BIOMETRICS 64, 1–36 DOI: 000
December 2022

Supporting Information for "Quantile Regression for Nonignorable Missing Data with its Application of Analyzing Electronic Medical Records" by Yu, Zhong, Feng and Wei

Aiai Yu[1,*], Yujie Zhong[1,], Xingdong Feng[1,***], and Ying Wei[2,****]**

[1]School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, China

[2]Department of Biostatistics, Columbia University, New York, NY 10032, U.S.A.

**email:* yu.aiai2018@163.sufe.edu.cn

***email:* zhong.yujie@mail.shufe.edu.cn

****email:* feng.xingdong@mail.shufe.edu.cn

*****email:* yw2148@cumc.columbia.edu

This paper has been submitted for consideration for publication in *Biometrics*

The Supporting Information includes an computing algorithm for the proposed method, the detailed proofs of Theorems 1–3, additional simulation results and empirical analysis results. The details of the proposed computing algorithm for proposed method are given in Web Appendix A. The proofs of Theorems 1–3 and related Lemmas are given in Web Appendices B, C and D, respectively. The additional simulation results are reported in Web Appendix E and the additional empirical analysis results are summarized in Web Appendix F.

Web Appendix A

The detailed computing algorithm to obtain the proposed estimates $\widehat{\boldsymbol{\xi}}_n$ is presented as follows.

Step 1. Initialize parameters $\boldsymbol{\phi}$ by solving the estimating equation $\sum_{i=1}^{n}\delta_i \boldsymbol{x}_i \Psi_\tau\left(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}\right) = 0$ with complete observations over the evenly-spaced find grid of quantile levels Ω, denoted as $\widehat{\boldsymbol{\phi}}^{(0)} = \left(\widehat{\boldsymbol{\beta}}_{\tau_1}^{(0)\top}, \ldots, \widehat{\boldsymbol{\beta}}_{\tau_{k_n}}^{(0)\top}\right)^\top$.

Step 2. Suppose $\widehat{\boldsymbol{\phi}}^{(t-1)} = \left(\widehat{\boldsymbol{\beta}}_{\tau_1}^{(t-1)\top}, \ldots, \widehat{\boldsymbol{\beta}}_{\tau_{k_n}}^{(t-1)\top}\right)^\top$ is the estimated quantile coefficients at the $(t-1)$th step, we approximate the conditional quantile functions of $\boldsymbol{x}_i^\top \boldsymbol{\beta}(\tau)$ by piece-wise linear spline expanding from $\widehat{\boldsymbol{\phi}}^{(t-1)}$, i.e.

$$\boldsymbol{x}_i^\top \widetilde{\boldsymbol{\beta}}^{(t-1)}(\tau) = \begin{cases} \boldsymbol{x}_i^\top \widehat{\boldsymbol{\beta}}_{\tau_1}^{(t-1)} & \tau < \tau_1 \\ \boldsymbol{x}_i^\top \left(\frac{\tau_{k+1}-\tau}{\tau_{k+1}-\tau_k}\widehat{\boldsymbol{\beta}}_{\tau_k}^{(t-1)} + \frac{\tau-\tau_k}{\tau_{k+1}-\tau_k}\widehat{\boldsymbol{\beta}}_{\tau_{k+1}}^{(t-1)}\right) & \tau \in [\tau_k, \tau_{k+1}) \text{ for } k = 1, 2, \ldots, (k_n - 1) \\ \boldsymbol{x}_i^\top \widehat{\boldsymbol{\beta}}_{\tau_{k_n}}^{(t-1)} & \tau \geq \tau_{k_n} \end{cases} \cdot$$

Step 3. Use Monte Carlo integration to approximate the conditional expectations in $\boldsymbol{S}_n(\boldsymbol{\theta}, \boldsymbol{\beta}(\tau))$. e.g.

$$E_y\{1 - \pi(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}) \mid \boldsymbol{x}_i\} \approx \frac{1}{m}\sum_{l=1}^{m}\{1 - \pi(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta})\},$$

where $\widetilde{y}_i^l$ $(l = 1, ..., m)$ are m random samples drawn from the quantile function $\boldsymbol{x}_i^\top \widetilde{\boldsymbol{\beta}}^{(t-1)}(\tau)$.

Step 4. Update $\widehat{\boldsymbol{\theta}}^{(t)}$ by solving the approximated $\boldsymbol{S}_n(\boldsymbol{\theta}, \widetilde{\boldsymbol{\beta}}^{(t-1)}(\tau)) = 0$ in Step 3. The approximated estimating equation used to estimate $\boldsymbol{\theta}$ is

$$\frac{1}{n}\sum_{i=1}^{n}\left[\delta_i \boldsymbol{s}\left(\boldsymbol{\theta}; \delta_i, \boldsymbol{x}_{1i}, y_i\right) + (1 - \delta_i)\frac{\sum_{l=1}^{m}\{1 - \pi(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \widehat{\boldsymbol{\theta}}^{(t-1)})\}\boldsymbol{s}\left(\boldsymbol{\theta}; \delta_i, \boldsymbol{x}_{1i}, \widetilde{y}_i^l\right)}{\sum_{l=1}^{m}\{1 - \pi(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \widehat{\boldsymbol{\theta}}^{(t-1)})\}}\right] = 0\,,$$

where

$$\boldsymbol{s}\left(\boldsymbol{\theta};\delta,\boldsymbol{x}_1,y\right)=\begin{cases}\pi\left(\boldsymbol{x}_1,y;\boldsymbol{\theta}\right)^{-1}\partial\pi\left(\boldsymbol{x}_1,y;\boldsymbol{\theta}\right)/\partial\boldsymbol{\theta} & \delta=1\\ -\left\{1-\pi\left(\boldsymbol{x}_1,y;\boldsymbol{\theta}\right)\right\}^{-1}\partial\pi\left(\boldsymbol{x}_1,y;\boldsymbol{\theta}\right)/\partial\boldsymbol{\theta} & \delta=0\end{cases}.$$

Here a weighted logistic regression is considered to obtain the estimates of $\widehat{\boldsymbol{\theta}}^{(t)}$, where the response is δ_i, the observations $(\boldsymbol{x}_{1i},y_i)$ are weighted with the constant 1 if $\delta_i=1$, and the simulated pairs $(\boldsymbol{x}_{1i},\widetilde{y}_i^l)$ are weighted with the ratios $\{1-\pi(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\widehat{\boldsymbol{\theta}}^{(t-1)})\}/\sum_{l=1}^m\{1-\pi(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\widehat{\boldsymbol{\theta}}^{(t-1)})\}$ if $\delta_i=0$.

Step 5. Update $\widehat{\boldsymbol{\phi}}^{(t)}$ by solving the equations

$$\boldsymbol{M}_n(\widehat{\boldsymbol{\theta}}^{(t)},\boldsymbol{\phi})=\frac{1}{n}\sum_{i=1}^n\frac{\delta_i}{\pi\left(\boldsymbol{x}_{1i},y_i;\widehat{\boldsymbol{\theta}}^{(t)}\right)}\Psi\left(y_i-\boldsymbol{x}_i^\top\boldsymbol{\phi}\right)\otimes\boldsymbol{x}_i=0\,,$$

where $\otimes$ is the Kronecker product and $\Psi\left(y_i-\boldsymbol{x}_i^\top\boldsymbol{\phi}\right)\otimes\boldsymbol{x}_i$ is a $(p\times k_n)$-dimensional vector, which consists of k_n components $\{\Psi_{\tau_k}\left(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k}\right)\boldsymbol{x}_i\}_{k=1,\ldots,k_n}$. We can use the R package *quantreg* to obtain the estimate $\widehat{\boldsymbol{\phi}}^{(t)}$ by minimizing the following function instead of directly solving the estimating equation

$$\sum_{i=1}^n\frac{\delta_i}{\pi(\boldsymbol{x}_{1i},y_i;\widehat{\boldsymbol{\theta}}^{(t)})}\rho_{\tau_k}\left(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}\right).$$

Step 6. Repeat Steps 2 to 5 until some stopping criteria are satisfied.

Web Appendix B

Let $\|\cdot\|$ be the standard L_2 norm, $\|\cdot\|_{\psi_1}$ be the sub-exponential norm, and q be the dimension of the vector $\boldsymbol{\xi}$, that is, $q=(d+1)+k_n\cdot p$. Denote $\mathcal{X},\mathcal{Y}$ as the support of $\boldsymbol{x}$, y, respectively. Without loss of generality, we assume in our proof that $\tau_k=k/(k_n+1)$ such that $\Omega=\{1/(k_n+1),2/(k_n+1),\ldots,k_n/(k_n+1)\}$. Denote

$$\begin{aligned}\widetilde{\boldsymbol{S}}_n(\boldsymbol{\xi})&=\boldsymbol{S}_n(\boldsymbol{\theta},\widetilde{\boldsymbol{\beta}}(\tau))\\&=\frac{1}{n}\sum_{i=1}^n\left[\delta_i s\left(\boldsymbol{\theta};\delta_i,\boldsymbol{x}_{1i},y_i\right)+(1-\delta_i)\frac{\int\left\{1-\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}\right)\right\}s\left(\boldsymbol{\theta};\delta_i,\boldsymbol{x}_{1i},y\right)\widetilde{f}(y\mid\boldsymbol{x}_i;\boldsymbol{\phi})dy}{\int\left\{1-\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}\right)\right\}\widetilde{f}(y\mid\boldsymbol{x}_i;\boldsymbol{\phi})dy}\right],\end{aligned}$$

where $\widetilde{f}(y\mid\boldsymbol{x}_i;\boldsymbol{\phi})=f(y\mid\boldsymbol{x}_i^\top\widetilde{\boldsymbol{\beta}}(\tau))$. Let

$$\widehat{\boldsymbol{S}}_n(\boldsymbol{\xi})=\frac{1}{n}\sum_{i=1}^n\left[\delta_i\boldsymbol{s}\left(\boldsymbol{\theta};\delta_i,\boldsymbol{x}_{1i},y_i\right)+(1-\delta_i)\frac{\sum_{l=1}^m\{1-\pi(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta})\}\boldsymbol{s}\left(\boldsymbol{\theta};\delta_i,\boldsymbol{x}_{1i},\widetilde{y}_i^l\right)}{\sum_{l=1}^m\{1-\pi(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta})\}}\right],$$

and $\widehat{\boldsymbol{H}}n(\boldsymbol{\xi}) = \left(\widehat{\boldsymbol{S}}_n^\top(\boldsymbol{\xi}), M_n^\top(\boldsymbol{\xi})\right)^\top$. The proposed estimate $\widehat{\boldsymbol{\xi}}_n = \left(\widehat{\boldsymbol{\theta}}_n^\top, \widehat{\boldsymbol{\phi}}_n^\top\right)^\top$ is obtained by solving the approximated estimating equation $\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}) = 0$. For simple presentation, we further let $\boldsymbol{H}_\tau^0(\boldsymbol{\theta}, \boldsymbol{\beta}) = \left(\boldsymbol{S}^{0\top}(\boldsymbol{\theta}), \boldsymbol{M}_\tau^\top(\boldsymbol{\theta}, \boldsymbol{\beta})\right)^\top$, $\boldsymbol{H}^0(\boldsymbol{\xi}) = \left(\boldsymbol{S}^{0\top}(\boldsymbol{\theta}), \boldsymbol{M}^\top(\boldsymbol{\xi})\right)^\top$, $\boldsymbol{H}(\boldsymbol{\xi}) = (\boldsymbol{S}^\top(\boldsymbol{\xi}), \boldsymbol{M}^\top(\boldsymbol{\xi}))^\top$ and $\widetilde{\boldsymbol{H}}(\boldsymbol{\xi}) = \left(\widetilde{\boldsymbol{S}}^\top(\boldsymbol{\xi}), \boldsymbol{M}^\top(\boldsymbol{\xi})\right)^\top$ be the expectations of $\boldsymbol{H}_{n,\tau}^0(\boldsymbol{\theta}, \boldsymbol{\beta})$, $\boldsymbol{H}_n^0(\boldsymbol{\xi})$, $\boldsymbol{H}_n(\boldsymbol{\xi})$ and $\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi})$, respectively. We first introduce Lemmas 1, 2 and 3 which will be used for the proof of Theorem 1.

LEMMA 1: *Under Assumptions 2-3, for $k_n \to \infty, k_n/n \to 0$, we have*

$$k_n^{-1} \left\| \boldsymbol{H}\left(\boldsymbol{\xi}_0\right) - \boldsymbol{H}^0\left(\boldsymbol{\xi}_0\right) \right\| = o(1) \tag{A.1}$$

Proof. Note that

$$\begin{aligned}
&k_n^{-1} \left\| \boldsymbol{H}\left(\boldsymbol{\xi}_0\right) - \boldsymbol{H}^0\left(\boldsymbol{\xi}_0\right) \right\| \\
&\le k_n^{-1} \left\| \boldsymbol{S}(\boldsymbol{\xi}_0) - \boldsymbol{S}^0(\boldsymbol{\theta}_0) \right\| + k_n^{-1} \left\| \boldsymbol{M}(\boldsymbol{\xi}_0) - \boldsymbol{M}(\boldsymbol{\xi}_0) \right\| = k_n^{-1} \left\| \boldsymbol{S}(\boldsymbol{\xi}_0) - \boldsymbol{S}^0(\boldsymbol{\theta}_0) \right\|
\end{aligned}$$

so it suffices to show that

$$k_n^{-1} \left\| \boldsymbol{S}(\boldsymbol{\xi}_0) - \boldsymbol{S}^0(\boldsymbol{\theta}_0) \right\| = o(1). \tag{A.2}$$

Let

$$\begin{aligned}
A_i(\boldsymbol{\xi}_0) &= \int_y \left\{1 - \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\right\} f\left(y | \boldsymbol{x}_i; \boldsymbol{\beta}_0(\tau)\right) dy, \\
\widetilde{A}_i(\boldsymbol{\xi}_0) &= \int_y \left\{1 - \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\right\} \widetilde{f}\left(y | \boldsymbol{x}_i; \boldsymbol{\phi}_0\right) dy.
\end{aligned}$$

Note that

$$\begin{aligned}
&\left\| \boldsymbol{S}(\boldsymbol{\xi}_0) - \boldsymbol{S}^0(\boldsymbol{\theta}_0) \right\| \\
=&\Bigg\| \frac{1}{n} \sum_{i=1}^n E_{\boldsymbol{x}_i} \Bigg\{ (1 - \delta_i) \int_y \left\{1 - \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\right\} s\left(\boldsymbol{\theta}_0; \delta_i, \boldsymbol{x}_{1i}, y\right) \\
&\cdot \left(\frac{f(y|\boldsymbol{x}_i; \boldsymbol{\beta}_0(\tau))}{A_i(\boldsymbol{\xi}_0)} - \frac{\widetilde{f}(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0)}{\widetilde{A}_i(\boldsymbol{\xi}_0)} \right) dy \Bigg\} \Bigg\| \\
=&\left\| \frac{1}{n} \sum_{i=1}^n E_{\boldsymbol{x}_i} \left\{ (1 - \delta_i) \int_y \partial_{\boldsymbol{\theta}_0} \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right) \left(\frac{f(y|\boldsymbol{x}_i; \boldsymbol{\beta}_0(\tau))}{A_i(\boldsymbol{\xi}_0)} - \frac{\widetilde{f}(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0)}{\widetilde{A}_i(\boldsymbol{\xi}_0)} \right) dy \right\} \right\| \\
\le& \frac{1}{n} \sum_{i=1}^n E_{\boldsymbol{x}_i} \left\{ \int_y \left\| \partial_{\boldsymbol{\theta}_0} \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right) \right\| \left| \frac{f(y|\boldsymbol{x}_i; \boldsymbol{\beta}_0(\tau))}{A_i(\boldsymbol{\xi}_0)} - \frac{\widetilde{f}(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0)}{\widetilde{A}_i(\boldsymbol{\xi}_0)} \right| \right\} dy \\
\le& \frac{1}{n} \sum_{i=1}^n E_{\boldsymbol{x}_i} \left\{ \frac{1}{\widetilde{A}_i(\boldsymbol{\xi}_0)} \int_y \left\| \partial_{\boldsymbol{\theta}_0} \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right) \right\| \left| f(y|\boldsymbol{x}_i; \boldsymbol{\beta}_0(\tau)) - \widetilde{f}(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0) \right| dy \right\}
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{n}\sum_{i=1}^{n}E_{\boldsymbol{x}_i}\left\{\frac{\left|A_i(\boldsymbol{\xi}_0)-\widetilde{A}_i(\boldsymbol{\xi}_0)\right|}{A_i(\boldsymbol{\xi}_0)\widetilde{A}_i(\boldsymbol{\xi}_0)}\int_y\left\|\partial_{\boldsymbol{\theta}_0}\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}_0\right)\right\|f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau))dy\right\}\\
\widehat{=}&SS_1+SS_2,
\end{aligned}
$$

so (A.2) holds if $k_n^{-1}SS_1=o(1)$ and $k_n^{-1}SS_2=o(1)$ hold.

Under Assumption 3, there exist positive constants $U_{\pi'}$, L_π and U_π such that the absolute value of each element of $\partial_{\boldsymbol{\theta}_0}\pi\left(\boldsymbol{x}_1,y;\boldsymbol{\theta}_0\right)$ is less than $U_{\pi'}$, and $L_\pi\le\pi\left(\boldsymbol{x}_1,y;\boldsymbol{\theta}_0\right)\le U_\pi$ for any $\boldsymbol{x}\in\mathcal{X}$ and $y\in\mathcal{Y}$. Therefore, we have $\left\|\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_1,y;\boldsymbol{\theta}_0\right)\right\|\le(d+1)U_{\pi'}$ and $\widetilde{A}_i(\boldsymbol{\xi}_0)\ge1-U_\pi$ for any $\boldsymbol{x}\in\mathcal{X}$ and $y\in\mathcal{Y}$. Thus,

$$
\begin{aligned}
SS_1\le&\frac{1}{n}\frac{1}{1-U_\pi}\sum_{i=1}^{n}E_{\boldsymbol{x}_i}\left\{\int_y\left\|\partial_{\boldsymbol{\theta}_0}\pi\left(\boldsymbol{x}_1,y;\boldsymbol{\theta}_0\right)\right\|\left|f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau))-\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0)\right|dy\right\}\\
\le&\frac{1}{n}\frac{(d+1)U_{\pi'}}{1-U_\pi}\sum_{i=1}^{n}E_{\boldsymbol{x}_i}\left\{\int_y\left|f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau))-\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0)\right|dy\right\}\\
\le&\frac{1}{n}\frac{(d+1)U_{\pi'}}{1-U_\pi}\sum_{i=1}^{n}E_{\boldsymbol{x}_i}\left\{\int_y\boldsymbol{I}\left(y>\boldsymbol{x}_i^T\boldsymbol{\beta}_{0,k_n/(k_n+1)}\right)f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau))dy\right\}\\
+&\frac{1}{n}\frac{(d+1)U_{\pi'}}{1-U_\pi}\sum_{i=1}^{n}E_{\boldsymbol{x}_i}\left\{\int_y\boldsymbol{I}\left(y<\boldsymbol{x}_i^T\boldsymbol{\beta}_{0,1/(k_n+1)}\right)f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau))dy\right\}\\
+&\frac{1}{n}\frac{(d+1)U_{\pi'}}{1-U_\pi}\sum_{i=1}^{n}E_{\boldsymbol{x}_i}\left\{\int_y\boldsymbol{I}\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_{0,1/(k_n+1)}\le y\le\boldsymbol{x}_i^T\boldsymbol{\beta}_{0,k_n/(k_n+1)}\right)\right.\\
&\left.\cdot\left|f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau))-\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0)\right|dy\right\}\\
\widehat{=}&SS_{11}+SS_{12}+SS_{13}.
\end{aligned}
$$

Since

$$
\begin{aligned}
&E_{\boldsymbol{x}_i}\left\{\int_y\boldsymbol{I}\left(y>\boldsymbol{x}_i^T\boldsymbol{\beta}_{0,k_n/(k_n+1)}\right)f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau))dy\right\}\\
=&E_{\boldsymbol{x}_i}\left[E_y\left\{\boldsymbol{I}(y>\boldsymbol{x}_i^T\boldsymbol{\beta}_{0,k_n/(k_n+1)})|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau)\right\}\right]\\
=&\frac{1}{1+k_n},
\end{aligned}
$$

we then have $SS_{11}=\frac{1}{1+k_n}\cdot\frac{(d+1)U_{\pi'}}{1-U_\pi}=o(1)$ and thus $k_n^{-1}SS_{11}=o(k_n^{-1})=o(1)$. With the similar argument, we can show that $k_n^{-1}SS_{12}=o(1)$. In what follows, we show that $k_n^{-1}SS_{13}=o(1)$.

Denote $g(\boldsymbol{x}_i)=\int_y|f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau))-\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0)|\times\boldsymbol{I}\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_{0,1/(k_n+1)}\le y\le\boldsymbol{x}_i^T\boldsymbol{\beta}_{0,k_n/(k_n+1)}\right)dy$. It is easy to show that $|g(\boldsymbol{x}_i)|\le2$ for any $\boldsymbol{x}_i\in\mathcal{X}$. Hence, by dominated convergence theorem, a sufficient condition for $k_n^{-1}SS_{13} = o(1)$ is that $g(\boldsymbol{x}_i) = o_p(1)$ holds for all i. In order to prove $g(\boldsymbol{x}_i) = o_p(1)$, by Scheffe's theorem, we only need to show that, for any $y \in \mathcal{Y}$, the following equation holds

$$\max_i \left| f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau)) - \widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0) \right| \times \boldsymbol{I}\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_{0,1/(k_n+1)} \le y \le \boldsymbol{x}_i^T\boldsymbol{\beta}_{0,k_n/(k_n+1)}\right) = o_p(1). \quad \text{(A.3)}$$

Note that $F_{\boldsymbol{x}_i}(y) = \inf\{\tau : \boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau} \ge y\}$ is the quantile level of y with respect to the probability measure induced by the quantile function $\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau}$, and $h_{\boldsymbol{x}_i}(\tau) = 1/\boldsymbol{x}_i^\top\boldsymbol{\beta}_0'(\tau)$ is the density of y at the τ-th quantile. For any y that is bounded between $\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,1/(k_n+1)}$ and $\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,k_n/(k_n+1)}$, there exists an integer k_i such that $\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,k_i/(k_n+1)} \le y \le \boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,(k_i+1)/(k_n+1)}$. Consequently, the left side of (A.3) is equivalent to

$$\begin{aligned}
&\max_i \left| \frac{1}{(k_n+1)\boldsymbol{x}_i^\top\left(\boldsymbol{\beta}_{0,(k_i+1)/(k_n+1)} - \boldsymbol{\beta}_{0,k_i/(k_n+1)}\right)} - \frac{1}{\boldsymbol{x}_i^\top\boldsymbol{\beta}_0'\{F_{\boldsymbol{x}_i}(y)\}} \right| \\
&= \max_i \left| h_{\boldsymbol{x}_i}(\tau_i^*) - h_{\boldsymbol{x}_i}\{F_{\boldsymbol{x}_i}(y)\} \right| \qquad \text{for some } k_i/(k_n+1) < \tau_i^* < (k_i+1)/(k_n+1) \\
&= \max_i \left| h'_{\boldsymbol{x}_i}\{k_i/(k_n+1)\} O\left(k_n^{-1}\right) \right|.
\end{aligned} \quad \text{(A.4)}$$

By Assumption 2, we then have

$$\begin{aligned}
h'_{\boldsymbol{x}_i}\{k_i/(k_n+1)\} &< M\left(\frac{k_i}{k_n+1}\right)^{v_1}\left(1 - \frac{k_i}{k_n+1}\right)^{v_2} \\
&< M\left(\frac{1}{k_n+1}\right)^{v_1}\left(1 - \frac{1}{k_n+1}\right)^{v_2} \\
&\quad + M\left(\frac{k_n}{k_n+1}\right)^{v_1}\left(1 - \frac{k_n}{k_n+1}\right)^{v_2} \\
&= O\left(k_n^{-v_1}\right) + O\left(k_n^{-v_2}\right).
\end{aligned} \quad \text{(A.5)}$$

Since $v_1, v_2 > -1$, (A.3) is implied by (A.4) and (A.5). Hence, we obtain $k_n^{-1}SS_1 = o(1)$. Moreover, as $k_n \to \infty$, we know that

$$\max_i \int_y \left| f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau)) - \widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0) \right| dy = \frac{2}{k_n+1} + o_p(1) = o_p(1). \quad \text{(A.6)}$$

In what follows, we show that $k_n^{-1}SS_2 = o(1)$. By (A.6), we have

$$\max_i \frac{\left|A_i(\boldsymbol{\xi}_0) - \widetilde{A}_i(\boldsymbol{\xi}_0)\right|}{A_i(\boldsymbol{\xi}_0)\widetilde{A}_i(\boldsymbol{\xi}_0)} \le \frac{1}{(1-U_\pi)^2} \max_i \int_y \left| f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau)) - \widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0) \right| dy = o_p(1),$$

and the right term of the above inequality is between 0 and 2. Thus, according to dominated convergence theorem, we have

$$
\begin{aligned}
SS_2 &= \frac{1}{n}\sum_{i=1}^{n} E_{\boldsymbol{x}_i}\left\{\frac{\left|A_i(\boldsymbol{\xi}_0)-\widetilde{A}_i(\boldsymbol{\xi}_0)\right|}{A_i(\boldsymbol{\xi}_0)\widetilde{A}_i(\boldsymbol{\xi}_0)}\int_y \left\|\partial_{\boldsymbol{\theta}_0}\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}_0\right)\right\| f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau))dy\right\}\\
&\leq \frac{(d+1)U_{\pi'}}{n}\sum_{i=1}^{n} E_{\boldsymbol{x}_i}\left\{\frac{\left|A_i(\boldsymbol{\xi}_0)-\widetilde{A}_i(\boldsymbol{\xi}_0)\right|}{A_i(\boldsymbol{\xi}_0)\widetilde{A}_i(\boldsymbol{\xi}_0)}\right\}\\
&\leq (d+1)U_{\pi'}E_{\boldsymbol{x}_i}\left\{\max_i \frac{\left|A_i(\boldsymbol{\xi}_0)-\widetilde{A}_i(\boldsymbol{\xi}_0)\right|}{A_i(\boldsymbol{\xi}_0)\widetilde{A}_i(\boldsymbol{\xi}_0)}\right\}\\
&= o(1).
\end{aligned}
$$

