left(\pi_j\right)}{U\left(t_j z_j\right)}\cdot\frac{\left(z_j-1\right)t_j}{\pi_j}\right),\end{aligned}$$

where $U'(\cdot)$ is the first derivative of $U(\cdot)$, $t_j < \pi_j < t_j z_j$, for $j = 1, \ldots, k_n$, and $t_1 > t_2 > \cdots > t_{k_n} \to \infty$ as $n \to \infty$. By condition (C7) ($\lim_{t\to\infty} \frac{tU'(t)}{U(t)} = \gamma$), we know that $\frac{U'(\pi_j)\pi_j}{U(\pi_j)} \geq \frac{\gamma}{2}$ when n is sufficiently large, for $j = 1, \ldots, k_n$. Hence, we have

$$\begin{aligned}\frac{U(t_j z_j)}{U(t_j)} &\geq \exp\left(\frac{\gamma}{2} \cdot \frac{U\left(t_j\right)}{U(t_j z_j)} \cdot \frac{\left(z_j - 1\right) t_j}{t_j z_j}\right) \\ &= \exp\left(\frac{\gamma}{2} \cdot \frac{U\left(t_j\right)}{U(t_j z_j)} \cdot \frac{z_j - 1}{z_j}\right).\end{aligned}$$

Let $\frac{U(t_j z_j)}{U(t_j)} = \Delta_j$, for $j = 1, \ldots, k_n$. Then

$$\Delta_j \log \Delta_j \geq \frac{\gamma}{2}\left(1 - \frac{1}{z_j}\right) \geq \frac{\gamma}{2}\left(1 - \frac{1}{\min_{1\leq i\leq k_n} z_i}\right),$$

uniformly in $j \in \{1, \ldots, k_n\}$. For $x \geq 1$, we have $\log x \leq x - 1$. Therefore,

$$\Delta_j(\Delta_j - 1) \geq \Delta_j \log \Delta_j \geq \frac{\gamma}{2}\left(1 - \frac{1}{\min_{1\leq i\leq k_n} z_i}\right),$$

and we have

$$\Delta_j \geq \sqrt{\frac{\gamma}{2}\left(1 - \frac{1}{\min_{1\leq i\leq k_n} z_i}\right) + \frac{1}{4}} + \frac{1}{2}.$$

Observe that $\log z_j = \log \frac{1}{1-v_{(n-j+1)}} - \log \frac{1}{1-v_{(n-j)}} \overset{d}{=} E_{(n-j+1)} - E_{(n-j)}$, where $E_{(1)} \leq \cdots \leq E_{(n)}$ are the order statistics of $E_1, \ldots, E_n$, and $E_1, \ldots, E_n$ are i.i.d. standard exponential random variables. Thus, by Rényi's representation (Rényi, 1953), we have

$$z_j \overset{d}{=} \exp\left(\frac{E_j}{j}\right), \text{ for } 1 \leq j \leq k_n.$$

Thus, we get

$$\begin{aligned}
\min_{1\le j\le k_n}\left[\frac{\alpha\left(v_{(n-j+1)}\right)}{\alpha\left(v_{(n-j)}\right)}-1\right] &= \min_{1\le j\le k_n}(\Delta_j-1)\\
&\ge \sqrt{\frac{\gamma}{2}\left(1-\frac{1}{\min_{1\le i\le k_n} z_i}\right)+\frac{1}{4}}-\frac{1}{2}\\
&\stackrel{d}{=} \sqrt{\frac{\gamma}{2}\left(1-\frac{1}{\exp\left(\min_{1\le i\le k_n}\frac{E_i}{i}\right)}\right)+\frac{1}{4}}-\frac{1}{2}\\
&\stackrel{d}{=} \sqrt{\frac{\gamma}{2}\left(1-\frac{1}{\exp\left(\frac{2E_1}{k_n(1+k_n)}\right)}\right)+\frac{1}{4}}-\frac{1}{2}\\
&\stackrel{d}{=} \frac{\gamma}{2}\left(1-\exp\left(-\frac{2E_1}{k_n(1+k_n)}\right)\right)(1+o_p(1))\\
&\stackrel{d}{=} \frac{2E_1}{k_n(1+k_n)}\left(C_0+o_p(1)\right), \qquad \text{(S3.19)}
\end{aligned}$$

where $C_0=\frac{\gamma}{2}>0$. By Corollary 1.2.10 in de Haan and Ferreira (2006), we have