Therefore, we have $k_n^{-1}SS_2 = o(1)$. The proof of Lemma 1 is hence completed. □

LEMMA 2: *Under Assumptions 1-3, for $k_n \to \infty, k_n/n \to 0$, we have*

$$
\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} k_n^{-1}\left\|\boldsymbol{H}_n(\boldsymbol{\xi})-\boldsymbol{H}(\boldsymbol{\xi})\right\| = o_p(1) \quad \text{as} \quad n\to\infty. \tag{A.7}
$$

Proof. For convenience, we assume that $\{y_i : i = 1,2,\ldots,n_0\}$ are observed, $\{y_i : i = n_0+1, n_0+2,\ldots,n\}$ are missing and denote $n_1 = n - n_0$. Note that for every $\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}$,

$$
\left\|\boldsymbol{H}_n(\boldsymbol{\xi})-\boldsymbol{H}(\boldsymbol{\xi})\right\| \leq \left\|\boldsymbol{S}_n(\boldsymbol{\xi})-\boldsymbol{S}(\boldsymbol{\xi})\right\| + \left\|\boldsymbol{M}_n(\boldsymbol{\xi})-\boldsymbol{M}(\boldsymbol{\xi})\right\|,
$$

so a sufficient condition for (A.7) is $\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} k_n^{-1}\|\boldsymbol{S}_n(\boldsymbol{\xi})-\boldsymbol{S}(\boldsymbol{\xi})\| = o_p(1)$ and $\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} k_n^{-1}\|\boldsymbol{M}_n(\boldsymbol{\xi})-\boldsymbol{M}(\boldsymbol{\xi})\| = o_p(1)$. In what follows, we will show $\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} k_n^{-1}\left\|\boldsymbol{S}_n(\boldsymbol{\xi})-\boldsymbol{S}(\boldsymbol{\xi})\right\| = o_p(1)$. First, we denote

$$
\begin{aligned}
\boldsymbol{S}_n(\boldsymbol{\xi}) &= \frac{1}{n}\sum_{i=1}^{n_0}\boldsymbol{s}\left(\boldsymbol{\theta};\delta_i=1,\boldsymbol{x}_{1i},y_i\right)+\frac{1}{n}\sum_{i=n_0+1}^{n} g(\boldsymbol{x_i};\boldsymbol{\xi})\\
&\widehat{=}\frac{n_0}{n}\boldsymbol{S}_n^{(1)}(\boldsymbol{\theta})+\frac{n_1}{n}\boldsymbol{S}_n^{(2)}(\boldsymbol{\xi}).
\end{aligned}
$$

where

$$
g(\boldsymbol{x}_i;\boldsymbol{\xi}) = \int s\left(\boldsymbol{\theta};\delta_i=0,\boldsymbol{x}_{1i},y\right)\frac{\left\{1-\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}\right)\right\}\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})}{\int\left\{1-\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}\right)\right\}\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})dy}dy.
$$

Similarly, we denote

$$
\boldsymbol{S}(\boldsymbol{\xi})\widehat{=}\frac{n_0}{n}\boldsymbol{S}^{(1)}(\boldsymbol{\theta})+\frac{n_1}{n}\boldsymbol{S}^{(2)}(\boldsymbol{\xi}),
$$

and $\boldsymbol{s}\left(\boldsymbol{\theta};\delta_i=1,\boldsymbol{x}_{1i},y_i\right) = (\boldsymbol{s}_{i,1}^1(\boldsymbol{\theta}),\ldots,\boldsymbol{s}_{i,d+1}^1(\boldsymbol{\theta}))^T$, $\boldsymbol{S}^{(1)}(\boldsymbol{\theta}) = (\boldsymbol{s}_1^1(\boldsymbol{\theta}),\ldots,\boldsymbol{s}_{d+1}^1(\boldsymbol{\theta}))^T$. Under Assumption 3, we have that $\boldsymbol{s}_{i,j}^1(\boldsymbol{\theta})$ is continuous at each $\boldsymbol{\theta}\in\Theta$ with $j = 1,\ldots,d+1$, $\sup_{\boldsymbol{\theta}\in\Theta}\left|\boldsymbol{s}_{i,j}^1(\boldsymbol{\theta})\right| \leq$

$\frac{U_{\pi'}}{L_\pi}$, and then $\sup_{\boldsymbol{\theta}\in\Theta} E\left|\boldsymbol{s}_{i,j}^1(\boldsymbol{\theta})\right| \leq \frac{U_{\pi'}}{L_\pi}$. According to Lemma 2.4 of Engle (1994), we have

$$\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}}\left|\frac{1}{n_0}\sum_{i=1}^{n_0}\boldsymbol{s}_{i,j}^1(\boldsymbol{\theta})-\boldsymbol{s}_j^1(\boldsymbol{\theta})\right| = \sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{n_0}\sum_{i=1}^{n_0}\boldsymbol{s}_{i,j}^1(\boldsymbol{\theta})-\boldsymbol{s}_j^1(\boldsymbol{\theta})\right| = o_p(1) \qquad \text{for} \quad j=1,\ldots,d+1.$$

Thus, we obtain that

$$\begin{aligned}
&\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}}\left\|\boldsymbol{S}_n^{(1)}(\boldsymbol{\theta})-\boldsymbol{S}^{(1)}(\boldsymbol{\theta})\right\| \\
&\leq \sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}}\sum_{j=1}^{d+1}\left|\frac{1}{n_0}\sum_{i=1}^{n_0}\boldsymbol{s}_{i,j}^1(\boldsymbol{\theta})-\boldsymbol{s}_j^1(\boldsymbol{\theta})\right| = o_p(1).
\end{aligned} \tag{A.8}$$

In what follows, we will show $\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} k_n^{-1}\left\|\boldsymbol{S}_n^{(2)}(\boldsymbol{\xi})-\boldsymbol{S}^{(2)}(\boldsymbol{\xi})\right\| = o_p(1)$, which is equivalent to

$$Pr\left(\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} k_n^{-1}\left\|\boldsymbol{S}_n^{(2)}(\boldsymbol{\xi})-\boldsymbol{S}^{(2)}(\boldsymbol{\xi})\right\| > \epsilon\right) \to 0 \tag{A.9}$$

for any $\epsilon > 0$ as $n \to \infty$. We will show (A.9) using Huber's chaining augment. Without loss of generality, we assume $\Theta_{\boldsymbol{\xi}} = \bigcup_{k=1}^{k_n}\{\boldsymbol{\xi}_{\tau_k} : \|\boldsymbol{\xi}_{\tau_k}-\boldsymbol{\xi}_{0,\tau_k}\| < 1\}$ where $\boldsymbol{\xi}_{\tau_k} = (\boldsymbol{\theta}^T, \boldsymbol{\beta}_{\tau_k}^T)^T$, and $\boldsymbol{\xi}_{0,\tau_k} = (\boldsymbol{\theta}_0^T, \boldsymbol{\beta}_{0,\tau_k}^T)^T$. We partition the parameter space $\Theta_{\boldsymbol{\xi}}$ into L_n disjoint small cubes Γ_l with diameters less than $q_n = C_1 k_n/n$, for some constant C_1. Let $\boldsymbol{\gamma}_l = (\boldsymbol{\theta}_l^\top, \boldsymbol{\phi}_l^\top)^\top$ be the center of the l-th cube Γ_l. The probability of the left side of (A.9) is bounded by the sum of the following two probabilities, $P_1 + P_2$, where

$$P_1 = Pr\left(\max_{1\leq l\leq L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}\left\|\boldsymbol{S}_n^{(2)}(\boldsymbol{\xi})-\boldsymbol{S}_n^{(2)}(\boldsymbol{\gamma}_l)-\boldsymbol{S}^{(2)}(\boldsymbol{\xi})+\boldsymbol{S}^{(2)}(\boldsymbol{\gamma}_l)\right\| > \epsilon/2\right),$$

and

$$P_2 = Pr\left(\max_{1\leq l\leq L_n} k_n^{-1}\left\|\boldsymbol{S}_n^{(2)}(\boldsymbol{\gamma}_l)-\boldsymbol{S}^{(2)}(\boldsymbol{\gamma}_l)\right\| > \epsilon/2\right).$$

We first note that

$$\begin{aligned}
&\left\|\boldsymbol{S}_n^{(2)}(\boldsymbol{\xi})-\boldsymbol{S}_n^{(2)}(\boldsymbol{\gamma}_l)\right\| \\
&= \left\|\frac{1}{n}\sum_{i=1}^n(1-\delta_i)\left\{\int_y \partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}\right)\frac{\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})}{\widetilde{A}_i(\boldsymbol{\xi})}dy - \int_y \partial_{\boldsymbol{\theta}_l}\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}_l\right)\frac{\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_l)}{\widetilde{A}_i(\boldsymbol{\gamma}_l)}dy\right\}\right\| \\
&\leq \frac{1}{n}\sum_{i=1}^n\int_y\left\|\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}\right)-\partial_{\boldsymbol{\theta}_l}\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}_l\right)\right\|\frac{\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})}{\widetilde{A}_i(\boldsymbol{\xi})}dy \\
&+ \frac{1}{n}\sum_{i=1}^n\int_y\left\|\partial_{\boldsymbol{\theta}_l}\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}_l\right)\right\|\left|\frac{\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})}{\widetilde{A}_i(\boldsymbol{\xi})}-\frac{\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_l)}{\widetilde{A}_i(\boldsymbol{\gamma}_l)}\right|dy \\
&\widehat{=} SS_{21}+SS_{22}.
\end{aligned}$$

Therefore,

$$\max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}\left\|\widetilde{\boldsymbol{S}}_n^{(2)}(\boldsymbol{\xi})-\widetilde{\boldsymbol{S}}_n^{(2)}(\boldsymbol{\gamma}_l)\right\|\le \max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}SS_{21}+\max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}SS_{22}. \tag{A.10}$$

Moreover, under Assumption 3, there exist constants $U_\pi, U_{\pi'}$ such that

$$\begin{aligned}\max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} SS_{21} &\le \max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l}\frac{1}{1-U_\pi}\times\frac{1}{n}\sum_{i=1}^n\int_y\left(\|\partial_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_{1i},y;\boldsymbol{\theta})\|+\|\partial_{\boldsymbol{\theta}_l}\pi(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}_l)\|\right)\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})dy\\ &\le \frac{2(d+1)U_{\pi'}}{1-U_\pi}\times\frac{1}{n}\sum_{i=1}^n\int_y\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})dy\\ &\le \frac{2(d+1)U_{\pi'}}{1-U_\pi}\\ &= O_p(1),\end{aligned}$$

so $\max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}SS_{21}=O_p(k_n^{-1})=o_p(1)$.

On the other hand, we have

$$\begin{aligned}\max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} SS_{22} &\le \max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l}(d+1)U_{\pi'}\times\frac{1}{n}\sum_{i=1}^n\int_y\left|\frac{\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})}{\widetilde{A}_i(\boldsymbol{\xi})}-\frac{\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_l)}{\widetilde{A}_i(\boldsymbol{\gamma}_l)}\right|dy\\ &\le \max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l}\frac{(d+1)U_{\pi'}}{1-U_\pi}\times\frac{1}{n}\sum_{i=1}^n\int_y\left\{\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})+\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_l)\right\}dy\\ &\le \frac{2(d+1)U_{\pi'}}{1-U_\pi}\\ &= O_p(1).\end{aligned}$$

It then follows that $\max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}SS_{22}=O_p(k_n^{-1})=o_p(1)$. Therefore, according to (A.10), we have

$$\max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}\left\|\boldsymbol{S}_n^{(2)}(\boldsymbol{\xi})-\boldsymbol{S}_n^{(2)}(\boldsymbol{\gamma}_l)\right\|=o_p(1).$$

With the similar argument, we can also show that

$$\max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}\left\|\boldsymbol{S}^{(2)}(\boldsymbol{\xi})-\boldsymbol{S}^{(2)}(\boldsymbol{\gamma}_l)\right\|=o(1).$$

It then follows that $P_1=o(1)$.

Let $\mathcal{X}_i(l,m)=(1-\delta_i)\int_y-\left(\partial_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_{1i},y;\boldsymbol{\theta})\right)_{(m)}\left(\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})/\widetilde{A}(\boldsymbol{\xi})\right)dy\cdot I(\boldsymbol{\xi}\in\Gamma_l)$, where $\left(\partial_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_{1i},y;\boldsymbol{\theta})\right)_{(m)}$ is the m-th element of $\partial_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_{1i},y;\boldsymbol{\theta})$. A sufficient condition for $P_2=o(1)$ is that, for any $\epsilon>0$,

$$Pr\left(\max_{1\le l\le L_n;1\le m\le d+1}\left|\frac{1}{n}\sum_{i=1}^n\{\mathcal{X}_i(l,m)-E\mathcal{X}_i(l,m)\}\right|>\epsilon\right)=o(1).$$

Since $|\mathcal{X}_i(l,m)| \le \frac{U_{\pi'}}{1-U_\pi}$ and $|\mathcal{X}_i(l,m) - E\mathcal{X}_i(l,m)| \le \frac{2U_{\pi'}}{1-U_\pi} \widehat{=} B$, by Bernstein's inequality we have

$$
\begin{aligned}
&Pr\left(\max_{1\le l\le L_n;1\le m\le d+1}\left|\frac{1}{n}\sum_{i=1}^{n}\{\mathcal{X}_i(l,m) - E\mathcal{X}_i(l,m)\}\right| > \epsilon\right)\\
\le&\sum_{l=1}^{L_n}\sum_{m=1}^{d+1} Pr\left(\left|\frac{1}{n}\sum_{i=1}^{n}\{\mathcal{X}_i(l,m) - E\mathcal{X}_i(l,m)\}\right| > \epsilon\right)\\
\le&2(d+1)L_n\exp\left(-\frac{n^2\epsilon^2}{2nB^2 + (2/3)Bn\epsilon}\right) = o(1).
\end{aligned}
$$

We now have shown that $P_1 = o(1)$ and $P_2 = o(1)$, which implies that the uniform convergence (A.9) holds. Hence, $k_n^{-1}\|\boldsymbol{S}_n(\boldsymbol{\xi}) - \boldsymbol{S}(\boldsymbol{\xi})\| = o_p(1)$ holds.

In what follows, we will show $\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} k_n^{-1}\|\boldsymbol{M}_n(\boldsymbol{\xi}) - \boldsymbol{M}(\boldsymbol{\xi})\| = o_p(1)$, which is equivalent to show that, for any $\epsilon > 0$,

$$
Pr\left(\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} k_n^{-1}\|\boldsymbol{M}_n(\boldsymbol{\xi}) - \boldsymbol{M}(\boldsymbol{\xi})\| > \epsilon\right) \to 0 \tag{A.11}
$$

as $n\to\infty$. The argument that proves (A.11) is similar to that of (A.9). The probability of the left side of (A.11) is bounded by the sum of the following two probabilities, $PP_1 + PP_2$, where

$$
PP_1 = Pr\left(\max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}\|\boldsymbol{M}_n(\boldsymbol{\xi}) - \boldsymbol{M}_n(\boldsymbol{\gamma}_l) - \boldsymbol{M}(\boldsymbol{\xi}) + \boldsymbol{M}(\boldsymbol{\gamma}_l)\| > \epsilon/2\right),
$$

and

$$
PP_2 = Pr\left(\max_{1\le l\le L_n} k_n^{-1}\|\boldsymbol{M}_n(\boldsymbol{\gamma}_l) - \boldsymbol{M}(\boldsymbol{\gamma}_l)\| > \epsilon/2\right).
$$

Firstly, note that

$$
\begin{aligned}
&\|\boldsymbol{M}_n(\boldsymbol{\xi}) - \boldsymbol{M}_n(\boldsymbol{\gamma}_l)\|\\
=&\left\|\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta})}\Psi\left(y_i - \boldsymbol{x}_i^\top\boldsymbol{\phi}\right)\otimes\boldsymbol{x}_i - \frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_l)}\Psi\left(y_i - \boldsymbol{x}_i^\top\boldsymbol{\phi}_l\right)\otimes\boldsymbol{x}_i\right\|\\
\le&\left\|\frac{1}{n}\sum_{i=1}^{n}\delta_i\left\{\frac{1}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta})} - \frac{1}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_l)}\right\}\Psi\left(y_i - \boldsymbol{x}_i^\top\boldsymbol{\phi}\right)\otimes\boldsymbol{x}_i\right\|\\
&+\left\|\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_l)}\left\{\Psi\left(y_i - \boldsymbol{x}_i^\top\boldsymbol{\phi}\right) - \Psi\left(y_i - \boldsymbol{x}_i^\top\boldsymbol{\phi}_l\right)\right\}\otimes\boldsymbol{x}_i\right\|\\
\widehat{=}&MM_1 + MM_2.
\end{aligned}
$$

Therefore,

$$
\max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}\|\boldsymbol{M}_n(\boldsymbol{\xi}) - \boldsymbol{M}_n(\boldsymbol{\gamma}_l)\| \le \max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}MM_1 + \max_{1\le l\le L_n}\sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1}MM_2. \tag{A.12}
$$

Morever, under Assumptions 1 and 3, there exist constants $K, U_{\pi'}$ such that $E_{\boldsymbol{x}_i}\|\boldsymbol{x}_i\| \le K$ and

$\|\partial_{\boldsymbol{\theta}} \pi(\boldsymbol{x}_{1i}, y_i; \boldsymbol{\theta})\| \le (d+1)U_{\pi'}$ for $\boldsymbol{x}_i \in \mathcal{X}, y_i \in \mathcal{Y}$. By Theorem 12.4 of Apostol (1974) and Assumption 3, there exists a number $\widetilde{\boldsymbol{\theta}} \in \Theta$ such that

$$
\begin{aligned}
&\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} MM_1 \\
\le &\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} \frac{1}{n} \sum_{k=1}^{k_n} \sum_{i=1}^{n} \delta_i \frac{|\pi(\boldsymbol{x}_{1i}, y_i; \boldsymbol{\theta}) - \pi(\boldsymbol{x}_{1i}, y_i; \boldsymbol{\theta}_l)|}{\pi(\boldsymbol{x}_{1i}, y_i; \boldsymbol{\theta})\, \pi(\boldsymbol{x}_{1i}, y_i; \boldsymbol{\theta}_l)} \|\Psi_{\tau_k}(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}_{\tau_k})\, \boldsymbol{x}_i\| \\
\le &\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} \frac{k_n}{L_\pi^2} \cdot \frac{1}{n} \sum_{i=1}^{n} \delta_i |\pi(\boldsymbol{x}_{1i}, y_i; \boldsymbol{\theta}) - \pi(\boldsymbol{x}_{1i}, y_i; \boldsymbol{\theta}_l)| \|\boldsymbol{x}_i\| \\
= &\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} \frac{k_n}{L_\pi^2} \cdot \frac{1}{n} \sum_{i=1}^{n} \delta_i \left| \left(\partial_{\boldsymbol{\theta}} \pi\left(\boldsymbol{x}_{1,i} y_i; \widetilde{\boldsymbol{\theta}}\right)\right)^\top (\boldsymbol{\theta} - \boldsymbol{\theta}_l) \right| \|\boldsymbol{x}_i\| \\
\le &\frac{(d+1)U_{\pi'}}{L_\pi^2} \cdot q_n k_n \cdot \frac{1}{n} \sum_{i=1}^{n} \|\boldsymbol{x}_i\| \\
= &O_p(k_n q_n).
\end{aligned}
$$

so $\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1} MM_1 = O(q_n) = o_p(1)$.

On the other hand, let $\boldsymbol{\phi}_l = \left(\boldsymbol{\beta}_{l,\tau_1}^\top, \boldsymbol{\beta}_{l,\tau_2}^\top, \ldots, \boldsymbol{\beta}_{l,\tau_{k_n}}^\top\right)^\top$, we have

$$
\begin{aligned}
&\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1} MM_2 \\
\le &\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} \frac{1}{k_n n} \sum_{k=1}^{k_n} \sum_{i=1}^{n} \frac{\delta_i}{\pi(\boldsymbol{x}_{1,i}, y_i; \boldsymbol{\theta}_l)} \|\{\Psi_{\tau_k}(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}_{\tau_k}) - \Psi_{\tau_k}(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}_{l,\tau_k})\}\, \boldsymbol{x}_i\| \\
\le &\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} \frac{1}{k_n n} \cdot \frac{1}{L_\pi} \sum_{k=1}^{k_n} \sum_{i=1}^{n} I\left\{|\boldsymbol{x}_i^\top \boldsymbol{\beta}_{l,\tau_k} - y_i| \le |\boldsymbol{x}_i^\top (\boldsymbol{\beta}_{\tau_k} - \boldsymbol{\beta}_{l,\tau_k})|\right\} \|\boldsymbol{x}_i\| \\
\le &\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} \frac{1}{k_n n} \cdot \frac{1}{L_\pi} \sum_{k=1}^{k_n} \sum_{i=1}^{n} I\left(|\boldsymbol{x}_i^\top \boldsymbol{\beta}_{l,\tau_k} - y_i| \le \|\boldsymbol{x}_i\|\, q_n\right) \|\boldsymbol{x}_i\|.
\end{aligned}
$$

Let $g_{l,k}(z_i)$ be the density of $z_i = (\boldsymbol{x}_i^\top \boldsymbol{\beta}_{l,\tau_k} - y_i)$ given $\boldsymbol{x}_i$ and $\boldsymbol{\beta}_{l,\tau_k}$. Then under Assumption 2, $g_{l,k}(z_i)$ is also continuous and bounded away from both zero and infinity for any $l \in \{1, 2, \ldots, L_n\}$, $k \in \{1, 2, \ldots, k_n\}$ and $i \in \{1, 2, \ldots, n\}$. Moreover, under Assumption 1, $E_{\boldsymbol{x}_i} \|\boldsymbol{x}_i\|^2$ is bounded.

Therefore, by the mean value theorem, for any i, there exists a value z_i^* such that

$$\begin{aligned}
& E\left\{I\left(\left|\boldsymbol{x}_i^\top \boldsymbol{\beta}_{l,\tau_k} - y_i\right| \le q_n \left\|\boldsymbol{x}_i\right\|\right) \left\|\boldsymbol{x}_i\right\|\right\} \\
=& E_{\boldsymbol{x}_i}\left[\left\|\boldsymbol{x}_i\right\| \cdot E_{y_i}\left\{I\left(\left|\boldsymbol{x}_i^\top \boldsymbol{\beta}_{l,\tau_k} - y_i\right| \le q_n \left\|\boldsymbol{x}_i\right\|\right) \mid \boldsymbol{x}_i; \boldsymbol{\beta}_{l,\tau_k}\right\}\right] \\
=& E_{\boldsymbol{x}_i}\left\{\left\|\boldsymbol{x}_i\right\| \cdot pr\left(\left|\boldsymbol{x}_i^\top \boldsymbol{\beta}_{l,\tau_k} - y_i\right| \le q_n \left\|\boldsymbol{x}_i\right\| \mid \boldsymbol{x}_i; \boldsymbol{\beta}_{l,\tau_k}\right)\right\} \\
=& 2q_n E\left\{\left\|\boldsymbol{x}_i\right\|^2 \cdot g_{l,k}(z_i^*)\right\} \\
=& O(q_n) = o(1).
\end{aligned}$$

Then by the law of large numbers, it follows that

$$\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1} MM_2 \le o_p(1).$$

According to (A.12), we have $\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1} \left\|\boldsymbol{M}_n(\boldsymbol{\xi}) - \boldsymbol{M}_n(\gamma_l)\right\| = o_p(1)$. With the similar argument, we can also show that, $\max_{1\le l\le L_n} \sup_{\boldsymbol{\xi}\in\Gamma_l} k_n^{-1} \left\|\boldsymbol{M}(\boldsymbol{\xi}) - \boldsymbol{M}(\gamma_l)\right\| = o(1)$. It then follows that $PP_1 = o(1)$.

Let $\mathcal{Y}_i(l,k,m) = \delta_i/\pi\left(\boldsymbol{x}_{1i}, y_i; \boldsymbol{\theta}_l\right) \cdot \Psi_{\tau_k}\left(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}_{\tau_k}\right) \boldsymbol{x}_{i(m)} \cdot I\left(\boldsymbol{\xi} \in \Gamma_l\right)$ where $\boldsymbol{x}_{i(m)}$ is the m-th element of $\boldsymbol{x}_i$. A sufficient condition for $PP_2 = o(1)$ is that, for any $\epsilon > 0$,

$$pr\left(\max_{1\le l\le L_n; 1\le k\le k_n; 1\le m\le p} \frac{1}{n}\left|\sum_{i=1}^n \left\{\mathcal{Y}_i(l,k,m) - E\mathcal{Y}_i(l,k,m)\right\}\right| > \epsilon\right) = o(1).$$

Under Assumption 1, for any $m \in \{1, 2, \ldots, p\}$, there exists a positive constant K_m such that $\max_{1\le i\le n} \left\|\boldsymbol{x}_{i(m)}\right\|_{\psi_1} \le K_m$. Under Assumption 3, there exists a positive constant L_π such that $|\mathcal{Y}_i(l,k,m)| \le |\boldsymbol{x}_{i(m)}|/L_\pi$. Let $K_0 = \max_{1\le m\le p} K_m$ and $K_{\mathcal{Y}}(l,k,m) = \max_{1\le i\le n} \|\mathcal{Y}_i(l,k,m) - E\mathcal{Y}_i(l,k,m)\|_{\psi_1}$. We then have

$$K_{\mathcal{Y}}(l,k,m) \le C_1 \max_{1\le i\le n} \left\|\mathcal{Y}_i(l,k,m)\right\|_{\psi_1} \le \frac{C_1}{L_\pi} \max_{1\le i\le n} \left\|\boldsymbol{x}_{i(m)}\right\|_{\psi_1} = \frac{C_1}{L_\pi} K_0,$$

where C_1 is a positive constant. By Bernstein's inequality, we obtain that

$$
\begin{aligned}
& pr\left(\max_{1\le l\le L_n;1\le k\le k_n;1\le m\le p}\frac{1}{n}\left|\sum_{i=1}^{n}\{\mathcal{Y}_i(l,k,m)-E\mathcal{Y}_i(l,k,m)\}\right|>\epsilon\right)\\
&\le \sum_{l=1}^{L_n}\sum_{k=1}^{k_n}\sum_{m=1}^{p} pr\left(\frac{1}{n}\left|\sum_{i=1}^{n}\{\mathcal{Y}_i(l,k,m)-E\mathcal{Y}_i(l,k,m)\}\right|>\epsilon\right)\\
&\le 2pL_nk_n\exp\left\{-\frac{C_2\epsilon^2 n}{K_{\mathcal{Y}}^2(l,k,m)}\right\}\\
&\le 2pL_nk_n\exp\left\{-\frac{C_2L_\pi^2\epsilon^2 n}{C_1^2K_0^2}\right\}\\
&= o(1).
\end{aligned}
$$

where C_2 is a positive constant. We now have shown that $PP_1 = o(1)$ and $PP_2 = o(1)$, which then implies the uniform convergence result (A.11). Lemma 2 is hence proved. □

LEMMA 3: *Under Assumption 3, for $k_n \to \infty, k_n/n \to 0$ and $m \to \infty$, we have*

$$
\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}}\left\|\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi})-\boldsymbol{H}_n(\boldsymbol{\xi})\right\| = o_p(1) \quad \textit{as } n\to\infty. \tag{A.13}
$$

Proof. Note that

$$
\begin{aligned}
\left\|\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi})-\boldsymbol{H}_n(\boldsymbol{\xi})\right\| &\le \left\|\widehat{\boldsymbol{S}}_n(\boldsymbol{\xi})-\widetilde{\boldsymbol{S}}_n(\boldsymbol{\xi})\right\| + \left\|\boldsymbol{M}_n(\boldsymbol{\xi})-\boldsymbol{M}_n(\boldsymbol{\xi})\right\|\\
&= \left\|\widehat{\boldsymbol{S}}_n(\boldsymbol{\xi})-\widetilde{\boldsymbol{S}}_n(\boldsymbol{\xi})\right\|,
\end{aligned}
$$

so to prove (A.13), it suffices to show that $\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}}\left\|\widehat{\boldsymbol{S}}_n(\boldsymbol{\xi})-\widetilde{\boldsymbol{S}}_n(\boldsymbol{\xi})\right\| = o_p(1)$.

For the subject i with $\delta_i = 0$, we let

$$
\begin{aligned}
\widehat{B}_{i,1}(\boldsymbol{\xi}) &= \frac{1}{m}\sum_{l=1}^{m}\left\{1-\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta}\right)\right\}s\left(\boldsymbol{\theta};\delta_i,\boldsymbol{x}_{1i},\widetilde{y}_i^l\right),\\
\widetilde{B}_{i,1}(\boldsymbol{\xi}) &= \int_y\left\{1-\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}\right)\right\}s\left(\boldsymbol{\theta};\delta_i,\boldsymbol{x}_{1i},y\right)\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})dy,\\
\widehat{B}_{i,2}(\boldsymbol{\xi}) &= \frac{1}{m}\sum_{l=1}^{m}\left\{1-\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta}\right)\right\},\\
\widetilde{B}_{i,2}(\boldsymbol{\xi}) &= \int_y\left\{1-\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}\right)\right\}\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})dy,
\end{aligned}
$$

where $\{\widetilde{y}_i^l\}_{l=1}^m$ are sampled from $\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})$. Under Assumption 3, we know that there exist positive constants $U_{\pi'}$, L_π and U_π such that $\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}}\left\|\widetilde{B}_{i,1}(\boldsymbol{\xi})\right\| \le (d+1)U_{\pi'}$ and $1-L_\pi > \widetilde{B}_{i,2}(\boldsymbol{\xi}) \ge 1-U_\pi$.

Then by Lemma 2.4 of Engle (1994), as $m \to \infty$,

$$\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} \left\| \widehat{B}_{i,1}(\boldsymbol{\xi}) - \widetilde{B}_{i,1}(\boldsymbol{\xi}) \right\| = o_p(1),$$

$$\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} \left\| \widehat{B}_{i,2}(\boldsymbol{\xi}) - \widetilde{B}_{i,2}(\boldsymbol{\xi}) \right\| = o_p(1).$$

Consequently,

$$\begin{aligned}
&\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} \|\widehat{\boldsymbol{S}}_n(\boldsymbol{\xi}) - \widetilde{\boldsymbol{S}}_n(\boldsymbol{\xi})\| \\
=&\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} \left\| \frac{1}{n}\sum_{i=1}^{n}(1-\delta_i)\left[\frac{\sum_{l=1}^{m}\left\{1-\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta}\right)\right\} s\left(\boldsymbol{\theta};\delta_i,\boldsymbol{x}_{1i},\widetilde{y}_i^l\right)}{\sum_{l=1}^{m}\left\{1-\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta}\right)\right\}} - \frac{\int_y \left\{1-\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}\right)\right\} s\left(\boldsymbol{\theta};\delta_i,\boldsymbol{x}_{1i},y\right)\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})dy}{\int_y \left\{1-\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}\right)\right\}\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})dy} \right] \right\| \\
\le&\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} \left\| \frac{1}{n}\sum_{i=1}^{n}(1-\delta_i)\frac{\widehat{B}_{i,1}(\boldsymbol{\xi}) - \widetilde{B}_{i,1}(\boldsymbol{\xi})}{\widehat{B}_{i,2}(\boldsymbol{\xi})} \right\| + \sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} \left\| \frac{1}{n}\sum_{i=1}^{n}(1-\delta_i)\frac{\widetilde{B}_{i,1}(\boldsymbol{\xi})}{\widehat{B}_{i,2}(\boldsymbol{\xi})\widetilde{B}_{i,2}(\boldsymbol{\xi})}\left(\widehat{B}_{i,2}(\boldsymbol{\xi}) - \widetilde{B}_{i,2}(\boldsymbol{\xi})\right) \right\| \\
\le&\frac{1}{1-U_\pi}\cdot\frac{1}{n}\sum_{i=1}^{n}\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}}\|\widehat{B}_{i,1}(\boldsymbol{\xi}) - \widetilde{B}_{i,1}(\boldsymbol{\xi})\| + \frac{(d+1)U_\pi^{'}}{(1-U_\pi)^2}\cdot\frac{1}{n}\sum_{i=1}^{n}\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}}|\widehat{B}_{i,2}(\boldsymbol{\xi}) - \widetilde{B}_{i,2}(\boldsymbol{\xi})| \\
=&o_p(1).
\end{aligned}$$

Thus, we have $\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} \left\|\widehat{\boldsymbol{S}}_n(\boldsymbol{\xi}) - \widetilde{\boldsymbol{S}}_n(\boldsymbol{\xi})\right\| = o_p(1)$ and (A.13) holds. □

Proof of Theorem 1

Proof. Recall that $\widehat{\boldsymbol{\xi}}_n = \left(\widehat{\boldsymbol{\theta}}_n^\top, \widehat{\boldsymbol{\phi}}_n^\top\right)^\top$ is the solution of the estimating equation $\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}) = 0$ with $\widehat{\boldsymbol{\phi}}_n = \left(\widehat{\boldsymbol{\beta}}_{\tau_1}^\top, \ldots, \widehat{\boldsymbol{\beta}}_{\tau_{k_n}}^\top\right)^\top$. We further define $\widehat{\boldsymbol{\beta}}_j(\tau)$ as the piece-wise linear function with $\widehat{\boldsymbol{\beta}}_j(\tau_k) = \widehat{\beta}_{j,\tau_k}$ and $\widehat{\beta}_{j,\tau_k}$ is the jth component of $\widehat{\boldsymbol{\beta}}_{\tau_k}$. Let $\boldsymbol{\xi}_\tau = (\boldsymbol{\theta}^\top, \boldsymbol{\beta}_\tau^\top)^\top$ and the estimate $\widehat{\boldsymbol{\xi}}_\tau = \left(\widehat{\boldsymbol{\theta}}_n^\top, \widehat{\boldsymbol{\beta}}_\tau^\top\right)^\top$ and its corresponding true value $\boldsymbol{\xi}_{0,\tau} = \left(\boldsymbol{\theta}_0^\top, \boldsymbol{\beta}_{0,\tau}^\top\right)^\top$, where $\boldsymbol{\beta}_{0,\tau}$ is the true coefficients at the quantile level τ.

For any $\gamma > 0$, we define a compact set $\mathcal{B}_\tau = \left\{\boldsymbol{\xi}_\tau \in \boldsymbol{R}^{p+d+1} : \|\boldsymbol{\xi}_\tau - \boldsymbol{\xi}_{0,\tau}\| < \gamma\right\}$ and $\mathcal{B}_\tau^c$ as its complementary set. We denote $\mathcal{B}_\tau \otimes \Omega = \{\boldsymbol{\xi} : \boldsymbol{\xi}_{\tau_k} \in \mathcal{B}_{\tau_k} \text{ for every } k = 1, \ldots, k_n\}$. We consider the distance

$$d_n(\gamma) = k_n^{-1}\left\{\min_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}\cap(\mathcal{B}_\tau^c\otimes\Omega)} \|\boldsymbol{H}(\boldsymbol{\xi})\| - \|\boldsymbol{H}(\boldsymbol{\xi}_0)\|\right\} \tag{A.14}$$

between the norm of the limiting working estimating equation $\boldsymbol{H}(\cdot)$ evaluated at the true coefficients $\boldsymbol{\xi}_0$ and the minimized norm when $\boldsymbol{\xi}$ stays in $\Theta_{\boldsymbol{\xi}} \cap (\mathcal{B}_\tau^c \otimes \Omega)$. In what follows, we show that $d_n(\gamma) > 0$ under Assumptions 1-5.

Recall in Assumption 4 that $\boldsymbol{\xi}_0$ is the unique solution of $\boldsymbol{H}^0(\boldsymbol{\xi}) = 0$. Therefore, the convergence of (A.1) stated in Lemma 1 is equivalent to $k_n^{-1}\left\|\boldsymbol{H}\left(\boldsymbol{\xi}_0\right)\right\| = o(1)$. Moreover, since $\boldsymbol{\xi}^\star = (\boldsymbol{\theta}^{\star\top}, \boldsymbol{\phi}^{\star\top})^\top$ is the unique solution of $\boldsymbol{H}(\boldsymbol{\xi}) = 0$ under Assumption 4, it follows that $k_n^{-1}\left\|\boldsymbol{H}\left(\boldsymbol{\xi}_0\right) - \boldsymbol{H}\left(\boldsymbol{\xi}^\star\right)\right\| = o(1)$. Due to the continuity of $\boldsymbol{H}(\cdot)$ and the uniqueness of $\boldsymbol{\xi}^\star$, we have $k_n^{-1}\left\|\boldsymbol{\xi}^\star - \boldsymbol{\xi}_0\right\| \to 0$ as $n \to \infty$. Consequently, there exists a constant K_γ, such that when $k_n > K_\gamma$, we have that $k_n^{-1}\left\|\boldsymbol{\xi}^\star - \boldsymbol{\xi}_0\right\| < \gamma/2$, so $\boldsymbol{\xi}^\star \in \Theta_{\boldsymbol{\xi}} \cap (\mathcal{B}_\tau \otimes \Omega)$ for $k_n > K_\gamma$. Due to the uniqueness of $\boldsymbol{\xi}^\star$, for any $k_n > K_\gamma$, we have

$$d_n^\star(\gamma) = k_n^{-1}\left\{\min_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}\cap(\mathcal{B}_\tau^c\otimes\Omega)}\|\boldsymbol{H}(\boldsymbol{\xi})\| - \|\boldsymbol{H}(\boldsymbol{\xi}^\star)\|\right\} > 0. \tag{A.15}$$

On the other hand, due to the continuity of $\boldsymbol{H}(\cdot)$, we also have

$$k_n^{-1}\left\|\boldsymbol{H}\left(\boldsymbol{\xi}_0\right) - \boldsymbol{H}\left(\boldsymbol{\xi}^\star\right)\right\| \le d_n^\star(\gamma)/2, \tag{A.16}$$

for sufficiently larger k_n. Combining (A.15) and (A.16), it follows that

$$d_n(\gamma) = k_n^{-1}\left\{\min_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}\cap(\mathcal{B}_\tau^c\otimes\Omega)}\|\boldsymbol{H}(\boldsymbol{\xi})\| - \|\boldsymbol{H}(\boldsymbol{\xi}_0)\|\right\} > d_n^\star(\gamma)/2 > 0 \tag{A.17}$$

for sufficiently large k_n.

We now consider an event

$$E_n = \left\{k_n^{-1}\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}}\left\|\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}) - \boldsymbol{H}(\boldsymbol{\xi})\right\| \le \frac{d_n(\gamma)}{3}\right\},$$

which implies that

$$k_n^{-1}\left\|\boldsymbol{H}\left(\widehat{\boldsymbol{\xi}}_n\right)\right\| \le k_n^{-1}\left\|\widehat{\boldsymbol{H}}_n\left(\widehat{\boldsymbol{\xi}}_n\right)\right\| + \frac{d_n(\gamma)}{3}, \tag{A.18}$$

$$k_n^{-1}\left\|\widehat{\boldsymbol{H}}_n\left(\boldsymbol{\xi}_0\right)\right\| \le k_n^{-1}\left\|\boldsymbol{H}\left(\boldsymbol{\xi}_0\right)\right\| + \frac{d_n(\gamma)}{3}. \tag{A.19}$$

Since $\widehat{\boldsymbol{\xi}}_n$ is the minimizer of $\left\|\widehat{\boldsymbol{H}}_n\left(\boldsymbol{\xi}\right)\right\|$, and thus $\left\|\widehat{\boldsymbol{H}}_n\left(\boldsymbol{\xi}_0\right)\right\| \ge \left\|\widehat{\boldsymbol{H}}_n\left(\widehat{\boldsymbol{\xi}}_n\right)\right\|$, which, together with (A.18) and (A.19), indicates that

$$k_n^{-1}\left\|\boldsymbol{H}\left(\widehat{\boldsymbol{\xi}}_n\right)\right\| \le k_n^{-1}\left\|\widehat{\boldsymbol{H}}_n\left(\boldsymbol{\xi}_0\right)\right\| + \frac{d_n(\gamma)}{3} \le k_n^{-1}\left\|\boldsymbol{H}\left(\boldsymbol{\xi}_0\right)\right\| + \frac{2d_n(\gamma)}{3}.$$

Combining Lemmas 2 and 3, we have $\sup_{\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}} k_n^{-1}\left\|\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}) - \boldsymbol{H}(\boldsymbol{\xi})\right\| = o_p(1)$ and thus $\lim_{n\to\infty} pr\left(E_n\right) = 1$, which implies

$$\lim_{n\to\infty} pr\left(k_n^{-1}\left\|\boldsymbol{H}\left(\widehat{\boldsymbol{\xi}}_n\right)\right\| \le k_n^{-1}\left\|\boldsymbol{H}\left(\boldsymbol{\xi}_0\right)\right\| + \frac{2d_n(\gamma)}{3}\right) \ge \lim_{n\to\infty} pr\left(E_n\right) = 1$$

By the definition of $\mathcal{B}_\tau$ and the fact that $d_n(\gamma) > 0$, this in turn implies that

$$\lim_{n\to\infty} pr\left(\widehat{\boldsymbol{\xi}}_n \in \Theta_{\boldsymbol{\xi}} \cap (\mathcal{B}_\tau \otimes \Omega)\right) = 1,$$

that is,

$$\left\|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\right\| = o_p(1),$$

$$\sup_{\tau\in[1/(k_n+1),k_n/(k_n+1)]} \left\|\widehat{\boldsymbol{\beta}}_n(\tau) - \boldsymbol{\beta}_0(\tau)\right\| = o_p(1)$$

The consistency of $\widehat{\boldsymbol{\theta}}_n$ and $\widehat{\boldsymbol{\beta}}_n(\tau)$ is then established. □

Web Appendix C

Denote

$$\boldsymbol{H}_n^0(\boldsymbol{\xi}) = (\boldsymbol{S}_n^\top(\boldsymbol{\theta}), \boldsymbol{M}_n^\top(\boldsymbol{\xi}))^\top = \frac{1}{n}\sum_{i=1}^n h_i(\boldsymbol{\xi}) = \left(\frac{1}{n}\sum_{i=1}^n h_{1i}^\top(\boldsymbol{\xi}), \frac{1}{n}\sum_{i=1}^n h_{2i}^\top(\boldsymbol{\xi})\right)^\top,$$

and

$$\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}) = \left(\widehat{\boldsymbol{S}}_n^\top(\boldsymbol{\xi}), \boldsymbol{M}_n^\top(\boldsymbol{\xi})\right)^\top = \frac{1}{n}\sum_{i=1}^n \widehat{h}_i(\boldsymbol{\xi}) = \left(\frac{1}{n}\sum_{i=1}^n \widehat{h}_{1i}^\top(\boldsymbol{\xi}), \frac{1}{n}\sum_{i=1}^n \widehat{h}_{2i}^\top(\boldsymbol{\xi})\right)^\top,$$

where

$$h_{1i}(\boldsymbol{\xi}) = \delta_i \boldsymbol{s}(\boldsymbol{\theta}; \delta_i, \boldsymbol{x}_{1i}, y_i) + (1-\delta_i)\frac{\int\{1-\pi(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta})\} s(\boldsymbol{\theta}; \delta_i, \boldsymbol{x}_{1i}, y) f(y|\boldsymbol{x}_i; \boldsymbol{\beta}(\boldsymbol{\tau}))dy}{\int\{1-\pi(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta})\} f(y|\boldsymbol{x}_i; \boldsymbol{\beta}(\boldsymbol{\tau}))dy},$$

$$\widehat{h}_{1i}(\boldsymbol{\xi}) = \delta_i \boldsymbol{s}(\boldsymbol{\theta}; \delta_i, \boldsymbol{x}_{1i}, y_i) + (1-\delta_i)\frac{\sum_{l=1}^m\{1-\pi(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta})\} s(\boldsymbol{\theta}; \delta_i, \boldsymbol{x}_{1i}, \widetilde{y}_i^l)}{\sum_{l=1}^m\{1-\pi(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta})\}},$$

and

$$h_{2i}(\boldsymbol{\xi}) = \widehat{h}_{2i}(\boldsymbol{\xi}) = \frac{\delta_i}{\pi(\boldsymbol{x}_{1i}, y_i; \boldsymbol{\theta})}\boldsymbol{\Psi}(y_i - \boldsymbol{x}_i^\top \boldsymbol{\phi}) \otimes \boldsymbol{x}_i.$$

Let $h_{i(j)}(\boldsymbol{\xi})$, $h_{1i(j)}(\boldsymbol{\xi})$ and $h_{2i(j)}(\boldsymbol{\xi})$ be the jth element of $h_i(\boldsymbol{\xi})$, $h_{1i}(\boldsymbol{\xi})$ and $h_{2i}(\boldsymbol{\xi})$, respectively, and let $\widehat{h}_{i(j)}(\boldsymbol{\xi})$, $\widehat{h}_{1i(j)}(\boldsymbol{\xi})$ and $\widehat{h}_{2i(j)}(\boldsymbol{\xi})$ be the jth element of $\widehat{h}_i(\boldsymbol{\xi})$, $\widehat{h}_{1i}(\boldsymbol{\xi})$ and $\widehat{h}_{2i}(\boldsymbol{\xi})$, respectively. Moreover, let

$$\eta_i(\boldsymbol{\xi}, \boldsymbol{\xi}_0) = \widehat{h}_i(\boldsymbol{\xi}) - \widehat{h}_i(\boldsymbol{\xi}_0) - E\widehat{h}_i(\boldsymbol{\xi}) + E\widehat{h}_i(\boldsymbol{\xi}_0). \tag{A.20}$$

LEMMA 4: *Under Assumptions 1-4 and 6-7, for $n\to\infty, k_n^{3+2v}/n \to \infty$ and $k_n m/n \to \infty$, we have*

$$\sqrt{\frac{n}{q}}\left\|\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}_0) - \boldsymbol{H}_n^0(\boldsymbol{\xi}_0)\right\| = o_p(1). \tag{A.21}$$

Proof. For convenience, we assume that $\{y_i : i = 1, 2, \ldots, n_0\}$ are observed and $\{y_i : i = n_0 + 1, n_0 + 2, \ldots, n\}$ are missing. Then, we have

$$\begin{aligned}
&\sqrt{\frac{n}{q}}\left\|\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}_0) - \boldsymbol{H}_n^0(\boldsymbol{\xi}_0)\right\| \\
=&\sqrt{\frac{n}{q}}\left\|\widehat{S}_n(\boldsymbol{\xi}_0) - S_n(\boldsymbol{\theta}_0)\right\| \\
\leq&\sqrt{\frac{n}{q}}\left\|\frac{1}{n}\sum_{i=n_0+1}^{n}\left[\frac{\frac{1}{m}\sum_{l=1}^{m}\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta}_0\right)}{\frac{1}{m}\sum_{l=1}^{m}\left\{1 - \pi\left(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta}_0\right)\right\}} - \frac{\int_y \partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\widetilde{f}\left(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0\right)dy}{\int_y\left\{1 - \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\right\}\widetilde{f}\left(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0\right)dy}\right]\right\| \\
&+\sqrt{\frac{n}{q}}\left\|\frac{1}{n}\sum_{i=n_0+1}^{n}\left[\frac{\int_y \partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\widetilde{f}\left(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0\right)dy}{\int_y\left\{1 - \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\right\}\widetilde{f}\left(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0\right)dy}\right.\right. \\
&\left.\left.-\frac{\int_y \partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)f\left(y|\boldsymbol{x}_i; \boldsymbol{\beta}_0(\tau)\right)dy}{\int_y\left\{1 - \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\right\}f\left(y|\boldsymbol{x}_i; \boldsymbol{\beta}_0(\tau)\right)dy}\right]\right\| \\
\widehat{=}&\sqrt{\frac{n}{q}}\left\|\frac{1}{n}\sum_{i=n_0+1}^{n}L_{1i}(\boldsymbol{\xi}_0)\right\| + \sqrt{\frac{n}{q}}\left\|\frac{1}{n}\sum_{i=n_0+1}^{n}L_{2i}(\boldsymbol{\xi}_0)\right\|
\end{aligned}$$

where $\{\widetilde{y}_i^l\}_{l=1}^m$ are randomly drawn from $\widetilde{f}(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0)$, and given $\boldsymbol{x}_i$, they are independent. By Assumption 3, $\|\partial_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_1, y; \boldsymbol{\theta}_0)\|^2$ is bounded. So by the central limit theorem, it is straightforward to show that

$$\left|\frac{1}{m}\sum_{l=1}^{m}\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta}_0\right) - \int_y \partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\widetilde{f}\left(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0\right)dy\right| = O_p\left(\frac{1}{\sqrt{m}}\right).$$

Moreover, by Assumption 3, there exist constants $0 < L_\pi \leq U_\pi < 1$ such that $1-U_\pi \leq |1 - \pi(\boldsymbol{x}_1, y; \boldsymbol{\theta}_0)| \leq 1-L_\pi$. Then, it is easy to show $\|L_{1i}(\boldsymbol{\xi}_0)\| = O_p\left(\frac{1}{\sqrt{m}}\right)$ and thus $\sqrt{\frac{n}{q}}\left\|\frac{1}{n}\sum_{i=n_0+1}^{n}L_{1i}(\boldsymbol{\xi}_0)\right\| = O_p\left(\sqrt{\frac{n}{qm}}\right) = O_p\left(\sqrt{\frac{n}{k_n m}}\right) = o_p(1)$. Therefore, it suffices to show that

$$\sqrt{\frac{n}{q}}\left\|\frac{1}{n}\sum_{i=n_0+1}^{n}L_{2i}(\boldsymbol{\xi}_0)\right\| = o_p(1).$$

Note that

$$\begin{aligned}
&\|L_{2i}(\boldsymbol{\xi}_0)\| \\
\leq&\left\|\frac{\int_y \partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\left\{\widetilde{f}\left(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0\right) - f\left(y|\boldsymbol{x}_i; \boldsymbol{\beta}_0(\tau)\right)\right\}dy}{\int_y\left\{1 - \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\right\}\widetilde{f}\left(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0\right)dy}\right\| \\
&+\left\|\frac{\int_y \partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)f\left(y|\boldsymbol{x}_i; \boldsymbol{\beta}_0(\tau)\right)dy}{\int_y\left\{1 - \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\right\}\widetilde{f}\left(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0\right)dy \times \int_y\left\{1 - \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\right\}f\left(y|\boldsymbol{x}_i; \boldsymbol{\beta}_0(\tau)\right)dy}\right\| \\
&\times\left|\int_y\left\{1 - \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\right\}\left[\widetilde{f}\left(y|\boldsymbol{x}_i; \boldsymbol{\phi}_0\right) - f\left(y|\boldsymbol{x}_i; \boldsymbol{\beta}_0(\tau)\right)\right]dy\right|
\end{aligned}$$

$$\le \left(\frac{(d+1)U_{\pi'}}{(\tau_{k_n} - \tau_{k_1})(1-U_\pi)} + \frac{(d+1)U_{\pi'}(1-L_\pi)}{(\tau_{k_n} - \tau_{k_1})(1-U_\pi)^2} \right) \int_y \left| \widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0) - f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau)) \right| dy$$

$$\widehat{=} C^* \int_y \left| \widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0) - f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau)) \right| dy. \tag{A.22}$$

According to (A.4), (A.5) and Assumption 7, we have that $\left| \widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0) - f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau)) \right| = O_p(k_n^{-(1+v)})$. Then with the similar argument as that of proving (A.6) in Lemma 1, we have

$$\int_y \left| \widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0) - f(y|\boldsymbol{x}_i;\boldsymbol{\beta}_0(\tau)) \right| dy = O_p\left(k_n^{-(v+1)}\right). \tag{A.23}$$

Combining (A.22) and (A.23), $\|L_{2i}(\boldsymbol{\xi}_0)\| = O_p(k_n^{-(v+1)})$ and then $\sqrt{n/q}\left\|n^{-1}\sum_{i=n_0+1}^n L_{2i}(\boldsymbol{\xi}_0)\right\| = O_p(\sqrt{n/k_n^{3+2v}}) = o_p(1)$. Hence, the conclusion of Lemma 4 is accomplished. □

LEMMA 5: *Under Assumptions 1-8, for $n \to \infty$, $k_n^{3+2v}/n \to \infty$ and $k_n m/n \to \infty$, we have*

$$\left\|\widehat{\boldsymbol{\xi}}_n - \boldsymbol{\xi}_0\right\| = O_p\left(\sqrt{\frac{q}{n}}\right) \tag{A.24}$$

Proof. Following the similar arguments as Lemma 4 and using the dominated convergence theorem, we obtain that

$$\left\|\boldsymbol{H}(\boldsymbol{\xi}_0) - \boldsymbol{H}^0(\boldsymbol{\xi}_0)\right\| = o\left(\sqrt{\frac{q}{n}}\right)$$

as $n \to \infty$ and $k_n^{3+2v}/n \to \infty$. Since $\boldsymbol{H}^0(\boldsymbol{\xi}_0) = 0$ holds, we then have $\|\boldsymbol{H}(\boldsymbol{\xi}_0)\| = o(1)$.