$$\begin{aligned}
&\frac{H_Y\left(\frac{n}{k_n}\cdot\frac{k_n}{\frac{n}{2}}|x=0\right)}{H_Y\left(\frac{n}{k_n}|x=0\right)}\cdot\left(\frac{n}{2k_n}\right)^{\gamma}\\
&=\frac{H_Y\left(2|x=0\right)}{H_Y\left(\frac{1}{1-(1-k_n/n)}|x=0\right)}\cdot\left(\frac{n}{2k_n}\right)^{\gamma}=\frac{H_Y\left(2|x=0\right)}{\alpha(1-k_n/n)}\cdot\left(\frac{n}{2k_n}\right)^{\gamma}\to 1
\end{aligned}$$

as $n\to\infty$, where $H_Y\left(2|x=0\right)=\inf\{y: F_Y(y|x=0)\ge 1-1/2\}$ is the $\frac{1}{2}$th quantile of $F_Y(\cdot|x=0)$, which is a constant. Hence, we have

$$\alpha\left(1-k_n/n\right)/\left(C\left(n/k_n\right)^{\gamma}\right)\to 1 \text{ as } n\to\infty.$$

From the proof of Theorem 1 in Wang et al. (2012), we have $v_{(n-k_n)}/(1-k_n/n)\to 1$ in probability and $\alpha\left(v_{(n-k_n)}\right)=\alpha(1-k_n/n)\left\{1+o_p(1)\right\}$. Thus, we get

$$\alpha\left(v_{(n-k_n)}\right)/\left(C(n/k_n)^{\gamma}\right)\to 1 \tag{S3.20}$$

in probability. Combining (S3.19) and (S3.20), we obtain that (S3.18) holds.

Therefore, from (S3.14) and (S3.15), we can rewrite (2.3) as

$$\begin{aligned}\widehat{\gamma} &= \frac{1}{k_n}\sum_{j=1}^{k_n}\log\frac{\alpha\left(v_{(n-j+1)}\right)+O_p\left(n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\right)}{\alpha\left(v_{(n-k_n)}\right)+O_p\left(n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\right)}\\ &= \widetilde{\gamma}+\varrho_n,\end{aligned}$$

where $\widetilde{\gamma}=\frac{1}{k_n}\sum_{j=1}^{k_n}\log\frac{\alpha\left(v_{(n-j+1)}\right)}{\alpha\left(v_{(n-k_n)}\right)}$ is the well-known Hill estimator, and

$$\varrho_n=\frac{1}{k_n}\sum_{j=1}^{k_n}\left[\log\left\{1+\frac{O_p\left(n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\right)}{\alpha\left(v_{(n-j+1)}\right)}\right\}-\log\left\{1+\frac{O_p\left(n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\right)}{\alpha\left(v_{(n-k_n)}\right)}\right\}\right].$$

Following Theorem 3.2.5 in de Haan and Ferreira (2006), under the second-order condition (3.3), we can show that there exist a sequence of Brownian motions $\{B_n(t):t\geq 0\}$ and a suitable function $A_0(\cdot)$, where $\lim_{t\to\infty}A_0(t)/A(t)=1$, such that

$$\begin{aligned}\sqrt{k_n}\left(\widetilde{\gamma}-\gamma\right) &= \frac{\sqrt{k_n}A_0(n/k_n)}{1-\delta}+\gamma\int_0^1\left\{t^{-1}B_n(t)-B_n(1)\right\}dt+o_p(1)\\ &=: \Gamma_n+o_p(1),\end{aligned}$$

where $\Gamma_n=\frac{\eta}{1-\delta}+\gamma\int_0^1\left\{t^{-1}B_n(t)-B_n(1)\right\}dt$ is normally distributed with $\mathbb{E}(\Gamma_n)=\frac{\eta}{1-\delta}$ and $\mathrm{Var}(\Gamma_n)=\gamma^2$. By Taylor expansion and $k_n^{\gamma+2}n^{-\gamma-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\to$

0 as $n \to \infty$,

$$\varrho_n = O_p\left(n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\right) / \alpha\left(v_{(n-k_n)}\right) = O_p\left(n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\left(n/k_n\right)^{-\gamma}\right) = o_p\left(k_n^{-1/2}\right).$$

Hence,

$$\sqrt{k_n}\left(\widehat{\gamma}-\gamma\right) = \sqrt{k_n}\left(\widetilde{\gamma}-\gamma\right) + o_p(1) \xrightarrow{d} N\left(\frac{\eta}{1-\delta}, \gamma^2\right),$$

and complete the proof of Theorem 2.