Moreover, note that $\boldsymbol{\xi}^\star = (\boldsymbol{\theta}^{\star\top}, \boldsymbol{\phi}^{\star\top})^\top$ is the unique solution of $\boldsymbol{H}(\boldsymbol{\xi}) = 0$ under Assumption 4, so it follows that $\|\boldsymbol{H}(\boldsymbol{\xi}^\star) - \boldsymbol{H}(\boldsymbol{\xi}_0)\| = o(1)$. Due to the continuity of $\boldsymbol{H}(\cdot)$ and the uniqueness of $\boldsymbol{\xi}^\star$, we have $\|\boldsymbol{\xi}^\star - \boldsymbol{\xi}_0\| = o(1)$

Since $\boldsymbol{H}^0(\boldsymbol{\xi}_0) = \boldsymbol{H}(\boldsymbol{\xi}^\star) = 0$ and $\sqrt{n/q}\left\|\boldsymbol{H}(\boldsymbol{\xi}_0) - \boldsymbol{H}^0(\boldsymbol{\xi}_0)\right\| = o(1)$, we have

$$\sqrt{\frac{n}{q}}\|\boldsymbol{H}(\boldsymbol{\xi}^\star) - \boldsymbol{H}(\boldsymbol{\xi}_0)\| = \sqrt{\frac{n}{q}}\left\|\boldsymbol{H}^0(\boldsymbol{\xi}_0) - \boldsymbol{H}(\boldsymbol{\xi}_0)\right\| = o(1). \tag{A.25}$$

As $n \to \infty$, by Assumption 8, we have

$$\begin{aligned}
\frac{n}{q}\|\boldsymbol{H}(\boldsymbol{\xi}^\star) - \boldsymbol{H}(\boldsymbol{\xi}_0)\|^2 &= \frac{n}{q}\|D_n(\boldsymbol{\xi}^\star - \boldsymbol{\xi}_0) + r_n\|^2 \\
&\ge \frac{1}{4}\frac{n}{q} \cdot \left[(\boldsymbol{\xi}^\star - \boldsymbol{\xi}_0)^\top D^\top D(\boldsymbol{\xi}^\star - \boldsymbol{\xi}_0)\right] \\
&\ge \frac{\lambda_{min}(D_n^\top D_n)}{4} \cdot \frac{n}{q}\|\boldsymbol{\xi}^\star - \boldsymbol{\xi}_0\|^2
\end{aligned} \tag{A.26}$$

where r_n is a q-dimensional vector with $\|r_n\| = o(\|\boldsymbol{\xi}^\star - \boldsymbol{\xi}_0\|)$ since $\|\boldsymbol{\xi}^\star - \boldsymbol{\xi}_0\| = o(1)$. Combining (A.25) and (A.26), we have $\|\boldsymbol{\xi}^\star - \boldsymbol{\xi}_0\| = o\left(\sqrt{\frac{q}{n}}\right)$. Since $\left\|\widehat{\boldsymbol{\xi}}_n - \boldsymbol{\xi}_0\right\| \le \left\|\widehat{\boldsymbol{\xi}}_n - \boldsymbol{\xi}^\star\right\| + \|\boldsymbol{\xi}^\star - \boldsymbol{\xi}_0\|$, it then suffices to show that $\left\|\widehat{\boldsymbol{\xi}}_n - \boldsymbol{\xi}^\star\right\| = O_p\left(\sqrt{\frac{q}{n}}\right)$.

Under Assumption 3, there exist constants $0 < L_\pi \le U_\pi < 1$ such that $L_\pi \le |\pi(\boldsymbol{x}_1, y; \boldsymbol{\theta})| \le U_\pi$ and a constant $0 < U_{\pi'} < \infty$ such that for the jth element of the gradient $\partial_{\boldsymbol{\theta}}\pi(\boldsymbol{\theta}; \boldsymbol{x}_1, y)$ $(j = 1, \dots, d+1)$, denoted as $(\partial_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_1, y; \boldsymbol{\theta}))_{(j)}$, satisfies $\left|(\partial_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_1, y; \boldsymbol{\theta}))_{(j)}\right| \le U_{\pi'}$. Therefore, for $i = 1, \dots, n$, the jth element of $\widehat{h}_{1i}(\boldsymbol{\xi})$ satisfies

$$
\begin{aligned}
\left|\widehat{h}_{1i(j)}(\boldsymbol{\xi})\right| &\le \left|\frac{(\partial_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_{1i}, y_i; \boldsymbol{\theta}))_{(j)}}{\pi(\boldsymbol{x}_{1i}, y_i; \boldsymbol{\theta})}\right| + \left|\frac{\frac{1}{m}\sum_{l=1}^m (\partial_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta}))_{(j)}}{\frac{1}{m}\sum_{l=1}^m \{1 - \pi(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta})\}}\right| \\
&\le \frac{U_{\pi'}}{L_\pi} + \frac{U_{\pi'}}{1 - U_\pi},
\end{aligned}
$$

and thus $E|\widehat{h}_{1i(j)}(\boldsymbol{\xi})|^2 < \infty$.

Furthermore, by Assumptions 1 and 3, we have $E|\widehat{h}_{2i(j)}(\boldsymbol{\xi})|^2 < Ex_{i(j)}^2 / L_\pi^2 < \infty$. It then follows from the central limit theorem that $\frac{1}{\sqrt{n}}\sum_{i=1}^n \widehat{h}_{i(j)}(\boldsymbol{\xi}) = \frac{1}{\sqrt{n}}\sum_{i=1}^n E\widehat{h}_{i(j)}(\boldsymbol{\xi}) + O_p(1)$ as $n \to \infty$, for all $\boldsymbol{\xi} \in \Theta_{\boldsymbol{\xi}}$, $j \in \{1, \dots, q\}$. Thus, we have

$$
\sup_{\boldsymbol{\xi} \in \Theta_{\boldsymbol{\xi}}} \left|\frac{1}{\sqrt{n}}\sum_{i=1}^n \left(\widehat{h}_{i(j)}(\boldsymbol{\xi}) - E\widehat{h}_{i(j)}(\boldsymbol{\xi})\right)\right|^2 = O_p(1),
$$

as $n \to \infty$. Following the similar argument as Lemma 3, we obtain

$$
\sup_{\boldsymbol{\xi} \in \Theta_{\boldsymbol{\xi}}} \left\|\widetilde{\boldsymbol{H}}(\boldsymbol{\xi}) - \boldsymbol{H}(\boldsymbol{\xi})\right\| = O\Big(\sqrt{\frac{1}{m}}\Big). \tag{A.27}
$$

Thus, we have

$$
\begin{aligned}
&\sup_{\boldsymbol{\xi} \in \Theta_{\boldsymbol{\xi}}} \sqrt{\frac{n}{q}} \left\|\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}) - \boldsymbol{H}(\boldsymbol{\xi})\right\| \\
&= \sup_{\boldsymbol{\xi} \in \Theta_{\boldsymbol{\xi}}} \sqrt{\frac{n}{q}} \left\|\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}) - \widetilde{\boldsymbol{H}}(\boldsymbol{\xi})\right\| + \sup_{\boldsymbol{\xi} \in \Theta_{\boldsymbol{\xi}}} \sqrt{\frac{n}{q}} \left\|\widetilde{\boldsymbol{H}}(\boldsymbol{\xi}) - \boldsymbol{H}(\boldsymbol{\xi})\right\| \\
&= \sup_{\boldsymbol{\xi} \in \Theta_{\boldsymbol{\xi}}} \sqrt{\frac{1}{q}\sum_{j=1}^q \left|\frac{1}{\sqrt{n}}\sum_{i=1}^n \left\{\widehat{h}_{i(j)}(\boldsymbol{\xi}) - E\widehat{h}_{i(j)}(\boldsymbol{\xi})\right\}\right|^2} + O\Big(\sqrt{\frac{n}{qm}}\Big) \\
&= O_p(1).
\end{aligned}
$$

By Assumptions 4 and 5 and the construction of the proposed estimator $\widehat{\boldsymbol{\xi}}_n$, we obtain that $\boldsymbol{H}(\boldsymbol{\xi}^\star) =$

$0, \sqrt{\frac{n}{q}}\left\|\widehat{\boldsymbol{H}}_n\left(\widehat{\boldsymbol{\xi}}_n\right)\right\| = o_p(1)$. Therefore, we have

$$\begin{aligned}\sqrt{\frac{n}{q}}\left\|\boldsymbol{H}\left(\widehat{\boldsymbol{\xi}}_n\right)-\boldsymbol{H}\left(\boldsymbol{\xi}^\star\right)\right\| \le& \sqrt{\frac{n}{q}}\left\|\boldsymbol{H}\left(\widehat{\boldsymbol{\xi}}_n\right)-\widehat{\boldsymbol{H}}_n\left(\widehat{\boldsymbol{\xi}}_n\right)\right\|+\sqrt{\frac{n}{q}}\left\|\widehat{\boldsymbol{H}}_n\left(\widehat{\boldsymbol{\xi}}_n\right)\right\| \\ =& O_p(1).\end{aligned} \tag{A.28}$$

Following the similar argument as that of proving (A.26) under Assumption 8, we obtain

$$\frac{n}{q}\left\|\boldsymbol{H}\left(\widehat{\boldsymbol{\xi}}_n\right)-\boldsymbol{H}\left(\boldsymbol{\xi}^\star\right)\right\|^2 \ge \frac{\lambda_{\min}(D_n^\top D_n)}{4}\cdot\frac{n}{q}\left\|\widehat{\boldsymbol{\xi}}_n-\boldsymbol{\xi}^\star\right\|^2. \tag{A.29}$$

Combining (A.28) and (A.29), we have $\left\|\widehat{\boldsymbol{\xi}}_n-\boldsymbol{\xi}^\star\right\| = O_p(\sqrt{\frac{q}{n}})$. The proof of Lemma 5 is hence completed. □

LEMMA 6: *Under Assumptions 1-8, as $n\to\infty, k_n^2/n\to 0, k_n^{3+2v}/n\to\infty$ and $k_n m/n\to\infty$, for any given $B>0$, we have*

$$E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\eta_i\left(\boldsymbol{\xi},\boldsymbol{\xi}_0\right)\right\|^2\right\} \le O\left(\frac{q}{\sqrt{n}}\right), \tag{A.30}$$

for $i=1,\dots,n$.

Proof. Note that

$$\begin{aligned}&E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\eta_i\left(\boldsymbol{\xi},\boldsymbol{\xi}_0\right)\right\|^2\right\}\\ =&E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\widehat{h}_i\left(\boldsymbol{\xi}\right)-\widehat{h}_i\left(\boldsymbol{\xi}_0\right)-E\widehat{h}_i\left(\boldsymbol{\xi}\right)+E\widehat{h}_i\left(\boldsymbol{\xi}_0\right)\right\|^2\right\}\\ \le&4E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\widehat{h}_i\left(\boldsymbol{\xi}\right)-\widehat{h}_i\left(\boldsymbol{\xi}_0\right)\right\|^2\right\}\\ =&4E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\widehat{h}_{1i}\left(\boldsymbol{\xi}\right)-\widehat{h}_{1i}\left(\boldsymbol{\xi}_0\right)\right\|^2\right\}+4E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\widehat{h}_{2i}\left(\boldsymbol{\xi}\right)-\widehat{h}_{2i}\left(\boldsymbol{\xi}_0\right)\right\|^2\right\}\\ =&4J_1+4J_2,\end{aligned}$$

so it suffices to show that $J_1 \le O\left(\frac{q}{\sqrt{n}}\right)$ and $J_2 \le O\left(\frac{q}{\sqrt{n}}\right)$.

We first show that that $J_1 \le O\left(\frac{q}{\sqrt{n}}\right)$. Note that

$$J_1 \le 2E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\frac{\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}\right)}{\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}\right)}-\frac{\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_0\right)}{\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_0\right)}\right\|^2\right\}$$

$$
\begin{aligned}
&+2E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\frac{\sum_{l=1}^m \partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta}\right)}{\sum_{l=1}^m\left\{1-\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta}\right)\right\}}-\frac{\sum_{l=1}^m \partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_{0,i}^l;\boldsymbol{\theta}_0\right)}{\sum_{l=1}^m\left\{1-\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_{0,i}^l;\boldsymbol{\theta}_0\right)\right\}}\right\|^2\right\}\\
&\widehat{=}2J_{11}+2J_{12}, \qquad\qquad (\text{A.31})
\end{aligned}
$$

where $\{\widetilde{y}_i^l\}_{l=1}^m$ are randomly drawn from $\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})$ and $\left\{\widetilde{y}_{0,i}^l\right\}_{l=1}^m$ are randomly drawn from $\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0)$, respectively. By Assumptions 3, there exists positive constants $U_{\pi'}$ and $0<L_\pi\le U_\pi<1$ such that $\|\partial_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta})\|\le(d+1)U_{\pi'}^2$ and $L_\pi\le|\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta})|\le U_\pi$. By Assumption 6, there exists a positive constant $U_{\pi''}$ such that $E\left\{\sum_{j=1}^{d+1}\|(\partial^2_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_1,y;\boldsymbol{\theta}))_{(.,j)}\|\right\}^2\le U_{\pi''}$ for every $\boldsymbol{\theta}\in\Theta$. Then, combining with Theorem 12.4 and the conclusion of Example 2 following this theorem of Apostol (1974), we have

$$
\begin{aligned}
J_{11}\le&2E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\frac{\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta};\right)\left\{\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}\right)-\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_0\right)\right\}}{\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}\right)\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_0\right)}\right\|^2\right\}\\
&+2E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\frac{1}{\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_0\right)}\left\{\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}\right)-\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_0\right)\right\}\right\|^2\right\}\\
\le&\frac{2(d+1)U_{\pi'}^2}{L_\pi^4}E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left|\left(\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_1\right)\right)^\top\left(\boldsymbol{\theta}-\boldsymbol{\theta}_0\right)\right|^2\right\}\\
&+\frac{2}{L_\pi^2}E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\partial^2_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_2\right)\left(\boldsymbol{\theta}-\boldsymbol{\theta}_0\right)\right\|^2\right\}\\
\le&\left(\frac{2(d+1)^2U_{\pi'}^4}{L_\pi^4}+\frac{2E\left\{\sum_{j=1}^{d+1}\|(\partial^2_{\boldsymbol{\theta}}\pi(\boldsymbol{x}_1,y;\boldsymbol{\theta}))_{(.,j)}\|\right\}^2}{L_\pi^2}\right)\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|^2\\
\le&\left(\frac{2(d+1)^2U_{\pi'}^4}{L_\pi^4}+\frac{2U_{\pi''}}{L_\pi^2}\right)\cdot B^2\frac{q}{n}=o\left(\frac{q}{\sqrt{n}}\right),
\end{aligned}
$$

where $\boldsymbol{\theta}_i=\boldsymbol{\theta}_0+c_i\left(\boldsymbol{\theta}-\boldsymbol{\theta}_0\right)$ and $c_i\in(0,1)$ for $i=1,2$.

For J_{12}, we have

$$
\begin{aligned}
J_{12}\le&2E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\frac{\frac{1}{m}\sum_{l=1}^m\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_{0,i}^l;\boldsymbol{\theta}_0\right)\cdot\frac{1}{m}\sum_{l=1}^m\left\{\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta}\right)-\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_{0,i}^l;\boldsymbol{\theta}_0\right)\right\}}{\frac{1}{m}\sum_{l=1}^m\left\{1-\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta}\right)\right\}\cdot\frac{1}{m}\sum_{l=1}^m\left\{1-\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_{0,i}^l;\boldsymbol{\theta}_0\right)\right\}}\right\|^2\right\}\\
&+2E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\frac{\frac{1}{m}\sum_{l=1}^m\left\{\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta}\right)-\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_{0,i}^l;\boldsymbol{\theta}_0\right)\right\}}{\frac{1}{m}\sum_{l=1}^m\left\{1-\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta}\right)\right\}}\right\|^2\right\}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{2(d+1)U_{\pi'}^2}{(1-U_\pi)^4} E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \left|\frac{1}{m}\sum_{l=1}^{m}\left\{\pi\left(\boldsymbol{x}_{1i}, \widetilde{y}_i^l\, \boldsymbol{\theta}\right) - \pi\left(\boldsymbol{x}_{1i}, \widetilde{y}_{0,i}^l; \boldsymbol{\theta}_0\right)\right\}\right|^2\right\} \\
&\quad + \frac{2}{(1-U_\pi)^2} E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \left\|\frac{1}{m}\sum_{l=1}^{m}\left\{\partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta}\right) - \partial_{\boldsymbol{\theta}}\pi\left(\boldsymbol{x}_{1i}, \widetilde{y}_{0,i}^l; \boldsymbol{\theta}_0\right)\right\}\right\|^2\right\} \\
&\widehat{=} \frac{2(d+1)U_{\pi'}^2}{(1-U_\pi)^4} JJ_1 + \frac{2}{(1-U_\pi)^2} JJ_2 \qquad \text{(A.32)}
\end{aligned}
$$

In what follows, we show that $JJ_1 \leq O(\frac{q}{\sqrt{n}})$. Let's first prove the following inequality

$$
\sup_{\boldsymbol{\phi}:\|\boldsymbol{\phi}-\boldsymbol{\phi}_0\|\leq B\sqrt{\frac{q}{n}}} \widetilde{f}(y|\boldsymbol{x};\boldsymbol{\phi}) < \infty. \qquad \text{(A.33)}
$$

for sufficiently large n. By Assumption 2, it is easy to show that $\widetilde{f}(y|\boldsymbol{x};\boldsymbol{\phi}_0) < \infty$. That is, for every given $\boldsymbol{x} \in \mathcal{X}$ and $k \in \{1, 2, \ldots, k_n - 1\}$, there esists a positive constant $U_{\widetilde{f}} < \infty$ such that $(\tau_{k+1} - \tau_k)/\boldsymbol{x}^\top(\boldsymbol{\beta}_{0,\tau_{k+1}} - \boldsymbol{\beta}_{0,\tau_k}) \leq U_{\widetilde{f}}$. Then, for every $\boldsymbol{\phi}$ satisfying $\|\boldsymbol{\phi} - \boldsymbol{\phi}_0\| \leq B\sqrt{\frac{q}{n}}$, when n is sufficiently large, for every $k \in \{1, 2, \ldots, k_n - 1\}$,

$$
\frac{(\tau_{k+1} - \tau_k)}{\boldsymbol{x}^\top(\boldsymbol{\beta}_{\tau_{k+1}} - \boldsymbol{\beta}_{\tau_k})} = \frac{(\tau_{k+1} - \tau_k)}{\boldsymbol{x}^\top(\boldsymbol{\beta}_{\tau_{0,k+1}} - \boldsymbol{\beta}_{\tau_{0,k}}) + \boldsymbol{x}^\top(\boldsymbol{\beta}_{\tau_{k+1}} - \boldsymbol{\beta}_{\tau_{0,k+1}}) - \boldsymbol{x}^\top(\boldsymbol{\beta}_{\tau_k} - \boldsymbol{\beta}_{\tau_{0,k}})} \leq 2U_{\widetilde{f}}.
$$

Thus, we have $\sup_{\boldsymbol{\phi}:\|\boldsymbol{\phi}-\boldsymbol{\phi}_0\|\leq B\sqrt{\frac{q}{n}}} \widetilde{f}(y|\boldsymbol{x};\boldsymbol{\phi}) \leq 2U_{\widetilde{f}} < \infty$.

Note that

$$
\begin{aligned}
JJ_1 \leq& 2E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \left|\frac{1}{m}\sum_{l=1}^{m}\left\{\pi\left(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta}\right) - \pi\left(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta}_0\right)\right\}\right|^2\right\} \\
&+2E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \left|\frac{1}{m}\sum_{l=1}^{m}\left\{\pi\left(\boldsymbol{x}_{1i}, \widetilde{y}_i^l; \boldsymbol{\theta}_0\right) - \pi\left(\boldsymbol{x}_{1i}, \widetilde{y}_{0,i}^l; \boldsymbol{\theta}_0\right)\right\}\right|^2\right\} \\
&\leq 2(d+1)U_{\pi'}^2 \sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|^2 \\
&+2E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \left|\int_y \pi\left(\boldsymbol{x}_{1i}, y; \boldsymbol{\theta}_0\right)\left\{\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}\right) - \widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0\right)\right\}dy + O_p\left(\frac{1}{\sqrt{m}}\right)\right|^2\right\} \\
&\leq 2(d+1)U_{\pi'}^2 \cdot B^2\frac{q}{n} \\
&\quad + 4U_\pi^2 E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \left(\int_y \left|\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}\right) - \widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0\right)\right|dy\right)^2\right\} + O\left(\frac{1}{m}\right)
\end{aligned}
$$

$$=O\left(\frac{q}{n}+\frac{1}{m}\right)+4U_\pi^2E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\int_y\left|\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}\right)-\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0\right)\right|dy\right)^2\right\}.$$

The first term of the the second inequality above is followed by Assumption 3, and the second term is followed by (A.33) and the central limit theorem. Note that $0<\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}_0\right)<1$, so it follows from the central limit theorem that

$$\begin{aligned}&\frac{1}{m}\sum_{l=1}^m\left[\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_i^l;\boldsymbol{\theta}_0\right)-\pi\left(\boldsymbol{x}_{1i},\widetilde{y}_{0,i}^l;\boldsymbol{\theta}_0\right)\right]\\&=\int_y\pi\left(\boldsymbol{x}_{1i},y;\boldsymbol{\theta}_0\right)\left[\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}\right)-\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0\right)\right]dy+O_p\left(\frac{1}{\sqrt{m}}\right).\end{aligned}$$

Moreover, the first term of the last equation above $O\left(\frac{q}{n}+\frac{1}{m}\right)=o\left(\frac{q}{\sqrt{n}}\right)$ as $n\to\infty$. Therefore, it suffices to show that

$$E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\int_y\left|\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}\right)-\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0\right)\right|dy\right)^2\right\}\le O\left(\frac{q}{\sqrt{n}}\right).$$

Consider the case where $\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\frac{k}{k_n+1}}\le\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\frac{k}{k_n+1}}<\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\frac{k+1}{k_n+1}}\le\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\frac{k+1}{k_n+1}}$ for $k=1,\dots k_n$, and the following argument can be similarly used for other cases. We then decompose $|\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}\right)-\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0\right)|$ as

$$\begin{aligned}&\left|\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}\right)-\widetilde{f}\left(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0\right)\right|\\
&=\frac{1}{k_n+1}\left|\frac{1}{\boldsymbol{x}_i^\top(\boldsymbol{\beta}_{0,\frac{2}{k_n+1}}-\boldsymbol{\beta}_{0,\frac{1}{k_n+1}})}\boldsymbol{I}(\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\frac{1}{k_n+1}}\le y\le\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\frac{1}{k_n+1}})\right|\\
&+\frac{1}{k_n+1}\left|\sum_{k=1}^{k_n-1}\left(\frac{1}{\boldsymbol{x}_i^\top(\boldsymbol{\beta}_{0,\frac{k+1}{k_n+1}}-\boldsymbol{\beta}_{0,\frac{k}{k_n+1}})}-\frac{1}{\boldsymbol{x}_i^\top(\boldsymbol{\beta}_{\frac{k+1}{k_n+1}}-\boldsymbol{\beta}_{\frac{k}{k_n+1}})}\right)\boldsymbol{I}(\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\frac{k}{k_n+1}}\le y\le\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\frac{k+1}{k_n+1}})\right|\\
&+\frac{1}{k_n+1}\left|\sum_{k=1}^{k_n-2}\left(\frac{1}{\boldsymbol{x}_i^\top(\boldsymbol{\beta}_{0,\frac{k+2}{k_n+1}}-\boldsymbol{\beta}_{0,\frac{k+1}{k_n+1}})}-\frac{1}{\boldsymbol{x}_i^\top(\boldsymbol{\beta}_{\frac{k+1}{k_n+1}}-\boldsymbol{\beta}_{\frac{k}{k_n+1}})}\right)\boldsymbol{I}(\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\frac{k+1}{k_n+1}}\le y\le\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\frac{k+1}{k_n+1}})\right|\\
&+\frac{1}{k_n+1}\left|\frac{1}{\boldsymbol{x}_i^\top(\boldsymbol{\beta}_{\frac{k_n}{k_n+1}}-\boldsymbol{\beta}_{\frac{k_n-1}{k_n+1}})}\boldsymbol{I}(\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\frac{k_n}{k_n+1}}\le y\le\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\frac{k_n}{k_n+1}})\right|\\
&\widehat{=}F_1+F_2+F_3+F_4.\end{aligned}\tag{A.35}$$

Thus we obtain that

$$
\begin{aligned}
&E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\int_y\left|\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi})-\widetilde{f}(y|\boldsymbol{x}_i;\boldsymbol{\phi}_0)\right|dy\right)^2\right\}\\
\le&4\sum_{j=1}^4 E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\int_y F_j dy\right)^2\right\}.
\end{aligned}
$$

In what follws, we will show $E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\int_y F_j dy\right)^2\right\}\le O\left(\frac{q}{\sqrt{n}}\right)$ for $j=1,2,3,$ and $4.$ Under Assumptions 1, there exist some positive constants Q that $E_{\boldsymbol{x}_i}\|\boldsymbol{x}_i\|^2<Q$. Hence, for $j=1$, we have

$$
\begin{aligned}
&E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\int_y F_1 dy\right)^2\right\}\\
=&E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\widetilde{f}(y_1|\boldsymbol{x}_i,\boldsymbol{\phi}_0)\int_{\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\frac{1}{k_n+1}}}^{\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\frac{1}{k_n+1}}}1dy\right)^2\right\}\\
\le&U_{\widetilde{f}}^2E_{\boldsymbol{x}_i}\left\{\|\boldsymbol{x}_i\|^2\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\boldsymbol{\beta}_{\frac{1}{k_n+1}}-\boldsymbol{\beta}_{0,\frac{1}{k_n+1}}\right\|^2\right\}\\
\le&QU_{\widetilde{f}}^2B^2\frac{q}{n}=o\left(\frac{q}{\sqrt{n}}\right),
\end{aligned}
$$

where $y_1=\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\frac{1}{k_n+1}}+c_1\boldsymbol{x}_i^\top\left(\boldsymbol{\beta}_{0,\frac{2}{k_n+1}}-\boldsymbol{\beta}_{0,\frac{1}{k_n+1}}\right)$ with $0<c_1<1$. With the similar argument, we can also show that $E_x\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\int_y F_4 dy\right)^2\right\}\le o\left(\frac{q}{\sqrt{n}}\right)$. For $j=2$, by Assumptions 1 and 2, we have

$$
\begin{aligned}
&E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\int_y F_2 dy\right)^2\right\}\\
\le&2E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\sum_{k=1}^{k_n-1}\widetilde{f}(y_{0,k}|\boldsymbol{x}_i,\boldsymbol{\phi}_0)\frac{\boldsymbol{x}_i^\top\left(\boldsymbol{\beta}_{\frac{k+1}{k_n+1}}-\boldsymbol{\beta}_{0,\frac{k+1}{k_n+1}}\right)}{\boldsymbol{x}_i^\top\left(\boldsymbol{\beta}_{\frac{k+1}{k_n+1}}-\boldsymbol{\beta}_{\frac{k}{k_n+1}}\right)}\int_{\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\frac{k}{k_n+1}}}^{\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\frac{k+1}{k_n+1}}}1dy\right)^2\right\}\\
&+2E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\sum_{k=1}^{k_n-1}\widetilde{f}(y_k|\boldsymbol{x}_i,\boldsymbol{\phi})\frac{\boldsymbol{x}_i^\top\left(\boldsymbol{\beta}_{\frac{k}{k_n+1}}-\boldsymbol{\beta}_{0,\frac{k}{k_n+1}}\right)}{\boldsymbol{x}_i^\top\left(\boldsymbol{\beta}_{\frac{k+1}{k_n+1}}-\boldsymbol{\beta}_{\frac{k}{k_n+1}}\right)}\int_{\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\frac{k}{k_n+1}}}^{\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\frac{k+1}{k_n+1}}}1dy\right)^2\right\}\\
\le&10U_{\widetilde{f}}^2E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\|\boldsymbol{x}_i\|^2\left\|\sum_{k=1}^{k_n}\left(\boldsymbol{\beta}_{\frac{k}{k_n+1}}-\boldsymbol{\beta}_{0,\frac{k}{k_n+1}}\right)\right\|^2\right\}\\
\le&10U_{\widetilde{f}}^2Q\cdot k_n\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|^2
\end{aligned}
$$

$$=o\left(\frac{q}{\sqrt{n}}\right),$$

where $y_{0,k}=\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\frac{k}{k_n+1}}+c_{0,k}\boldsymbol{x}_i^\top(\boldsymbol{\beta}_{0,\frac{k+1}{k_n+1}}-\boldsymbol{\beta}_{0,\frac{k}{k_n+1}})$ and $y_k=\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\frac{k}{k_n+1}}+c_k\boldsymbol{x}_i^\top(\boldsymbol{\beta}_{\frac{k+1}{k_n+1}}-\boldsymbol{\beta}_{\frac{k}{k_n+1}})$ with $c_{0,k}\in(0,1),c_k\in(0,1)$ for $k=1,\ldots,k_n-1$. With the similar argument, we can also show that $E_{\boldsymbol{x}_i}\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left(\int_y F_3dy\right)^2\right\}\le o\left(\frac{q}{\sqrt{n}}\right)$. Hence, we have shown that $JJ_1=o\left(\frac{q}{\sqrt{n}}\right)$. Similarly, we also can show that $JJ_2=o\left(\frac{q}{\sqrt{n}}\right)$. Therefore, by (A.33), we have shown that $J_{12}\le o\left(\frac{q}{\sqrt{n}}\right)$. The proof of $J_1\le O\left(\frac{q}{\sqrt{n}}\right)$ is hence completed.

In what follows, we show that $J_2\le O\left(\frac{q}{\sqrt{n}}\right)$. Note that

$$\begin{aligned}
&E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\|h_{2i}(\boldsymbol{\xi})-h_{2i}(\boldsymbol{\xi}_0)\|^2\right\}\\
=&E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\delta_i\left(\frac{\Psi(y_i-\boldsymbol{x}_i^\top\boldsymbol{\phi})}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta})}-\frac{\Psi(y_i-\boldsymbol{x}_i^\top\boldsymbol{\phi}_0)}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_0)}\right)\otimes\boldsymbol{x}_i\right\|^2\right\}\\
\le&E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\sum_{k=1}^{k_n}\left|\frac{\Psi(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k})}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta})}-\frac{\Psi(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k})}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_0)}\right|^2\cdot\|\boldsymbol{x}_i\|^2\right\}\\
\le&2E\left\{\|\boldsymbol{x}_i\|^2\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\sum_{k=1}^{k_n}\left|\frac{1}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta})}-\frac{1}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_0)}\right|^2\cdot|\Psi(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k})|^2\right\}\\
&+2E\left\{\|\boldsymbol{x}_i\|^2\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\sum_{k=1}^{k_n}\left|\frac{1}{\pi(\boldsymbol{x}_{1i},y_i;\boldsymbol{\theta}_0)}\right|^2\cdot|\Psi(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k})-\Psi(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k})|^2\right\}\\
\widehat{=}&2(MM_1+MM_2).
\end{aligned}\tag{A.36}$$

By Assumptions 1 and 3, for $k_n^2/n\to0$ and $n\to\infty$, we have that

$$MM_1=O\left(k_n\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|^2\right)\le O\left(\frac{q}{\sqrt{n}}\right).\tag{A.37}$$

On the other hand, note that

$$\begin{aligned}
&|\Psi(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k})-\Psi(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k})|^2\\
=&|\boldsymbol{I}(y_i<\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k})-\boldsymbol{I}(y_i<\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k})|^2\\
=&\boldsymbol{I}(\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k}<y_i<\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k};\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k}\le\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k})+\boldsymbol{I}(\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k}<y_i<\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k};\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k}\le\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k}).
\end{aligned}$$

Under Assumptions 1 and 2, there exists a positive constant M such that $E_{\boldsymbol{x}_i}\|\boldsymbol{x}_i\|^3<M$, and

a constant U_f such that $f\left(y|\boldsymbol{x},\boldsymbol{\beta}_0\left(\tau\right)\right) < U_f$ for any $y \in \mathcal{Y}$. By the mean value theorem and Assumptions 1 and 2, we obtain that

$$
\begin{aligned}
MM_2 \leq & \frac{1}{L_\pi^2} E\left\{\|\boldsymbol{x}_i\|^2 \sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \sum_{k=1}^{k_n} \boldsymbol{I}\left(\boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k} < y_i < \boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k}; \boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k} \leq \boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k}\right)\right\} \\
& + \frac{1}{L_\pi^2} E\left\{\|\boldsymbol{x}_i\|^2 \sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \sum_{k=1}^{k_n} \boldsymbol{I}\left(\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k} < y_i < \boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k}; \boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k} \leq \boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k}\right)\right\} \\
= & \frac{1}{L_\pi^2} E_{\boldsymbol{x}_i}\left\{\|\boldsymbol{x}_i\|^2 \sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \sum_{k=1}^{k_n} f\left(y_{i,k}|\boldsymbol{x}_i;\boldsymbol{\beta}_0\left(\tau\right)\right)\left|\boldsymbol{x}_i^\top\left(\boldsymbol{\beta}_{\tau_k}-\boldsymbol{\beta}_{0,\tau_k}\right)\right|\right\} \\
\leq & \frac{U_f}{L_\pi^2} E_{\boldsymbol{x}_i}\left\{\|\boldsymbol{x}_i\|^3 \sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \sum_{k=1}^{k_n} \|\boldsymbol{\beta}_{\tau_k}-\boldsymbol{\beta}_{0,\tau_k}\|\right\} \\
\leq & \frac{U_f M}{L_\pi^2}\sqrt{k_n} \sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \|\boldsymbol{\xi}-\boldsymbol{\xi}_0\| \\
= & O\left(\frac{q}{\sqrt{n}}\right),
\end{aligned} \tag{A.38}
$$

where $y_{i,k} = \boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k} + c_k\left(\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\tau_k} - \boldsymbol{x}_i^\top\boldsymbol{\beta}_{0,\tau_k}\right)$ with $0 < c_k < 1$. By (A.36), (A.37) and (A.38), we have $J_2 \leq O\left(\frac{q}{\sqrt{n}}\right)$. The proof of Lemma 6 is hence completed. □

Proof of Theorem 2

Proof. Since $k_n^{3+2v}/n \to \infty, k_n^2\log^2 n/n \to 0, k_n m/n \to \infty$ as $n \to \infty$, by Lemma 2.1 of He and Shao (2000) and Lemma 6, for any given B, and any $\boldsymbol{\alpha} \in \mathcal{R}^q$, $\|\boldsymbol{\alpha}\| = 1$, we have

$$
\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \sum_{i=1}^{n} E\left|\boldsymbol{\alpha}^\top\eta_i\left(\boldsymbol{\xi},\boldsymbol{\xi}_0\right)\right|^2 = O\left(q\sqrt{n}\right),
$$

and

$$
\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}} \sum_{i=1}^{n} \left(\boldsymbol{\alpha}^\top\eta_i\left(\boldsymbol{\xi},\boldsymbol{\xi}_0\right)\right)^2 = O_p\left(q\sqrt{n}\right).
$$

Then by Lemma 3.3 of He and Shao (2000), we have

$$\begin{aligned}&\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\leq B\sqrt{\frac{q}{n}}}\sqrt{\frac{n}{q}}\left|\boldsymbol{\alpha}^{\top}\left\{\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi})-\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}_0)-\widetilde{\boldsymbol{H}}(\boldsymbol{\xi})+\widetilde{\boldsymbol{H}}(\boldsymbol{\xi}_0)\right\}\right| \\ &=O_p\left(\left(\frac{q^2\log^2 n}{n}\right)^{\frac{1}{4}}\right)=o_p(1).\end{aligned} \tag{A.39}$$

Combining (A.39) with (A.27) and Lemma 5, we further obtain

$$\sqrt{\frac{n}{q}}\left|\boldsymbol{\alpha}^{\top}\left\{\widehat{\boldsymbol{H}}_n\left(\widehat{\boldsymbol{\xi}}_n\right)-\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}_0)-\boldsymbol{H}\left(\widehat{\boldsymbol{\xi}}_n\right)+\boldsymbol{H}(\boldsymbol{\xi}_0)\right\}\right|=o_p(1). \tag{A.40}$$

Note that $\sqrt{n/q}\left\|\boldsymbol{\alpha}^{\top}\widehat{\boldsymbol{H}}_n\left(\widehat{\boldsymbol{\xi}}_n\right)\right\|=o_p(1)$ and $\sqrt{n/q}\left\|\boldsymbol{\alpha}^{\top}\left\{\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi}_0)-\boldsymbol{H}_n^0(\boldsymbol{\xi}_0)\right\}\right\|=o_p(1)$, which is implied by the construction of the proposed estimator and Lemma 4, respectively. It follows that

$$\sqrt{\frac{n}{q}}\left|\boldsymbol{\alpha}^{\top}\left\{\boldsymbol{H}_n^0(\boldsymbol{\xi}_0)+\boldsymbol{H}\left(\widehat{\boldsymbol{\xi}}_n\right)-\boldsymbol{H}(\boldsymbol{\xi}_0)\right\}\right|=o_p(1).$$

Then by Assumption 8, we have

$$\sqrt{\frac{n}{q}}\boldsymbol{\alpha}^{\top}D_n\left(\widehat{\boldsymbol{\xi}}_n-\boldsymbol{\xi}_0\right)=-\sqrt{\frac{n}{q}}\boldsymbol{\alpha}^{\top}\boldsymbol{H}_n^0(\boldsymbol{\xi}_0)+o_p(1).$$

Theorem 2 follows immediately from the central limit theorem. □

Web Appendix D

Suppose that P_B is the bootstrap probability given the observed data $\{\boldsymbol{x}_i, y_i, \delta_i\}_{i=1}^n$. Here, we further denote E^* and P^* be the expectation and probability on $\{y_i, \boldsymbol{x}_i, \delta_i, w_i\}_{i=1}^n$. Let

$$\boldsymbol{H}_{w,n}^0(\boldsymbol{\xi})=\left(\boldsymbol{S}_{w,n}^{\top}(\boldsymbol{\theta}),\boldsymbol{M}_{w,n}^{\top}(\boldsymbol{\xi})\right)^{\top}=\frac{1}{n}\sum_{i=1}^n w_i h_i(\boldsymbol{\xi})=\left(\frac{1}{n}\sum_{i=1}^n w_i h_{1i}^{\top}(\boldsymbol{\xi}),\frac{1}{n}\sum_{i=1}^n w_i h_{2i}^{\top}(\boldsymbol{\xi})\right)^{\top},$$

$$\widehat{\boldsymbol{H}}_{w,n}(\boldsymbol{\xi})=\left(\widehat{\boldsymbol{S}}_{w,n}^{\top}(\boldsymbol{\xi}),\boldsymbol{M}_{w,n}^{\top}(\boldsymbol{\xi})\right)^{\top}=\frac{1}{n}\sum_{i=1}^n w_i\widehat{h}_i(\boldsymbol{\xi})=\left(\frac{1}{n}\sum_{i=1}^n w_i\widehat{h}_{1i}^{\top}(\boldsymbol{\xi}),\frac{1}{n}\sum_{i=1}^n w_i\widehat{h}_{2i}^{\top}(\boldsymbol{\xi})\right)^{\top},$$

and $\widehat{\boldsymbol{\xi}}_n^*$ be the weighted bootstrap estimator of $\boldsymbol{\xi}$. Define

$$\eta_{w,i}(\boldsymbol{\xi},\boldsymbol{\xi}_0)=w_i\widehat{h}_i(\boldsymbol{\xi})-w_i\widehat{h}_i(\boldsymbol{\xi}_0)-E^*\left[w_i\widehat{h}_i(\boldsymbol{\xi})\right]+E^*\left[w_i\widehat{h}_i(\boldsymbol{\xi}_0)\right]. \tag{A.41}$$

Note that by Assumption 10, we have $E^*\{\boldsymbol{H}_{w,n}^0(\boldsymbol{\xi})\}=E\{\boldsymbol{H}_n^0(\boldsymbol{\xi})\}=\boldsymbol{H}^0(\boldsymbol{\xi})$ and $E^*\{\widehat{\boldsymbol{H}}_{w,n}(\boldsymbol{\xi})\}=E\{\widehat{\boldsymbol{H}}_n(\boldsymbol{\xi})\}=\widetilde{\boldsymbol{H}}(\boldsymbol{\xi})$ for any $\boldsymbol{\xi}\in\Theta_{\boldsymbol{\xi}}$.

LEMMA 7: *Under Assumptions 1-4, 6-7 and 9, for $n \to \infty, k_n^{3+2v}/n \to \infty$ and $k_n m/n \to \infty$, we have*

$$\sqrt{\frac{n}{q}}\left\|\widehat{\boldsymbol{H}}_{w,n}(\boldsymbol{\xi}_0) - \boldsymbol{H}^0_{w,n}(\boldsymbol{\xi}_0)\right\| = o_{p^*}(1). \tag{A.42}$$

Proof. This Lemma can be shown with the similar argument of the proof of Lemma 4. □

LEMMA 8: *Under Assumptions 1-9, for $n \to \infty, k_n^{3+2v}/n \to \infty$ and $k_n m/n \to \infty$, we have*

$$\left\|\widehat{\boldsymbol{\xi}}^*_n - \boldsymbol{\xi}_0\right\| = O_{p^*}\left(\sqrt{\frac{q}{n}}\right). \tag{A.43}$$

Proof. This Lemma can be shown with the similar argument of the proof of Lemma 5. □

LEMMA 9: *Under Assumptions 1-9 as $n \to \infty, k_n^2/n \to 0, k_n^{3+2v}/n \to \infty, k_n m/n \to \infty$, for any given $B > 0$, we have*

$$E^\star\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\|\eta_{w,i}(\boldsymbol{\xi},\boldsymbol{\xi}_0)\|^2\right\} \le O\left(\frac{q}{\sqrt{n}}\right), \tag{A.44}$$

for $i = 1,\dots,n$.

Proof. Note that

$$\begin{aligned}
&E^*\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\|\eta_{w,i}(\boldsymbol{\xi},\boldsymbol{\xi}_0)\|^2\right\}\\
=&E^*\left[\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|w_i\widehat{h}_i(\boldsymbol{\xi}) - w_i\widehat{h}_i(\boldsymbol{\xi}_0) - E^*\left\{w_i\widehat{h}_i(\boldsymbol{\xi})\right\} + E^*\left\{w_i\widehat{h}_i(\boldsymbol{\xi}_0)\right\}\right\|^2\right]\\
\le&2E^*\left\{w_i^2\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\widehat{h}_i(\boldsymbol{\xi}) - \widehat{h}_i(\boldsymbol{\xi}_0)\right\|^2\right\} + 2\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|E\left\{\widehat{h}_i(\boldsymbol{\xi}) - \widehat{h}_i(\boldsymbol{\xi}_0)\right\}\right\|^2\\
\le&6E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\widehat{h}_i(\boldsymbol{\xi}) - \widehat{h}_i(\boldsymbol{\xi}_0)\right\|^2\right\}\\
=&6E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\widehat{h}_{1i}(\boldsymbol{\xi}) - \widehat{h}_{1i}(\boldsymbol{\xi}_0)\right\|^2\right\} + 6E\left\{\sup_{\boldsymbol{\xi}:\|\boldsymbol{\xi}-\boldsymbol{\xi}_0\|\le B\sqrt{\frac{q}{n}}}\left\|\widehat{h}_{2i}(\boldsymbol{\xi}) - \widehat{h}_{2i}(\boldsymbol{\xi}_0)\right\|^2\right\}\\
=&6J_1 + 6J_2.
\end{aligned}$$

The last inequality follows under Assumption 9. We have shown that $J_1 \leq O\left(q/\sqrt{n}\right)$ and $J_2 \leq O\left(q/\sqrt{n}\right)$ in Lemma 6, so (A.44) hence holds. □

Proof of Theorem 3

Proof. For any $\boldsymbol{\alpha} \in \mathcal{R}^q$ with $\|\boldsymbol{\alpha}\| = 1$, by some algebra, we obtain that

$$\begin{aligned}
&\sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\left\{\widetilde{\boldsymbol{H}}\left(\widehat{\boldsymbol{\xi}}_n^*\right) - \widetilde{\boldsymbol{H}}\left(\widehat{\boldsymbol{\xi}}_n\right)\right\} + \sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\left\{\boldsymbol{H}_{w,n}^0\left(\boldsymbol{\xi}_0\right) - \boldsymbol{H}_n^0\left(\boldsymbol{\xi}_0\right)\right\}\\
&= -\sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\left\{\widehat{\boldsymbol{H}}_{w,n}\left(\widehat{\boldsymbol{\xi}}_n^*\right) - \widehat{\boldsymbol{H}}_{w,n}\left(\boldsymbol{\xi}_0\right) - \widetilde{\boldsymbol{H}}\left(\widehat{\boldsymbol{\xi}}_n^*\right) + \widetilde{\boldsymbol{H}}\left(\boldsymbol{\xi}_0\right)\right\}\\
&\quad + \sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\left\{\widehat{\boldsymbol{H}}_n\left(\widehat{\boldsymbol{\xi}}_n\right) - \widehat{\boldsymbol{H}}_n\left(\boldsymbol{\xi}_0\right) - \widetilde{\boldsymbol{H}}\left(\widehat{\boldsymbol{\xi}}_n\right) + \widetilde{\boldsymbol{H}}\left(\boldsymbol{\xi}_0\right)\right\}\\
&\quad + \sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\left\{\widehat{\boldsymbol{H}}_n\left(\boldsymbol{\xi}_0\right) - \boldsymbol{H}_n^0\left(\boldsymbol{\xi}_0\right)\right\} - \sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\left\{\widehat{\boldsymbol{H}}_{w,n}\left(\boldsymbol{\xi}_0\right) - \boldsymbol{H}_{w,n}^0\left(\boldsymbol{\xi}_0\right)\right\}\\
&\quad - \sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\widehat{\boldsymbol{H}}_n\left(\widehat{\boldsymbol{\xi}}_n\right) + \sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\widehat{\boldsymbol{H}}_{w,n}\left(\widehat{\boldsymbol{\xi}}_n^*\right)\\
&\;\widehat{=} \sum_{i=1}^{6} T_i.
\end{aligned} \tag{A.45}$$

Note that $|T_2| = o_p(1)$, $|T_3| = o_p(1)$ and $|T_4| = o_{p^*}(1)$ which is implied by (A.40), Lemmas 4 and 7, respectively. Moreover, $|T_5| = o_p(1)$ and $|T_6| = o_{p^*}(1)$ according to the construction of the proposed estimator $\widehat{\boldsymbol{\xi}}_n$ and the weighted bootstrap estimator $\widehat{\boldsymbol{\xi}}_n^*$. Moreover, by Lemmas 8 and 9 and the similar argument as that of proving (A.40), we obtain that $|T_1| = o_{p^*}(1)$. Therefore, we have

$$\begin{aligned}
&\left|\sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\left\{\widetilde{\boldsymbol{H}}\left(\widehat{\boldsymbol{\xi}}_n^*\right) - \widetilde{\boldsymbol{H}}\left(\widehat{\boldsymbol{\xi}}_n\right)\right\} + \sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\left\{\boldsymbol{H}_{w,n}^0\left(\boldsymbol{\xi}_0\right) - \boldsymbol{H}_n^0\left(\boldsymbol{\xi}_0\right)\right\}\right|\\
&\leq \sum_{i=1}^{6}|T_i| = o_{p^*}(1) + o_p(1) = o_{p^*}(1).
\end{aligned} \tag{A.46}$$

By (A.27) and Assumption 8, we have

$$\begin{aligned}
&\sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top D_n\left(\widehat{\boldsymbol{\xi}}_n^* - \widehat{\boldsymbol{\xi}}_n\right)\\
&= -\sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\left\{\boldsymbol{H}_{w,n}^0\left(\boldsymbol{\xi}_0\right) - \boldsymbol{H}_n^0\left(\boldsymbol{\xi}_0\right)\right\} + o_{p^*}(1)\\
&= -\sqrt{\frac{n}{q}}\boldsymbol{\alpha}^\top\sum_{i=1}^{n}\left(w_i - 1\right)h_i\left(\boldsymbol{\xi}_0\right) + o_{p^*}(1)
\end{aligned}$$

Under Assumption 9, by Lemma 4.6 of Præstgaard and Wellner (1993) and Lemma 3 of Cheng and Huang (2010), we obtain

$$\sup_{z\in\mathcal{R}}\left|P_B\left(\frac{\sqrt{n}\boldsymbol{\alpha}^\top\left(\widehat{\boldsymbol{\xi}}_n^* - \widehat{\boldsymbol{\xi}}_n\right)}{\sigma\left(\boldsymbol{\alpha}\right)} \le z\right) - P\left(Z \le z\right)\right| = o_p(1), \tag{A.47}$$

where Z is a standard normal random variable. Theorem 2 together with Lemma 2.11 of Van der Vaart (2000) implies that

$$\sup_{z\in\mathcal{R}}\left|P\left(\frac{\sqrt{n}\boldsymbol{\alpha}^\top\left(\widehat{\boldsymbol{\xi}}_n - \boldsymbol{\xi}_0\right)}{\sigma\left(\boldsymbol{\alpha}\right)} \le z\right) - P\left(Z \le z\right)\right| = o(1) \tag{A.48}$$

Combining (A.47) and (A.48), we complete the proof of Theorem 3. □

Web Appendix E

E.1 Simulation results for different settings, sample sizes and missing rates

[Web Table 1 about here.]