S4 Proof of Theorem 3

Following Theorem 2.4.2 in de Haan and Ferreira (2006) and Theorem 2.1 in Drees (1998), we have

$$\begin{aligned}
&\alpha\left(v_{(n-k_n)}\right) \\
&= H_Y\left(n/k_n|X=0\right)\left[1 + \frac{f_0\left(n/k_n\right)}{H_Y\left(n/k_n|X=0\right)}\left\{k_n^{-1/2}B_n(1) + o_p\left(k_n^{-1/2}\right)\right\}\right],
\end{aligned}$$

where the definition of $f_0(\cdot)$ can be found in de Haan and Ferreira (2006) and $0 < \lim_{t\to\infty} f_0(t)/H_Y(t|X=0) =: C_0 < \infty$. From $n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}(n/k_n)^{-\gamma}k_n^{1/2} \to 0$ and (2.4), we have

$$\begin{aligned}
&\widehat{e}_{(n-k_n)} \\
&= \alpha\left(v_{(n-k_n)}\right) + O_p\left(n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\right) \\
&= H_Y(n/k_n|X=0)\left\{1 + C_0 k_n^{-1/2}B_n(1) + o_p\left(k_n^{-1/2}\right)\right\} + O_p\left(n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\right)
\end{aligned}$$

$$= H_Y(n/k_n|X=0)\left\{1 + C_0 k_n^{-1/2} B_n(1) + o_p\left(k_n^{-1/2}\right)\right\}. \tag{S4.1}$$

Recall that $\widehat{\gamma} = \gamma + k_n^{-1/2}\{\Gamma_n + o_p(1)\} =: \gamma + k_n^{-1/2}\widetilde{\Gamma}_n$, where $\widetilde{\Gamma}_n = \Gamma_n + o_p(1)$. By $|\log(nq_n)| = o\left(\sqrt{k_n}\right)$, we have $k_n^{-1/2}\left|\log\{k_n/(nq_n)\}\right| = o(1)$. Thus, by Taylor expansion, we obtain

$$\begin{aligned}\left(\frac{k_n}{nq_n}\right)^{\widehat{\gamma}} &= \exp\left[\widehat{\gamma}\log\{k_n/(nq_n)\}\right]\\ &= \exp\left[\gamma\log\{k_n/(nq_n)\} + k_n^{-1/2}\log\{k_n/(nq_n)\}\widetilde{\Gamma}_n\right]\\ &= \left(\frac{k_n}{nq_n}\right)^{\gamma}\left[1 + k_n^{-1/2}\log\{k_n/(nq_n)\}\widetilde{\Gamma}_n + o_p\left(k_n^{-1/2}\log\{k_n/(nq_n)\}\right)\right].\end{aligned} \tag{S4.2}$$

From the second-order condition (3.3), we have

$$\begin{aligned}&H_Y(1/q_n|X=0)\\ &= H_Y(n/k_n|X=0)\left(\frac{k_n}{nq_n}\right)^{\gamma}\left[1 + A(n/k_n)\frac{\{k_n/(nq_n)\}^{\delta} - 1}{\delta} + o\{A(n/k_n)\}\right].\end{aligned} \tag{S4.3}$$

By $\sqrt{k_n}A(n/k_n) \to \eta \in \mathbb{R}$, $nq_n = o(k_n)$, $\delta < 0$ and $k_n^{-1/2}\left|\log\{k_n/(nq_n)\}\right| = o(1)$, we get

$$\frac{1 + k_n^{-1/2}\log\{k_n/(nq_n)\}\widetilde{\Gamma}_n + o_p\left(k_n^{-1/2}\log\{k_n/(nq_n)\}\right)}{1 + A(n/k_n)\left[\{k_n/(nq_n)\}^{\delta} - 1\right]\delta^{-1} + o\{A(n/k_n)\}} = 1 + o_p(1). \tag{S4.4}$$