[Web Table 2 about here.]

[Web Table 3 about here.]

[Web Table 4 about here.]

[Web Table 5 about here.]

[Web Table 6 about here.]

[Web Table 7 about here.]

[Web Table 8 about here.]

[Web Table 9 about here.]

[Web Table 10 about here.]

[Web Table 11 about here.]

E.2 Selection of turning parameters k_n and m

Here, we conduct numerical investigations to explore the effect of the numbers of k_n and m on the performance of the proposed method by using different combinations of the numbers of k_n ($k_n = 20, 30, 40, 50, 100$) and m ($m = 10, 20, 50, 100$).

The resulting RMS, 95ECP and CI length from 500 Monte-Carlo replicates from the proposed method with different numbers of k_n and m under Setting 3 with different sample sizes are presented in Web Tables 12–14. We find that, when taking the number of k_n between $\lfloor n^{0.4} \rfloor$ and $\lfloor 3 * n^{0.4} + 4 \rfloor$, our proposed method performs well under Setting 3 with different sample sizes. When k_n increases, the asymptotic standard errors estimated via bootstrapping will be larger. This is reasonable, because when k_n increases, the number of unknown parameters to be estimated will also increase, which reduces the estimation efficiency of the proposed estimates. These conclusions can also be confirmed under Settings 1–2. Therefore, we recommend taking the number of k_n between $\lfloor n^{0.4} \rfloor$ and $\lfloor 3 * n^{0.4} + 4 \rfloor$ in practical implementation. Besides, when the number of random draws m increases, the RMS, 95ECP and CI length of quantile regression coefficients under all the three settings with different sample sizes remain nearly unchanged. A small m between 10 and 20 is sufficient to stabilize the estimated coefficients. Bigger m does not further improve the accuracy in our simulations. Considering the computational cost, we recommend taking the number of m between 10 and 20 in practical implementation.

[Web Table 12 about here.]

[Web Table 13 about here.]

[Web Table 14 about here.]

E.3 Comparison of computing time

Web Table 15 displays the average computing time (in seconds) for different methods to estimate the coefficients $\boldsymbol{\beta}_\tau$ at the quantile levels $\tau = 0.25, 0.5$ and 0.75 from 500 Monte-Carlo replicates under three settings with various sample sizes and missing rates as considered in our paper. For the proposed method (proIpwQr), we also investigate the effect of the numbers of k_n and m on its average computing time by using different combinations of $k_n (k_n = 20, 40)$ and $m (m = 10, 20)$.

Under Setting 1 with sample size $n = 500$ and 20% missing rate in y, the average computing time of the proposed proIpwQr with $(k_n, m) = (20, 10)$ is 0.840s, and the average computing time of elmIpwQr, elsIpwQr and swelQr is 35.232s, 4.911s and 7.573s, respectively. When increasing m to 20 or k_n to 40, the average computing time of the proposed proIpwQr will also increases slightly. When the sample size n is increased from 500 to 1000, the average computing time of the proposed proIpwQr approximately doubles, the average computing time of elsIpwQr and swelQr will both increase to about 1.25 times, and the average computing time of elmIpwQr will increase to about two to three times. These conclusions can also be confirmed under Settings 2 and 3. Comparing the results under the three settings, the conditional distribution of the response y given covariate $\boldsymbol{x}$, $F(y \mid \boldsymbol{x})$, has a small effect on the average computing time of elsIpwQr, swelQr and proIpwQr, and has a more significant effect on the average computing time of elmIpwQr.

Under all these considered settings, the proposed proIpwQr takes less average computing time than its competitors (except NaiveQr). When the sample size n is doubled, the average computing time of the proposed proIpwQr also doubles. The numbers of k_n and m also affect the average computing time of the proposed proIpwQr. When k_n or m increases, the average computing time of the proposed proIpwQr also increases. The conditional distribution of the response y given covariate $\boldsymbol{x}$, $F(y \mid \boldsymbol{x})$, has a small effect on the average computing time of the proposed proIpwQr.

[Web Table 15 about here.]

Web Appendix F

F.1 Forest plots for 40% and 60% Missing Rates

[Web Figure 1 about here.]

[Web Figure 2 about here.]

[Web Figure 3 about here.]

[Web Figure 4 about here.]

[Web Figure 5 about here.]

F.2 Estimation Results of Real Data for 40% and 60% Missing Rates

[Web Table 16 about here.]

[Web Table 17 about here.]

F.3 Estimation Results of θ from Real Data

The following tables are the $\boldsymbol{\theta}$ estimates when analysing the EMR data from 18744 ICU patients. Table 18 are obtained based on the EMR data from 18744 ICU patients, 7148 of whom had glucose measurements at their admission. Based on those 7148 ICU completely observed data (as a full dataset), Tables 19 and 20 are the estimated results of $\boldsymbol{\theta}$ based on 40% and 60% artificially missing data, respectively.

[Web Table 18 about here.]

[Web Table 19 about here.]

[Web Table 20 about here.]

References

Apostol, T. (1974). Mathematical analysis, addison-welsey. *Reading, MA* **2,**.

Cheng, G. and Huang, J. Z. (2010). Bootstrap consistency for general semiparametric m-estimation. *The Annals of Statistics* **38,** 2884–2915.

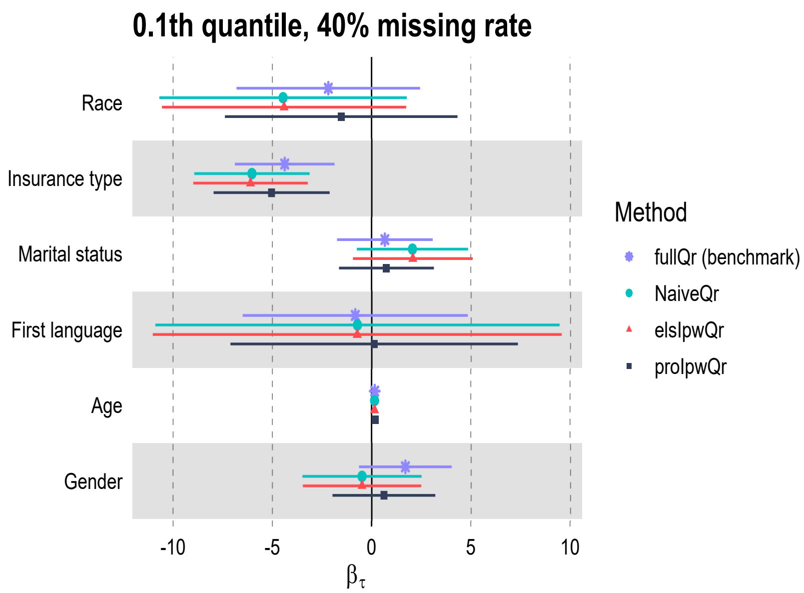
Engle, R. F. (1994). *Handbook of econometrics: volume 4*. Number 330.015195 E53 v. 4.

He, X. and Shao, Q.-M. (2000). On parameters of increasing dimensions. *Journal of multivariate analysis* **73,** 120–135.

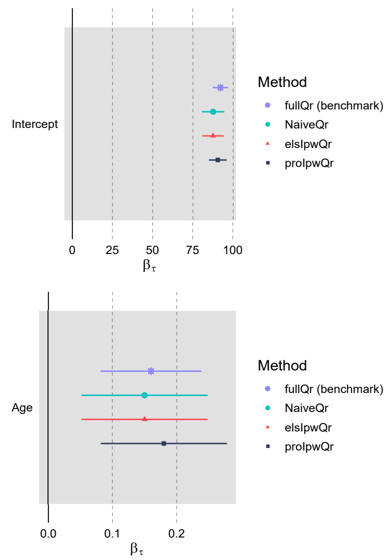
Præstgaard, J. and Wellner, J. A. (1993). Exchangeably weighted bootstraps of the general empirical process. *The Annals of Probability* pages 2053–2086.

Van der Vaart, A. W. (2000). *Asymptotic statistics*, volume 3. Cambridge university press.

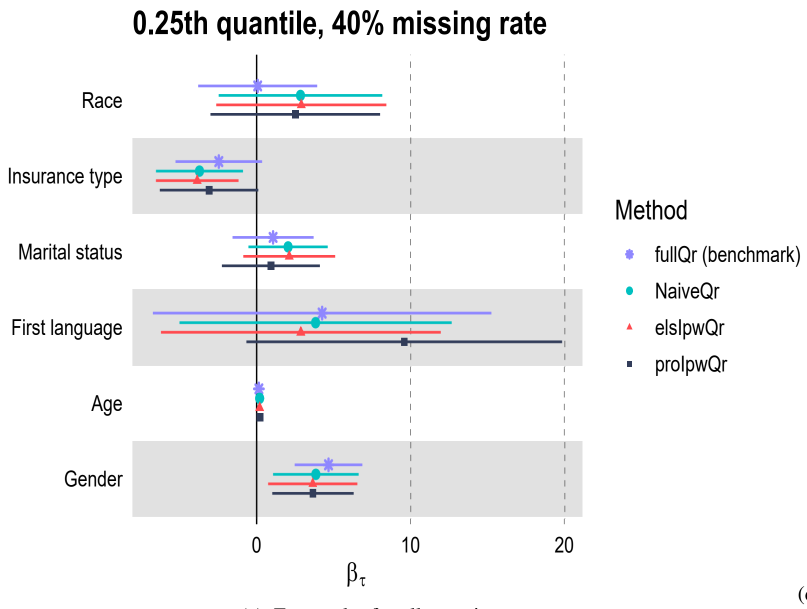
Received December 2021. *Revised February* 2021. *Accepted March* 2022.



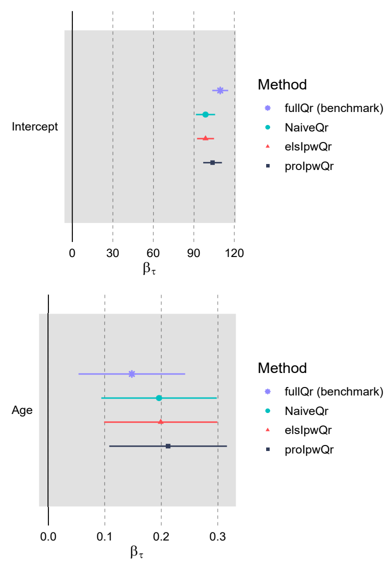
(a) Forest plot for all covariates



(b) Forest plot for intercept term and variable "Age"

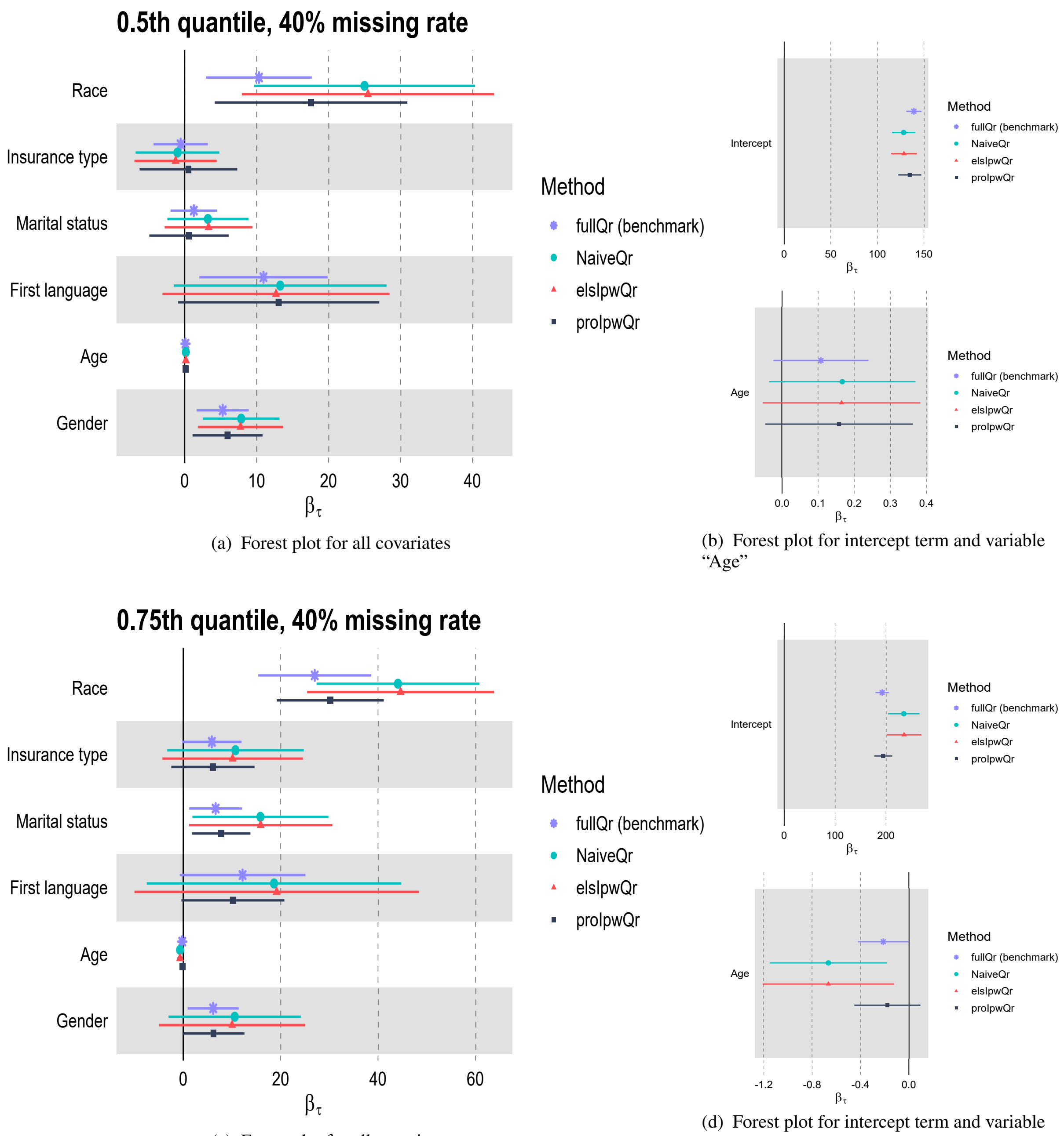


(c) Forest plot for all covariates



(d) Forest plot for intercept term and variable "Age"

Figure 1: The forest plots for quantile regression coefficient estimates at 0.1th and 0.25th quantiles and 40% missing rate. The points with different shapes represent the quantile regression coefficient estimates from different methods. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.



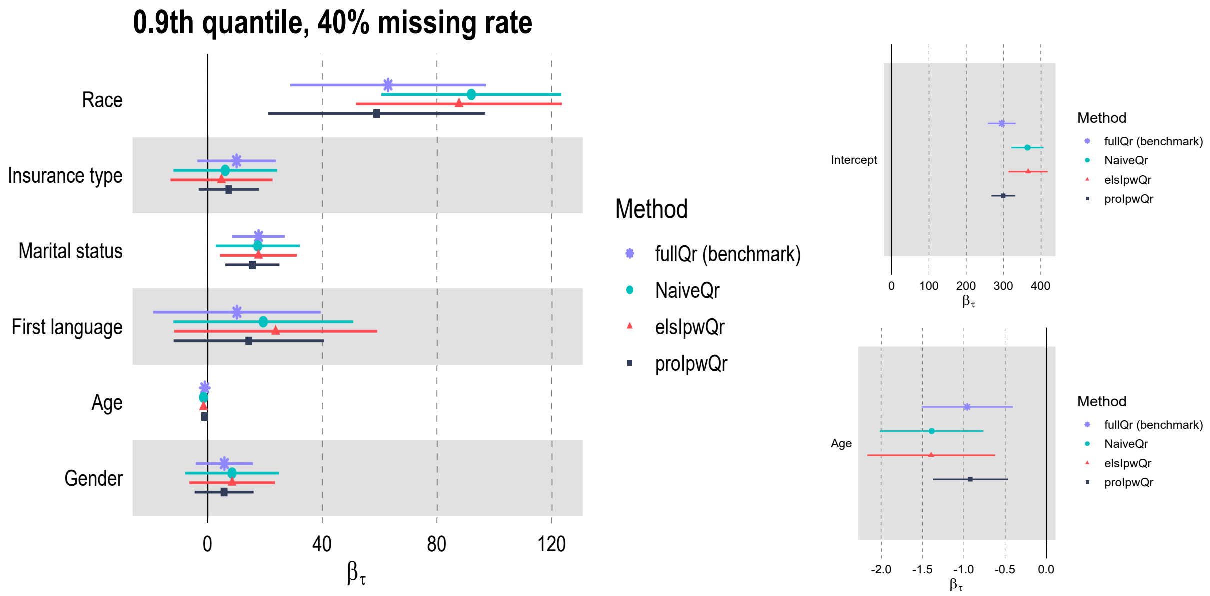
(a) Forest plot for all covariates

(b) Forest plot for intercept term and variable "Age"

(c) Forest plot for all covariates

(d) Forest plot for intercept term and variable "Age"

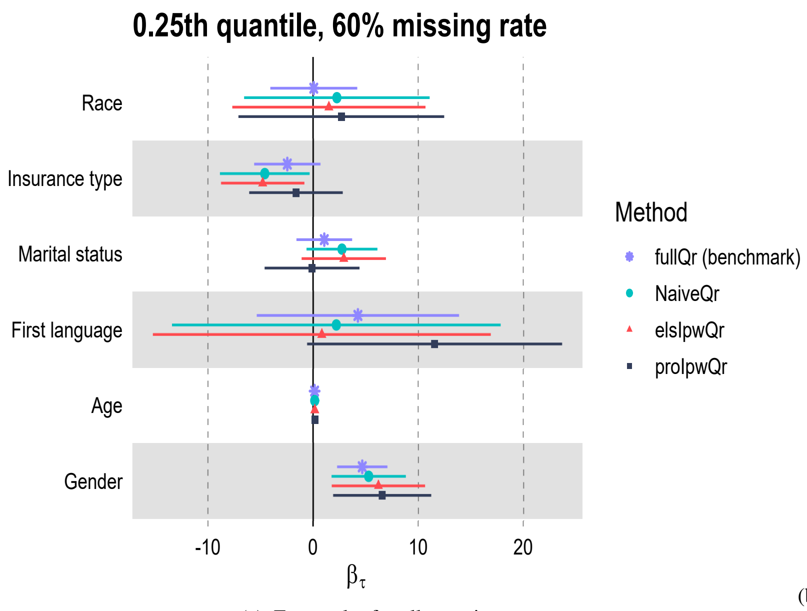
Figure 2: The forest plots for quantile regression coefficient estimates at 0.5th and 0.75th quantiles and 40% missing rate. The points with different shapes represent the quantile regression coefficient estimates from different methods. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.



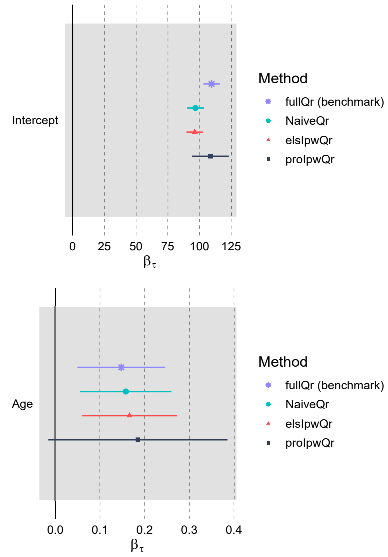
(a) Forest plot for all covariates

(b) Forest plot for intercept term and variable "Age"

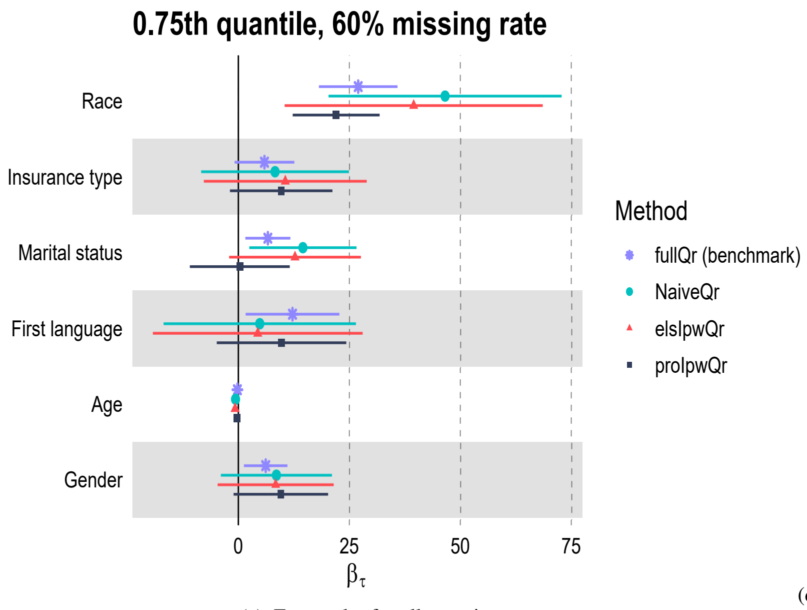
Figure 3: The forest plots for quantile regression coefficient estimates at 0.9th quantiles and 40% missing rate. The points with different shapes represent the quantile regression coefficient estimates from different methods. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.



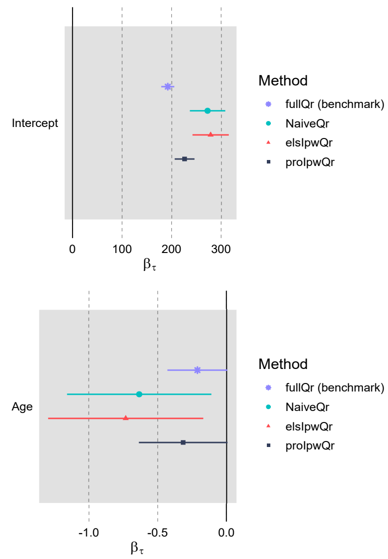
(a) Forest plot for all covariates



(b) Forest plot for intercept term and variable "Age"

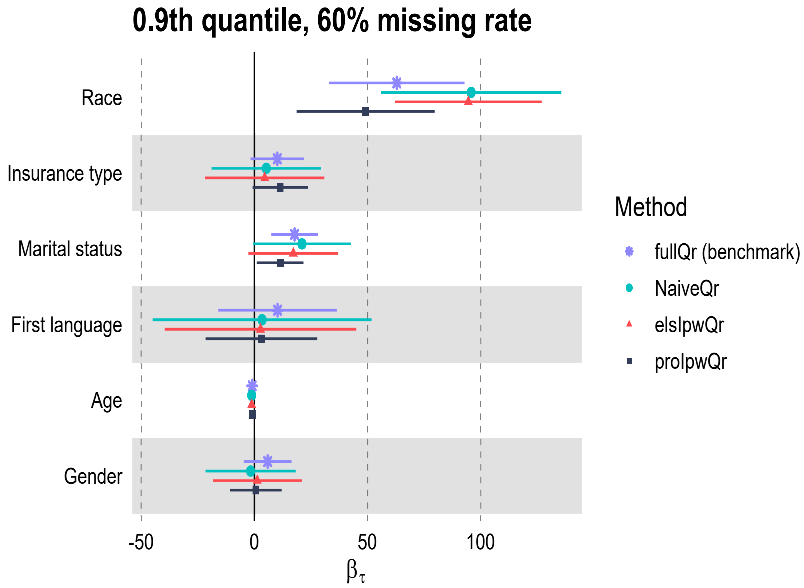


(c) Forest plot for all covariates

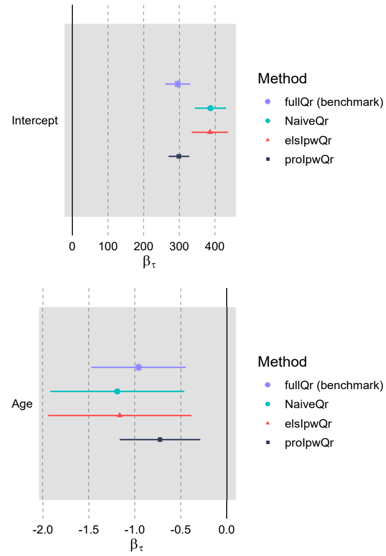


(d) Forest plot for intercept term and variable "Age"

Figure 4: The forest plots for quantile regression coefficient estimates at 0.25th and 0.75th quantiles and 60% missing rate. The points with different shapes represent the quantile regression coefficient estimates from different methods. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.



(a) Forest plot for all covariates



(b) Forest plot for intercept term and variable "Age"

Figure 5: The forest plots for quantile regression coefficient estimates at 0.9th quantiles and 60% missing rate. The points with different shapes represent the quantile regression coefficient estimates from different methods. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.

Table 1: (Setting 1) The mean biases, root of mean square errors (RMS), coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) and confidence interval (CI) lengths of different estimators for quantile regression coefficient with τ = $(0.1, 0.25, 0.5, 0.75, 0.9)$, n = 500 and 20% missing in y_i

	τ = 0.1				τ = 0.25				τ = 0.5				τ = 0.75				τ = 0.9			
Method	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
$\beta_{j,\tau}^{true}$	0.680	-2.320	1.872	0.500	0.831	-2.169	1.933	0.500	1.000	-2.000	2.000	0.500	1.169	-1.831	2.067	0.500	1.320	-1.680	2.128	0.500
Mean Biases																				
fullQr	0.001	0.004	-0.001	0.002	0.003	0.003	-0.001	-0.003	0.005	0.000	0.000	-0.005	0.001	-0.002	0.002	-0.004	0.007	0.000	-0.003	-0.002
NaiveQr	0.068	0.050	-0.025	-0.001	0.070	0.046	-0.024	-0.007	0.072	0.037	-0.023	-0.009	0.068	0.033	-0.021	-0.010	0.069	0.031	-0.024	-0.008
elmIpwQr	-0.059	-0.084	-0.014	0.059	-0.086	-0.074	0.003	0.062	-0.065	-0.040	0.004	0.052	-0.027	-0.002	-0.003	0.037	-0.012	0.014	-0.002	0.025
elsIpwQr	-0.013	-0.021	-0.010	0.020	-0.015	-0.008	-0.003	0.015	-0.004	-0.002	-0.003	0.010	0.007	0.011	-0.006	0.008	0.009	0.016	-0.006	0.007
swelQr	-0.029	-0.010	-0.015	0.035	-0.021	-0.004	-0.006	0.021	-0.001	0.000	-0.007	0.015	0.014	0.008	-0.009	0.013	0.025	0.012	-0.012	0.015
proIpwQr	0.065	0.044	-0.026	0.000	0.059	0.039	-0.022	-0.004	0.057	0.029	-0.019	-0.008	0.049	0.026	-0.014	-0.009	0.047	0.022	-0.016	-0.006
RMS																				
fullQr	0.186	0.088	0.092	0.145	0.148	0.073	0.077	0.114	0.131	0.065	0.067	0.099	0.147	0.071	0.073	0.113	0.178	0.091	0.090	0.143
NaiveQr	0.231	0.114	0.113	0.157	0.182	0.094	0.090	0.126	0.169	0.083	0.081	0.112	0.188	0.087	0.090	0.128	0.221	0.106	0.107	0.161
elmIpwQr	0.372	0.266	0.189	0.287	0.360	0.257	0.170	0.268	0.331	0.221	0.175	0.256	0.285	0.200	0.153	0.232	0.278	0.191	0.176	0.234
elsIpwQr	0.300	0.173	0.142	0.190	0.252	0.134	0.113	0.140	0.206	0.106	0.092	0.115	0.207	0.108	0.097	0.127	0.240	0.131	0.114	0.170
swelQr	0.284	0.159	0.124	0.165	0.233	0.128	0.106	0.124	0.189	0.098	0.085	0.105	0.189	0.100	0.089	0.116	0.215	0.118	0.105	0.148
proIpwQr	0.245	0.114	0.118	0.164	0.191	0.095	0.096	0.130	0.171	0.084	0.084	0.114	0.187	0.086	0.092	0.129	0.221	0.106	0.109	0.164
95ECP																				
fullQr	0.936	0.958	0.951	0.947	0.949	0.939	0.947	0.959	0.958	0.952	0.961	0.959	0.942	0.948	0.958	0.951	0.950	0.949	0.962	0.950
NaiveQr	0.925	0.916	0.932	0.960	0.939	0.916	0.933	0.955	0.932	0.925	0.951	0.960	0.937	0.931	0.944	0.958	0.948	0.946	0.946	0.949
elmIpwQr	0.885	0.800	0.899	0.903	0.893	0.801	0.910	0.916	0.912	0.860	0.932	0.944	0.919	0.874	0.950	0.942	0.915	0.894	0.934	0.938
elsIpwQr	0.901	0.883	0.934	0.955	0.918	0.881	0.918	0.966	0.898	0.904	0.939	0.967	0.933	0.921	0.938	0.963	0.932	0.921	0.948	0.946
swelQr	0.893	0.923	0.917	0.957	0.936	0.942	0.937	0.960	0.949	0.967	0.951	0.971	0.946	0.964	0.939	0.965	0.937	0.936	0.940	0.938
proIpwQr	0.916	0.927	0.933	0.953	0.942	0.922	0.938	0.955	0.933	0.939	0.945	0.965	0.947	0.941	0.948	0.956	0.948	0.947	0.947	0.950
CI lengths																				
fullQr	0.743	0.366	0.381	0.587	0.585	0.287	0.303	0.469	0.539	0.266	0.279	0.431	0.588	0.290	0.306	0.467	0.738	0.366	0.380	0.591
NaiveQr	0.890	0.421	0.446	0.666	0.704	0.333	0.355	0.528	0.649	0.305	0.325	0.486	0.708	0.333	0.356	0.529	0.884	0.420	0.440	0.663
elmIpwQr	1.068	0.524	0.537	0.805	0.906	0.456	0.464	0.686	0.827	0.410	0.416	0.615	0.829	0.408	0.421	0.623	0.930	0.455	0.477	0.722
elsIpwQr	1.007	0.496	0.504	0.761	0.797	0.388	0.400	0.599	0.697	0.336	0.349	0.521	0.730	0.349	0.367	0.548	0.878	0.428	0.445	0.673
swelQr	1.074	0.635	0.520	0.770	0.958	0.533	0.429	0.597	0.826	0.434	0.369	0.507	0.800	0.419	0.374	0.520	0.854	0.463	0.425	0.610
proIpwQr	0.908	0.438	0.455	0.685	0.762	0.359	0.375	0.552	0.712	0.331	0.347	0.508	0.748	0.351	0.370	0.546	0.887	0.425	0.443	0.674

Table 2: (Setting 1) The mean biases, root of mean square errors (RMS), coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) and confidence interval(CI) lengths of different estimators for quantile regression coefficient with $\tau = (0.1, 0.25, 0.5, 0.75, 0.9)$, $n = 1000$ and 20% missing in y_i

	$\tau = 0.1$				$\tau = 0.25$				$\tau = 0.5$				$\tau = 0.75$				$\tau = 0.9$			
Method	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
$\beta_{j,\tau}^{true}$	0.680	-2.320	1.872	0.500	0.831	-2.169	1.933	0.500	1.000	-2.000	2.000	0.500	1.169	-1.831	2.067	0.500	1.320	-1.680	2.128	0.500
									Mean Biases											
fullQr	-0.002	0.002	0.001	-0.001	0.003	0.001	-0.002	-0.001	0.003	-0.001	-0.002	0.001	0.002	-0.001	-0.002	0.001	0.004	-0.002	-0.003	0.003
NaiveQr	0.062	0.049	-0.020	-0.007	0.063	0.045	-0.022	-0.005	0.061	0.040	-0.021	-0.003	0.062	0.037	-0.023	-0.002	0.063	0.032	-0.023	-0.002
elmIpwQr	-0.066	-0.099	-0.001	0.039	-0.095	-0.086	0.010	0.053	-0.082	-0.052	0.009	0.054	-0.054	-0.021	0.002	0.046	-0.031	-0.006	-0.001	0.033
elsIpwQr	-0.017	-0.015	0.000	0.010	-0.013	-0.006	-0.001	0.009	-0.014	-0.001	0.001	0.010	-0.003	0.006	-0.003	0.008	-0.001	0.009	-0.004	0.010
swelQr	-0.028	-0.009	0.000	0.012	-0.018	-0.003	-0.001	0.011	-0.011	0.000	0.000	0.010	0.003	0.004	-0.004	0.009	0.012	0.005	-0.006	0.009
proIpwQr	0.055	0.041	-0.018	-0.007	0.052	0.037	-0.018	-0.005	0.046	0.032	-0.016	-0.004	0.046	0.029	-0.017	-0.002	0.043	0.025	-0.016	-0.001
									RMS											
fullQr	0.130	0.066	0.067	0.102	0.101	0.051	0.051	0.080	0.092	0.046	0.048	0.075	0.099	0.049	0.051	0.081	0.125	0.063	0.065	0.102
NaiveQr	0.168	0.089	0.079	0.112	0.133	0.075	0.062	0.090	0.124	0.066	0.058	0.083	0.134	0.067	0.064	0.093	0.164	0.079	0.079	0.112
elmIpwQr	0.318	0.256	0.141	0.224	0.334	0.256	0.133	0.216	0.329	0.201	0.128	0.218	0.293	0.178	0.133	0.223	0.227	0.148	0.113	0.169
elsIpwQr	0.209	0.125	0.095	0.128	0.165	0.093	0.076	0.095	0.150	0.074	0.067	0.084	0.140	0.071	0.068	0.091	0.162	0.082	0.078	0.110
swelQr	0.196	0.114	0.088	0.120	0.158	0.090	0.069	0.089	0.141	0.071	0.063	0.080	0.132	0.067	0.064	0.085	0.152	0.078	0.073	0.103
proIpwQr	0.170	0.087	0.082	0.117	0.135	0.072	0.064	0.092	0.123	0.063	0.059	0.084	0.130	0.065	0.064	0.093	0.156	0.076	0.077	0.112
									95ECP											
fullQr	0.937	0.946	0.945	0.946	0.951	0.944	0.950	0.960	0.949	0.952	0.954	0.945	0.958	0.951	0.957	0.951	0.945	0.953	0.946	0.946
NaiveQr	0.915	0.878	0.936	0.942	0.914	0.871	0.936	0.948	0.930	0.891	0.945	0.958	0.938	0.911	0.950	0.943	0.935	0.933	0.935	0.954
elmIpwQr	0.889	0.766	0.901	0.919	0.863	0.764	0.924	0.936	0.881	0.799	0.933	0.940	0.896	0.854	0.942	0.951	0.920	0.899	0.952	0.946
elsIpwQr	0.896	0.879	0.920	0.960	0.899	0.859	0.927	0.962	0.896	0.880	0.925	0.963	0.921	0.902	0.936	0.961	0.925	0.920	0.952	0.952
swelQr	0.902	0.919	0.923	0.945	0.928	0.944	0.940	0.963	0.947	0.962	0.947	0.957	0.946	0.961	0.954	0.957	0.928	0.944	0.937	0.939
proIpwQr	0.913	0.922	0.931	0.951	0.916	0.934	0.939	0.958	0.931	0.914	0.945	0.956	0.944	0.922	0.950	0.947	0.944	0.940	0.944	0.952
									CI lengths											
fullQr	0.516	0.256	0.266	0.412	0.406	0.204	0.213	0.332	0.378	0.187	0.196	0.303	0.409	0.203	0.213	0.330	0.514	0.255	0.267	0.412
NaiveQr	0.608	0.292	0.307	0.461	0.485	0.232	0.245	0.372	0.454	0.215	0.228	0.340	0.491	0.233	0.248	0.368	0.621	0.293	0.311	0.463
elmIpwQr	0.804	0.401	0.397	0.617	0.673	0.340	0.336	0.502	0.592	0.293	0.293	0.437	0.580	0.286	0.290	0.436	0.654	0.320	0.332	0.501
elsIpwQr	0.698	0.341	0.347	0.531	0.539	0.262	0.273	0.412	0.473	0.226	0.239	0.357	0.495	0.236	0.251	0.373	0.603	0.289	0.305	0.458
swelQr	0.718	0.439	0.336	0.502	0.633	0.373	0.293	0.406	0.577	0.300	0.256	0.342	0.552	0.282	0.256	0.351	0.590	0.309	0.287	0.410
proIpwQr	0.633	0.303	0.316	0.480	0.514	0.247	0.256	0.382	0.480	0.227	0.237	0.348	0.507	0.241	0.253	0.374	0.619	0.294	0.309	0.465

Table 3: (Setting 2) The mean biases, root of mean square errors (RMS), coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) and confidence interval(CI) lengths of different estimators for quantile regression coefficient with τ = $(0.1, 0.25, 0.5, 0.75, 0.9)$, n = 500 and 20% missing in y_i

	τ = 0.1				τ = 0.25				τ = 0.5				τ = 0.75				τ = 0.9			
Method	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
$\beta_{j,\tau}^{true}$	0.657	-2.343	1.863	0.500	0.742	-2.258	1.897	0.500	0.889	-2.111	1.956	0.500	1.127	-1.873	2.051	0.500	1.464	-1.536	2.185	0.500
Mean Biases																				
fullQr	0.003	0.003	-0.001	0.000	0.004	0.002	-0.001	-0.002	0.009	0.001	-0.001	-0.005	0.010	-0.002	0.001	-0.008	0.033	-0.002	-0.012	-0.003
NaiveQr	0.032	0.020	-0.011	-0.001	0.056	0.033	-0.018	-0.007	0.112	0.053	-0.037	-0.013	0.206	0.094	-0.068	-0.022	0.361	0.140	-0.125	-0.026
elmIpwQr	0.039	0.025	-0.015	0.001	0.059	0.042	-0.026	0.003	0.122	0.073	-0.050	0.009	0.230	0.143	-0.087	0.008	0.429	0.242	-0.149	-0.014
elsIpwQr	0.013	0.006	-0.008	0.006	0.013	0.008	-0.010	0.008	0.029	0.018	-0.017	0.011	0.078	0.075	-0.044	0.029	0.142	0.180	-0.083	0.080
swelQr	-0.017	0.009	-0.003	0.009	0.001	0.002	-0.007	0.012	0.018	0.005	-0.015	0.020	0.056	0.056	-0.042	0.046	0.068	0.134	-0.070	0.117
proIpwQr	0.010	0.003	-0.005	0.002	0.005	0.002	-0.003	0.000	0.010	-0.004	-0.002	-0.003	0.007	-0.007	0.001	-0.004	-0.001	-0.011	-0.004	0.014
RMS																				
fullQr	0.084	0.040	0.041	0.065	0.101	0.050	0.053	0.078	0.144	0.072	0.073	0.109	0.260	0.125	0.130	0.200	0.483	0.248	0.244	0.385
NaiveQr	0.111	0.053	0.052	0.076	0.135	0.069	0.065	0.093	0.216	0.104	0.099	0.132	0.391	0.183	0.174	0.240	0.707	0.333	0.328	0.459
elmIpwQr	0.135	0.063	0.072	0.088	0.433	0.173	0.329	0.174	0.546	0.273	0.430	0.420	0.663	0.417	0.364	0.458	0.926	0.555	0.417	0.596
elsIpwQr	0.117	0.057	0.055	0.079	0.147	0.077	0.072	0.095	0.232	0.125	0.106	0.128	0.491	0.340	0.233	0.245	0.820	0.779	0.408	0.549
swelQr	0.105	0.052	0.051	0.068	0.132	0.071	0.066	0.083	0.211	0.116	0.100	0.115	0.518	0.367	0.246	0.228	0.736	0.709	0.383	0.505
proIpwQr	0.111	0.049	0.054	0.079	0.127	0.060	0.064	0.094	0.179	0.085	0.090	0.127	0.296	0.143	0.153	0.221	0.523	0.267	0.268	0.408
95ECP																				
fullQr	0.940	0.955	0.952	0.950	0.946	0.944	0.940	0.959	0.956	0.952	0.961	0.960	0.941	0.949	0.955	0.955	0.940	0.946	0.955	0.946
NaiveQr	0.930	0.940	0.941	0.958	0.949	0.916	0.946	0.956	0.918	0.925	0.937	0.954	0.913	0.922	0.934	0.962	0.930	0.939	0.929	0.957
elmIpwQr	0.912	0.916	0.927	0.946	0.899	0.892	0.904	0.943	0.901	0.898	0.932	0.953	0.912	0.902	0.949	0.954	0.922	0.930	0.932	0.950
elsIpwQr	0.914	0.909	0.937	0.953	0.921	0.893	0.919	0.960	0.911	0.908	0.938	0.965	0.926	0.920	0.935	0.967	0.924	0.923	0.939	0.941
swelQr	0.919	0.938	0.918	0.958	0.954	0.945	0.937	0.961	0.946	0.956	0.951	0.978	0.936	0.957	0.926	0.962	0.898	0.918	0.907	0.909
proIpwQr	0.929	0.953	0.932	0.951	0.944	0.947	0.940	0.951	0.954	0.948	0.953	0.965	0.949	0.950	0.942	0.960	0.937	0.947	0.955	0.948
CI lengths																				
fullQr	0.334	0.163	0.172	0.263	0.399	0.197	0.207	0.319	0.592	0.293	0.306	0.475	1.035	0.516	0.536	0.829	1.973	1.002	1.016	1.603
NaiveQr	0.425	0.200	0.211	0.314	0.518	0.245	0.258	0.382	0.767	0.364	0.379	0.567	1.351	0.642	0.672	0.996	2.575	1.245	1.263	1.922
elmIpwQr	0.453	0.212	0.242	0.334	0.599	0.280	0.308	0.440	0.900	0.416	0.446	0.646	1.544	0.730	0.766	1.112	2.901	1.464	1.421	2.172
elsIpwQr	0.422	0.201	0.211	0.319	0.529	0.249	0.265	0.392	0.788	0.393	0.393	0.584	1.414	0.774	0.707	1.058	2.597	1.480	1.315	2.057
swelQr	0.400	0.206	0.189	0.276	0.565	0.296	0.262	0.359	0.929	0.504	0.420	0.563	1.719	1.065	0.828	1.132	2.764	1.840	1.439	2.203
proIpwQr	0.443	0.208	0.225	0.328	0.536	0.250	0.268	0.393	0.763	0.357	0.379	0.558	1.230	0.596	0.626	0.934	2.196	1.087	1.124	1.732

Table 4: (Setting 2) The mean biases, root of mean square errors (RMS), coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) and confidence interval(CI) lengths of different estimators for quantile regression coefficient with τ = $(0.1, 0.25, 0.5, 0.75, 0.9)$, n = 1000 and 20% missing in y_i

	τ = 0.1				τ = 0.25				τ = 0.5				τ = 0.75				τ = 0.9			
Method	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
$\beta_{j,\tau}^{true}$	0.657	-2.343	1.863	0.500	0.742	-2.258	1.897	0.500	0.889	-2.111	1.956	0.500	1.127	-1.873	2.051	0.500	1.464	-1.536	2.185	0.500
Mean Biases																				
fullQr	0.001	0.001	0.000	0.000	0.003	0.000	-0.001	-0.001	0.005	-0.001	-0.003	0.001	0.008	0.000	-0.004	0.001	0.017	-0.008	-0.009	0.007
NaiveQr	0.027	0.020	-0.008	-0.003	0.050	0.031	-0.017	-0.004	0.097	0.056	-0.034	-0.006	0.188	0.098	-0.067	-0.010	0.330	0.146	-0.120	-0.013
elmIpwQr	0.027	0.023	-0.009	-0.002	0.038	0.034	-0.013	0.001	0.091	0.066	-0.033	0.001	0.209	0.129	-0.076	-0.004	0.401	0.218	-0.143	-0.013
elsIpwQr	0.004	0.003	-0.003	0.002	0.006	0.005	-0.006	0.006	0.010	0.013	-0.010	0.013	0.048	0.047	-0.029	0.020	0.114	0.114	-0.059	0.038
swelQr	-0.012	0.005	0.000	0.003	0.000	0.002	-0.004	0.007	0.007	0.006	-0.008	0.013	0.039	0.034	-0.025	0.024	0.092	0.082	-0.054	0.048
proIpwQr	0.002	0.002	0.000	-0.001	-0.002	-0.001	-0.001	0.002	-0.009	-0.004	0.002	0.004	-0.006	-0.003	0.001	0.004	-0.015	-0.010	0.003	0.011
RMS																				
fullQr	0.058	0.029	0.029	0.046	0.069	0.035	0.035	0.055	0.102	0.050	0.053	0.082	0.176	0.088	0.091	0.142	0.338	0.173	0.176	0.277
NaiveQr	0.079	0.041	0.037	0.053	0.100	0.055	0.046	0.064	0.160	0.085	0.072	0.098	0.294	0.145	0.131	0.175	0.560	0.260	0.255	0.329
elmIpwQr	0.094	0.045	0.042	0.056	0.138	0.063	0.059	0.074	0.197	0.100	0.083	0.111	0.344	0.178	0.146	0.190	0.674	0.331	0.290	0.356
elsIpwQr	0.080	0.041	0.038	0.054	0.103	0.056	0.048	0.064	0.170	0.090	0.078	0.094	0.365	0.241	0.172	0.171	0.638	0.520	0.300	0.332
swelQr	0.074	0.038	0.035	0.047	0.095	0.051	0.045	0.059	0.159	0.084	0.074	0.089	0.355	0.220	0.169	0.163	0.605	0.413	0.289	0.318
proIpwQr	0.075	0.036	0.037	0.055	0.087	0.044	0.043	0.064	0.124	0.061	0.062	0.093	0.204	0.101	0.104	0.158	0.376	0.184	0.193	0.289
95ECP																				
fullQr	0.936	0.949	0.953	0.945	0.948	0.946	0.942	0.946	0.945	0.951	0.958	0.946	0.954	0.954	0.955	0.944	0.949	0.952	0.934	0.949
NaiveQr	0.926	0.899	0.942	0.948	0.907	0.887	0.937	0.954	0.909	0.884	0.938	0.951	0.896	0.873	0.918	0.948	0.892	0.914	0.906	0.948
elmIpwQr	0.872	0.878	0.906	0.937	0.871	0.840	0.915	0.950	0.878	0.825	0.915	0.951	0.882	0.828	0.917	0.945	0.880	0.891	0.922	0.954
elsIpwQr	0.908	0.900	0.937	0.949	0.908	0.885	0.932	0.961	0.893	0.863	0.925	0.961	0.906	0.900	0.924	0.953	0.919	0.914	0.927	0.956
swelQr	0.919	0.942	0.938	0.946	0.937	0.958	0.946	0.961	0.943	0.953	0.940	0.964	0.937	0.966	0.942	0.957	0.895	0.938	0.900	0.922
proIpwQr	0.931	0.932	0.937	0.947	0.939	0.949	0.942	0.952	0.955	0.945	0.955	0.954	0.943	0.953	0.952	0.943	0.945	0.956	0.945	0.953
CI lengths																				
fullQr	0.229	0.113	0.119	0.183	0.277	0.139	0.145	0.225	0.417	0.205	0.216	0.332	0.722	0.360	0.376	0.583	1.392	0.700	0.723	1.122
NaiveQr	0.290	0.140	0.145	0.217	0.358	0.173	0.179	0.270	0.541	0.257	0.267	0.399	0.943	0.446	0.465	0.696	1.808	0.865	0.890	1.341
elmIpwQr	0.296	0.140	0.148	0.220	0.388	0.179	0.193	0.279	0.611	0.277	0.298	0.424	1.062	0.491	0.520	0.750	2.066	0.979	1.003	1.478
elsIpwQr	0.295	0.141	0.148	0.223	0.360	0.173	0.182	0.272	0.537	0.259	0.270	0.397	0.944	0.490	0.475	0.698	1.794	0.960	0.902	1.386
swelQr	0.290	0.150	0.138	0.194	0.412	0.220	0.188	0.252	0.713	0.382	0.321	0.388	1.486	0.848	0.661	0.750	2.136	1.301	1.072	1.365
proIpwQr	0.294	0.141	0.148	0.222	0.356	0.171	0.179	0.268	0.516	0.246	0.258	0.384	0.847	0.409	0.432	0.643	1.538	0.755	0.787	1.196

Table 5: (Setting 3) The mean biases, root of mean square errors (RMS), coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) and confidence interval(CI) lengths of different estimators for quantile regression coefficient with τ = $(0.1, 0.25, 0.5, 0.75, 0.9)$, n = 500 and 20% missing in y_i