Hence, by (S4.1)–(S4.4),

$$\frac{\widehat{\alpha}(\tau_n)}{\alpha(\tau_n)} = \left(\frac{k_n}{nq_n}\right)^{\widehat{\gamma}}\frac{\widehat{e}_{(n-k_n)}}{H_Y(1/q_n|X=0)}$$

$$
\begin{aligned}
&= \left(\frac{k_n}{nq_n}\right)^{\widehat{\gamma}-\gamma} \frac{\widehat{e}_{(n-k_n)}}{H_Y(n/k_n|X=0)\left(1+A(n/k_n)\left[\{k_n/(nq_n)\}^{\delta}-1\right]\delta^{-1}+o\{A(n/k_n)\}\right)} \\
&= \frac{1+k_n^{-1/2}\log\{k_n/(nq_n)\}\widetilde{\Gamma}_n + o_p\left(k_n^{-1/2}\log\{k_n/(nq_n)\}\right)}{1+A(n/k_n)\left[\{k_n/(nq_n)\}^{\delta}-1\right]\delta^{-1}+o\{A(n/k_n)\}} \cdot \frac{\widehat{e}_{(n-k_n)}}{H_Y(n/k_n|X=0)} \\
&= 1+k_n^{-1/2}\log\{k_n/(nq_n)\}\widetilde{\Gamma}_n + o_p\left(k_n^{-1/2}\log\{k_n/(nq_n)\}\right) + \Big[o\{A(n/k_n)\} \\
&\quad - A(n/k_n)\frac{\{k_n/(nq_n)\}^{\delta}-1}{\delta} + C_0 k_n^{-1/2}B_n(1) + o_p\left(k_n^{-1/2}\right)\Big]\{1+o_p(1)\}.
\end{aligned}
$$

Therefore,

$$
\frac{\sqrt{k_n}}{\log\{k_n/(nq_n)\}}\left\{\frac{\widehat{\alpha}(\tau_n)}{\alpha(\tau_n)}-1\right\} = \Gamma_n + o_p(1), \tag{S4.5}
$$

and complete the proof of Theorem 3.

S5 Proof of Theorem 4

Let $c_n = \sqrt{k_n}/\log\{k_n/(nq_n)\}$. Following (S4.5), we have

$$
\begin{aligned}
\widehat{Q}_Y(\tau_n|X) &= \widehat{\alpha}(\tau_n) + \langle X, \widehat{\beta}_0\rangle \\
&= \alpha(\tau_n) + \frac{\alpha(\tau_n)}{c_n}\{\Gamma_n + o_p(1)\} + \langle X, \beta_0\rangle + \left(\langle X, \widehat{\beta}_0\rangle - \langle X, \beta_0\rangle\right) \\
&= Q_Y(\tau_n|X) + \frac{\alpha(\tau_n)}{c_n}\{\Gamma_n + o_p(1)\} + O_p\left\{n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}}\right\}.
\end{aligned}
$$

By $\alpha(\tau_n) = H_Y(1/q_n|X=0) \approx Cq_n^{-\gamma}$, where $a_n \approx b_n (b_n \neq 0)$ means that $a_n/b_n \to 1$ when $n \to \infty$, and $c_n q_n^{\gamma} n^{-\frac{2\zeta-2-\nu}{4(\nu+2\zeta)}} \to 0$, we obtain

$$
\frac{c_n}{\alpha(\tau_n)}\left\{\widehat{Q}_Y(\tau_n|X) - Q_Y(\tau_n|X)\right\} = \Gamma_n + o_p(1),
$$

From $\alpha(\tau_n) = H_Y(1/q_n|X = 0) \approx Cq_n^{-\gamma}$, $||X|| = O_p(1)$ and $q_n^{\gamma} \to 0$, we have

$$\frac{Q_Y(\tau_n|X)}{\alpha(\tau_n)} = \frac{\alpha(\tau_n) + \langle X, \beta_0 \rangle}{\alpha(\tau_n)} = 1 + O_p(q_n^{\gamma}) = 1 + o_p(1),$$

and complete the proof of Theorem 4.

Bibliography

de Haan, L. and A. Ferreira (2006). *Extreme value theory: an introduction*. Springer Science & Business Media.

Drees, H. (1998). On smooth statistical tail functionals. *Scandinavian Journal of Statistics 25*(1), 187–210.

Hall, P. and J. L. Horowitz (2007). Methodology and convergence rates for functional linear regression. *The Annals of Statistics 35*(1), 70–91.

Kong, D., K. Xue, F. Yao, and H. H. Zhang (2016). Partially functional linear regression in high dimensions. *Biometrika 103*(1), 147–159.

Rényi, A. (1953). On the theory of order statistics. *Acta Mathematica Academiae Scientiarum Entiarum Hungarica 4*, 191–231.

Wang, H. J., D. Li, and X. He (2012). Estimation of high conditional quantiles for heavy-tailed distributions. *Journal of the American Statistical Association 107*(500), 1453–1464.

Yao, F., S. Sue-Chee, and F. Wang (2017). Regularized partially functional quantile regression. *Journal of Multivariate Analysis 156*, 39–56.