	τ = 0.1				τ = 0.25				τ = 0.5				τ = 0.75				τ = 0.9			
Method	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
$\beta_{j,\tau}^{true}$	0.507	-2.493	1.803	0.500	0.635	-2.365	1.854	0.500	0.871	-2.129	1.948	0.500	1.226	-1.774	2.090	0.500	1.664	-1.336	2.265	0.500
									Mean Biases											
fullQr	-0.001	0.001	-0.001	0.007	-0.003	0.003	-0.002	0.010	-0.001	0.004	-0.004	0.009	0.006	0.002	-0.004	0.008	-0.009	0.003	-0.003	0.017
NaiveQr	0.056	0.042	-0.018	0.000	0.114	0.078	-0.038	-0.006	0.211	0.126	-0.074	-0.011	0.342	0.174	-0.116	-0.028	0.417	0.196	-0.151	-0.015
elmIpwQr	0.056	0.046	-0.021	0.006	0.061	0.070	-0.020	0.012	0.140	0.123	-0.046	0.010	0.304	0.202	-0.102	-0.003	0.414	0.255	-0.147	0.011
elsIpwQr	0.018	0.009	-0.016	0.016	0.011	0.014	-0.018	0.027	0.031	0.045	-0.032	0.040	0.106	0.127	-0.064	0.060	0.117	0.227	-0.074	0.099
swelQr	-0.006	0.009	-0.010	0.020	0.001	0.007	-0.016	0.031	0.014	0.030	-0.029	0.051	0.089	0.101	-0.065	0.077	0.104	0.189	-0.085	0.133
proIpwQr	0.024	0.015	-0.010	0.004	0.027	0.019	-0.010	0.002	0.036	0.026	-0.012	0.000	0.063	0.029	-0.019	-0.006	0.035	0.027	-0.014	0.008
									RMS											
fullQr	0.113	0.057	0.059	0.089	0.161	0.077	0.081	0.124	0.231	0.114	0.117	0.183	0.357	0.178	0.186	0.286	0.603	0.302	0.304	0.478
NaiveQr	0.158	0.085	0.080	0.114	0.238	0.126	0.111	0.152	0.363	0.191	0.166	0.219	0.558	0.281	0.253	0.337	0.871	0.406	0.403	0.575
elmIpwQr	0.187	0.094	0.091	0.122	0.310	0.146	0.136	0.178	0.450	0.224	0.190	0.264	0.620	0.326	0.271	0.374	0.935	0.474	0.419	0.630
elsIpwQr	0.170	0.090	0.084	0.113	0.260	0.142	0.118	0.150	0.445	0.239	0.198	0.218	0.800	0.471	0.374	0.364	1.010	0.780	0.491	0.636
swelQr	0.154	0.083	0.076	0.100	0.245	0.134	0.113	0.138	0.437	0.230	0.199	0.204	0.821	0.466	0.384	0.358	1.054	0.761	0.508	0.636
proIpwQr	0.149	0.076	0.080	0.113	0.212	0.101	0.104	0.151	0.290	0.142	0.148	0.215	0.420	0.211	0.217	0.310	0.667	0.326	0.338	0.521
									95ECP											
fullQr	0.943	0.947	0.936	0.952	0.935	0.944	0.954	0.957	0.931	0.956	0.948	0.958	0.955	0.954	0.944	0.952	0.937	0.944	0.946	0.944
NaiveQr	0.931	0.916	0.932	0.951	0.920	0.883	0.928	0.943	0.903	0.870	0.918	0.950	0.895	0.883	0.926	0.949	0.913	0.909	0.916	0.937
elmIpwQr	0.899	0.880	0.896	0.949	0.849	0.841	0.882	0.921	0.893	0.836	0.921	0.929	0.899	0.864	0.925	0.936	0.912	0.889	0.932	0.937
elsIpwQr	0.922	0.882	0.915	0.954	0.880	0.846	0.927	0.952	0.876	0.838	0.916	0.947	0.889	0.889	0.909	0.950	0.898	0.876	0.915	0.947
swelQr	0.922	0.925	0.919	0.941	0.928	0.937	0.945	0.953	0.921	0.951	0.926	0.956	0.920	0.951	0.918	0.943	0.844	0.904	0.872	0.876
proIpwQr	0.938	0.937	0.927	0.943	0.946	0.932	0.944	0.948	0.955	0.947	0.939	0.955	0.949	0.950	0.946	0.957	0.937	0.940	0.941	0.940
									CI lengths											
fullQr	0.458	0.229	0.238	0.369	0.625	0.309	0.325	0.502	0.929	0.463	0.480	0.744	1.462	0.722	0.750	1.170	2.362	1.187	1.223	1.913
NaiveQr	0.610	0.301	0.302	0.456	0.833	0.399	0.415	0.616	1.229	0.578	0.602	0.900	1.861	0.881	0.924	1.371	2.963	1.406	1.469	2.223
elmIpwQr	0.613	0.299	0.305	0.464	0.886	0.412	0.438	0.642	1.405	0.630	0.676	0.964	2.086	0.967	1.020	1.489	3.200	1.533	1.588	2.410
elsIpwQr	0.588	0.288	0.296	0.453	0.832	0.405	0.418	0.627	1.278	0.623	0.647	0.944	2.013	1.029	0.997	1.518	3.015	1.533	1.511	2.356
swelQr	0.615	0.341	0.284	0.410	1.068	0.591	0.471	0.644	1.905	1.017	0.868	1.042	2.985	1.699	1.370	1.832	3.422	2.220	1.813	2.584
proIpwQr	0.599	0.295	0.303	0.454	0.830	0.397	0.415	0.617	1.203	0.572	0.596	0.883	1.758	0.839	0.877	1.321	2.655	1.281	1.337	2.059

Table 6: (Setting 3) The mean biases, root of mean square errors (RMS), coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) and confidence interval(CI) lengths of different estimators for quantile regression coefficient with $\tau = (0.1, 0.25, 0.5, 0.75, 0.9)$, $n = 1000$ and 20% missing in y_i

	$\tau = 0.1$				$\tau = 0.25$				$\tau = 0.5$				$\tau = 0.75$				$\tau = 0.9$			
Method	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
$\beta_{j,\tau}^{true}$	0.507	-2.493	1.803	0.500	0.635	-2.365	1.854	0.500	0.871	-2.129	1.948	0.500	1.226	-1.774	2.090	0.500	1.664	-1.336	2.265	0.500
									Mean Biases											
fullQr	-0.001	0.003	0.001	0.000	0.001	0.002	0.001	0.000	0.005	0.006	-0.002	-0.001	0.011	0.008	-0.006	0.001	0.009	0.009	-0.005	0.004
NaiveQr	0.062	0.042	-0.018	-0.009	0.126	0.075	-0.040	-0.013	0.217	0.130	-0.074	-0.017	0.342	0.182	-0.120	-0.024	0.438	0.206	-0.151	-0.032
elmIpwQr	0.058	0.044	-0.017	-0.007	0.087	0.072	-0.027	0.000	0.178	0.135	-0.061	0.000	0.330	0.212	-0.119	-0.006	0.473	0.258	-0.165	-0.021
elsIpwQr	0.017	0.007	-0.008	0.000	0.019	0.012	-0.013	0.010	0.031	0.033	-0.021	0.016	0.073	0.082	-0.042	0.024	0.124	0.135	-0.066	0.042
swelQr	0.005	0.007	-0.005	0.003	0.010	0.006	-0.010	0.012	0.015	0.021	-0.017	0.023	0.044	0.064	-0.033	0.034	0.083	0.116	-0.056	0.058
proIpwQr	0.022	0.013	-0.005	-0.007	0.033	0.015	-0.009	-0.006	0.044	0.027	-0.014	-0.006	0.060	0.035	-0.024	-0.002	0.066	0.029	-0.023	-0.003
									RMS											
fullQr	0.083	0.041	0.043	0.066	0.112	0.057	0.056	0.087	0.165	0.081	0.083	0.133	0.258	0.124	0.129	0.203	0.414	0.201	0.214	0.335
NaiveQr	0.127	0.068	0.057	0.078	0.195	0.104	0.083	0.107	0.309	0.164	0.128	0.163	0.467	0.238	0.198	0.237	0.681	0.320	0.300	0.391
elmIpwQr	0.156	0.073	0.066	0.086	0.273	0.117	0.107	0.130	0.376	0.187	0.148	0.185	0.518	0.275	0.214	0.262	0.752	0.372	0.321	0.421
elsIpwQr	0.130	0.063	0.057	0.078	0.196	0.102	0.083	0.106	0.319	0.167	0.133	0.158	0.510	0.309	0.218	0.230	0.709	0.446	0.328	0.423
swelQr	0.119	0.057	0.052	0.070	0.183	0.096	0.079	0.101	0.301	0.157	0.127	0.152	0.472	0.293	0.205	0.222	0.662	0.465	0.316	0.385
proIpwQr	0.116	0.053	0.055	0.078	0.153	0.073	0.073	0.106	0.215	0.102	0.101	0.155	0.312	0.146	0.151	0.222	0.470	0.221	0.238	0.357
									95ECP											
fullQr	0.937	0.944	0.934	0.937	0.938	0.944	0.954	0.957	0.946	0.953	0.954	0.937	0.940	0.950	0.948	0.953	0.946	0.948	0.956	0.945
NaiveQr	0.896	0.869	0.931	0.946	0.866	0.832	0.914	0.957	0.822	0.770	0.895	0.937	0.837	0.787	0.891	0.959	0.876	0.883	0.900	0.951
elmIpwQr	0.838	0.843	0.895	0.929	0.781	0.772	0.877	0.945	0.791	0.711	0.888	0.942	0.823	0.759	0.889	0.954	0.857	0.849	0.915	0.953
elsIpwQr	0.886	0.887	0.932	0.948	0.862	0.826	0.918	0.962	0.820	0.804	0.900	0.950	0.850	0.843	0.901	0.955	0.897	0.886	0.904	0.954
swelQr	0.915	0.947	0.935	0.948	0.932	0.947	0.939	0.966	0.925	0.959	0.936	0.953	0.925	0.959	0.921	0.943	0.902	0.945	0.895	0.927
proIpwQr	0.915	0.946	0.939	0.945	0.930	0.943	0.942	0.953	0.936	0.944	0.954	0.947	0.938	0.953	0.946	0.954	0.941	0.948	0.939	0.949
									CI lengths											
fullQr	0.321	0.160	0.168	0.257	0.439	0.219	0.229	0.356	0.649	0.323	0.334	0.521	1.013	0.505	0.528	0.819	1.674	0.837	0.867	1.353
NaiveQr	0.427	0.208	0.214	0.320	0.589	0.284	0.294	0.438	0.842	0.404	0.417	0.624	1.315	0.625	0.652	0.978	2.105	1.002	1.040	1.580
elmIpwQr	0.434	0.210	0.218	0.326	0.645	0.297	0.317	0.458	0.988	0.445	0.473	0.675	1.481	0.676	0.720	1.046	2.273	1.070	1.115	1.679
elsIpwQr	0.418	0.204	0.212	0.320	0.583	0.279	0.292	0.438	0.841	0.407	0.423	0.625	1.306	0.639	0.658	0.981	2.050	1.023	1.038	1.601
swelQr	0.444	0.250	0.207	0.286	0.764	0.426	0.332	0.434	1.258	0.694	0.542	0.657	2.159	1.222	0.962	1.106	2.527	1.587	1.227	1.884
proIpwQr	0.421	0.204	0.212	0.319	0.585	0.279	0.292	0.434	0.821	0.394	0.409	0.608	1.217	0.583	0.612	0.919	1.871	0.903	0.949	1.452

Table 7: (Setting 1) The mean biases, root of mean square errors (RMS), coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) and confidence interval(CI) lengths of different estimators for quantile regression coefficient with τ = $(0.1, 0.25, 0.5, 0.75, 0.9)$, n = 500 and 40% missing in y_i

	τ = 0.1				τ = 0.25				τ = 0.5				τ = 0.75				τ = 0.9			
Method	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
$\beta_{j,\tau}^{true}$	0.680	-2.320	1.872	0.500	0.831	-2.169	1.933	0.500	1.000	-2.000	2.000	0.500	1.169	-1.831	2.067	0.500	1.320	-1.680	2.128	0.500
Mean Biases																				
fullQr	0.001	0.004	-0.001	0.002	0.003	0.003	-0.001	-0.003	0.005	0.000	0.000	-0.005	0.001	-0.002	0.002	-0.004	0.007	0.000	-0.003	-0.002
NaiveQr	0.092	0.081	-0.024	-0.008	0.100	0.074	-0.029	-0.008	0.105	0.066	-0.032	-0.008	0.104	0.060	-0.031	-0.010	0.107	0.057	-0.034	-0.012
elmIpwQr	0.030	-0.033	-0.040	0.019	-0.020	-0.063	-0.022	0.046	-0.065	-0.049	0.002	0.055	-0.059	-0.016	0.010	0.049	-0.049	-0.002	0.011	0.047
elsIpwQr	0.042	0.002	-0.033	0.005	0.008	-0.010	-0.015	0.013	-0.009	0.002	-0.005	0.016	-0.012	0.015	0.000	0.015	-0.009	0.024	0.003	0.008
swelQr	0.013	0.016	-0.033	0.030	-0.003	0.002	-0.020	0.030	-0.001	0.010	-0.013	0.026	0.006	0.018	-0.008	0.022	0.015	0.022	-0.005	0.015
proIpwQr	0.085	0.061	-0.031	-0.009	0.064	0.048	-0.024	0.000	0.053	0.042	-0.019	0.000	0.046	0.036	-0.014	-0.001	0.028	0.033	-0.005	-0.003
RMS																				
fullQr	0.186	0.088	0.092	0.145	0.148	0.073	0.077	0.114	0.131	0.065	0.067	0.099	0.147	0.071	0.073	0.113	0.178	0.091	0.090	0.143
NaiveQr	0.274	0.147	0.127	0.185	0.224	0.121	0.106	0.146	0.212	0.108	0.096	0.133	0.232	0.113	0.106	0.148	0.272	0.129	0.129	0.187
elmIpwQr	0.426	0.283	0.209	0.291	0.417	0.332	0.191	0.269	0.435	0.360	0.190	0.274	0.399	0.316	0.186	0.279	0.428	0.305	0.214	0.349
elsIpwQr	0.360	0.219	0.176	0.248	0.306	0.197	0.136	0.178	0.281	0.164	0.120	0.145	0.268	0.151	0.120	0.149	0.293	0.162	0.139	0.198
swelQr	0.327	0.201	0.154	0.202	0.288	0.186	0.125	0.153	0.267	0.153	0.111	0.131	0.247	0.142	0.111	0.138	0.269	0.153	0.129	0.178
proIpwQr	0.324	0.165	0.153	0.207	0.273	0.139	0.127	0.163	0.255	0.127	0.113	0.143	0.266	0.127	0.118	0.159	0.289	0.135	0.136	0.202
95ECP																				
fullQr	0.931	0.954	0.948	0.946	0.944	0.944	0.938	0.955	0.959	0.948	0.958	0.956	0.945	0.947	0.960	0.951	0.948	0.952	0.956	0.958
NaiveQr	0.920	0.893	0.936	0.949	0.917	0.868	0.938	0.958	0.918	0.888	0.940	0.960	0.926	0.913	0.942	0.954	0.939	0.931	0.947	0.949
elmIpwQr	0.858	0.766	0.870	0.903	0.872	0.744	0.904	0.934	0.877	0.736	0.919	0.953	0.910	0.811	0.941	0.964	0.909	0.861	0.927	0.930
elsIpwQr	0.875	0.815	0.893	0.927	0.889	0.826	0.930	0.950	0.908	0.826	0.938	0.971	0.898	0.855	0.939	0.961	0.914	0.887	0.935	0.947
swelQr	0.864	0.862	0.893	0.932	0.908	0.916	0.926	0.960	0.930	0.961	0.941	0.973	0.931	0.942	0.939	0.961	0.914	0.924	0.906	0.938
proIpwQr	0.909	0.900	0.919	0.943	0.930	0.936	0.936	0.961	0.954	0.944	0.953	0.971	0.941	0.947	0.961	0.965	0.937	0.938	0.939	0.948
CI lengths																				
fullQr	0.741	0.366	0.382	0.588	0.584	0.287	0.303	0.470	0.537	0.266	0.279	0.430	0.588	0.290	0.305	0.465	0.735	0.365	0.379	0.591
NaiveQr	1.036	0.486	0.509	0.760	0.837	0.387	0.408	0.604	0.775	0.354	0.380	0.557	0.843	0.388	0.414	0.602	1.053	0.485	0.515	0.760
elmIpwQr	1.211	0.597	0.613	0.938	1.084	0.531	0.546	0.823	1.026	0.498	0.516	0.752	1.043	0.503	0.522	0.755	1.131	0.544	0.568	0.845
elsIpwQr	1.159	0.585	0.581	0.900	0.989	0.484	0.489	0.734	0.878	0.424	0.438	0.648	0.899	0.424	0.448	0.652	1.039	0.493	0.521	0.773
swelQr	1.201	0.769	0.611	0.928	1.155	0.730	0.537	0.755	1.110	0.665	0.486	0.676	1.056	0.597	0.468	0.663	1.083	0.603	0.520	0.757
proIpwQr	1.172	0.606	0.580	0.872	1.084	0.564	0.511	0.736	1.080	0.539	0.494	0.671	1.111	0.539	0.516	0.708	1.179	0.568	0.570	0.830

Table 8: (Setting 1) The mean biases, root of mean square errors (RMS), coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) and confidence interval(CI) lengths of different estimators for quantile regression coefficient with $\tau = (0.1, 0.25, 0.5, 0.75, 0.9)$, $n = 1000$ and 40% missing in y_i

	$\tau = 0.1$				$\tau = 0.25$				$\tau = 0.5$				$\tau = 0.75$				$\tau = 0.9$			
Method	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
$\beta_{j,\tau}^{true}$	0.680	-2.320	1.872	0.500	0.831	-2.169	1.933	0.500	1.000	-2.000	2.000	0.500	1.169	-1.831	2.067	0.500	1.320	-1.680	2.128	0.500
									Mean Biases											
fullQr	-0.002	0.002	0.001	-0.001	0.003	0.001	-0.002	-0.001	0.003	-0.001	-0.002	0.001	0.002	-0.001	-0.002	0.001	0.004	-0.002	-0.003	0.003
NaiveQr	0.090	0.077	-0.022	-0.012	0.100	0.074	-0.028	-0.012	0.098	0.066	-0.029	-0.007	0.093	0.060	-0.027	-0.008	0.094	0.055	-0.029	-0.007
elmIpwQr	-0.007	-0.079	-0.022	0.012	-0.045	-0.101	-0.008	0.034	-0.086	-0.091	0.010	0.051	-0.078	-0.058	0.011	0.051	-0.061	-0.041	0.008	0.047
elsIpwQr	0.013	-0.019	-0.012	-0.004	-0.007	-0.016	-0.003	0.002	-0.016	-0.010	0.002	0.007	-0.021	-0.002	0.006	0.006	-0.019	0.001	0.007	0.005
swelQr	0.002	-0.008	-0.013	0.005	-0.007	-0.007	-0.006	0.006	-0.009	-0.006	-0.001	0.008	-0.009	-0.001	0.002	0.007	-0.005	0.000	0.004	0.006
proIpwQr	0.073	0.052	-0.023	-0.009	0.064	0.047	-0.019	-0.009	0.057	0.040	-0.017	-0.005	0.044	0.036	-0.011	-0.005	0.039	0.031	-0.008	-0.006
									RMS											
fullQr	0.130	0.066	0.067	0.102	0.101	0.051	0.051	0.080	0.092	0.046	0.048	0.075	0.099	0.049	0.051	0.081	0.125	0.063	0.065	0.102
NaiveQr	0.201	0.114	0.091	0.129	0.171	0.100	0.073	0.102	0.167	0.090	0.071	0.094	0.170	0.088	0.075	0.106	0.194	0.099	0.091	0.129
elmIpwQr	0.330	0.258	0.172	0.248	0.315	0.299	0.166	0.222	0.370	0.285	0.176	0.235	0.362	0.219	0.166	0.265	0.328	0.196	0.170	0.311
elsIpwQr	0.254	0.167	0.126	0.176	0.212	0.139	0.098	0.125	0.188	0.112	0.084	0.100	0.183	0.094	0.083	0.106	0.198	0.097	0.091	0.129
swelQr	0.235	0.152	0.111	0.152	0.192	0.125	0.085	0.114	0.173	0.104	0.077	0.094	0.172	0.089	0.077	0.099	0.184	0.092	0.085	0.119
proIpwQr	0.222	0.120	0.102	0.145	0.182	0.101	0.083	0.114	0.171	0.089	0.077	0.100	0.170	0.086	0.078	0.108	0.192	0.095	0.092	0.130
									95ECP											
fullQr	0.935	0.946	0.952	0.946	0.943	0.947	0.947	0.955	0.952	0.956	0.954	0.947	0.959	0.950	0.958	0.944	0.949	0.953	0.941	0.948
NaiveQr	0.911	0.828	0.936	0.945	0.909	0.810	0.941	0.951	0.886	0.820	0.920	0.958	0.910	0.868	0.939	0.944	0.942	0.910	0.939	0.952
elmIpwQr	0.884	0.776	0.892	0.929	0.892	0.715	0.910	0.933	0.870	0.711	0.907	0.945	0.874	0.782	0.930	0.946	0.903	0.856	0.947	0.947
elsIpwQr	0.911	0.850	0.908	0.944	0.915	0.819	0.937	0.957	0.898	0.813	0.926	0.964	0.898	0.861	0.934	0.962	0.927	0.911	0.944	0.953
swelQr	0.893	0.907	0.914	0.949	0.937	0.929	0.933	0.967	0.939	0.954	0.941	0.968	0.947	0.963	0.943	0.961	0.936	0.947	0.933	0.939
proIpwQr	0.932	0.921	0.928	0.946	0.931	0.917	0.942	0.960	0.937	0.927	0.941	0.966	0.954	0.949	0.962	0.957	0.949	0.941	0.951	0.956
									CI lengths											
fullQr	0.512	0.256	0.266	0.412	0.407	0.204	0.212	0.330	0.377	0.187	0.197	0.302	0.408	0.202	0.212	0.327	0.514	0.255	0.267	0.411
NaiveQr	0.723	0.337	0.357	0.530	0.581	0.270	0.286	0.427	0.536	0.248	0.262	0.389	0.586	0.268	0.285	0.422	0.745	0.341	0.360	0.535
elmIpwQr	0.940	0.456	0.470	0.723	0.809	0.385	0.402	0.600	0.732	0.348	0.362	0.528	0.720	0.336	0.360	0.521	0.807	0.375	0.401	0.591
elsIpwQr	0.873	0.434	0.435	0.667	0.694	0.337	0.347	0.511	0.602	0.282	0.298	0.438	0.599	0.280	0.301	0.442	0.721	0.335	0.359	0.528
swelQr	0.923	0.600	0.472	0.702	0.777	0.522	0.363	0.515	0.754	0.446	0.318	0.435	0.710	0.389	0.310	0.427	0.755	0.399	0.347	0.485
proIpwQr	0.844	0.427	0.406	0.610	0.746	0.384	0.350	0.497	0.711	0.357	0.325	0.444	0.736	0.354	0.346	0.469	0.830	0.391	0.401	0.564

Table 9: (Setting 2) The mean biases, root of mean square errors (RMS), coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) and confidence interval(CI) lengths of different estimators for quantile regression coefficient with τ = $(0.1, 0.25, 0.5, 0.75, 0.9)$, n = 500 and 40% missing in y_i

	τ = 0.1				τ = 0.25				τ = 0.5				τ = 0.75				τ = 0.9			
Method	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
$\beta_{j,\tau}^{true}$	0.657	-2.343	1.863	0.500	0.742	-2.258	1.897	0.500	0.889	-2.111	1.956	0.500	1.127	-1.873	2.051	0.500	1.464	-1.536	2.185	0.500
Mean Biases																				
fullQr	0.003	0.003	-0.001	0.000	0.004	0.002	-0.001	-0.002	0.009	0.001	-0.001	-0.005	0.010	-0.002	0.001	-0.008	0.033	-0.002	-0.012	-0.003
NaiveQr	0.047	0.036	-0.013	-0.004	0.089	0.057	-0.027	-0.008	0.180	0.106	-0.056	-0.015	0.355	0.186	-0.111	-0.031	0.629	0.292	-0.196	-0.059
elmIpwQr	-0.230	0.675	-0.011	-0.140	0.087	0.629	-0.056	-0.074	0.301	0.659	-0.040	-0.067	0.678	0.715	-0.071	-0.141	1.035	0.807	-0.125	-0.177
elsIpwQr	0.036	0.018	-0.018	0.005	0.046	0.025	-0.024	0.008	0.093	0.064	-0.044	0.013	0.189	0.182	-0.081	0.038	0.295	0.301	-0.099	0.064
swelQr	0.000	0.016	-0.011	0.008	0.022	0.015	-0.018	0.014	0.062	0.041	-0.038	0.028	0.128	0.133	-0.074	0.073	0.174	0.236	-0.111	0.165
proIpwQr	0.009	-0.003	-0.008	0.004	-0.008	-0.017	-0.002	0.004	-0.039	-0.030	0.008	0.008	-0.107	-0.056	0.031	0.024	-0.178	-0.062	0.055	0.033
RMS																				
fullQr	0.084	0.040	0.041	0.065	0.101	0.050	0.053	0.078	0.144	0.072	0.073	0.109	0.260	0.125	0.130	0.200	0.483	0.248	0.244	0.385
NaiveQr	0.138	0.071	0.063	0.093	0.177	0.095	0.080	0.112	0.293	0.154	0.122	0.165	0.560	0.273	0.225	0.300	1.022	0.468	0.432	0.565
elmIpwQr	6.275	3.166	3.250	4.731	4.416	2.486	2.500	3.580	3.439	2.168	1.784	2.678	2.849	1.979	1.317	2.271	3.042	1.864	1.300	2.293
elsIpwQr	0.154	0.079	0.073	0.101	0.219	0.117	0.097	0.116	0.476	0.240	0.191	0.174	1.214	0.706	0.375	0.484	1.652	1.071	0.596	0.884
swelQr	0.135	0.071	0.064	0.086	0.189	0.105	0.087	0.103	0.426	0.219	0.176	0.154	1.188	0.686	0.361	0.472	1.555	0.991	0.567	0.876
proIpwQr	0.139	0.063	0.069	0.099	0.166	0.077	0.082	0.113	0.226	0.109	0.109	0.151	0.369	0.171	0.174	0.235	0.603	0.290	0.296	0.440
95ECP																				
fullQr	0.940	0.961	0.956	0.953	0.946	0.946	0.943	0.958	0.949	0.948	0.963	0.960	0.943	0.948	0.956	0.955	0.939	0.952	0.959	0.953
NaiveQr	0.934	0.906	0.941	0.950	0.918	0.872	0.933	0.949	0.908	0.874	0.937	0.962	0.884	0.865	0.915	0.950	0.897	0.909	0.918	0.958
elmIpwQr	0.838	0.804	0.875	0.906	0.823	0.758	0.854	0.887	0.807	0.727	0.865	0.888	0.858	0.763	0.904	0.880	0.843	0.791	0.869	0.895
elsIpwQr	0.906	0.878	0.911	0.933	0.899	0.853	0.925	0.956	0.900	0.849	0.940	0.969	0.869	0.850	0.909	0.968	0.864	0.876	0.913	0.947
swelQr	0.915	0.935	0.918	0.941	0.930	0.943	0.945	0.968	0.935	0.956	0.938	0.975	0.882	0.927	0.918	0.962	0.838	0.891	0.862	0.892
proIpwQr	0.919	0.935	0.923	0.937	0.932	0.933	0.928	0.949	0.941	0.930	0.942	0.960	0.918	0.934	0.949	0.971	0.921	0.935	0.944	0.947
CI lengths																				
fullQr	0.333	0.163	0.172	0.263	0.397	0.196	0.207	0.321	0.591	0.292	0.307	0.474	1.040	0.518	0.539	0.834	1.979	1.006	1.017	1.605
NaiveQr	0.520	0.245	0.254	0.373	0.646	0.304	0.310	0.460	0.974	0.458	0.469	0.692	1.720	0.810	0.828	1.213	3.227	1.558	1.544	2.319
elmIpwQr	2.764	1.189	1.434	2.013	2.569	1.055	1.295	1.843	2.661	1.145	1.285	1.856	3.407	1.622	1.612	2.386	4.875	2.507	2.311	3.634
elsIpwQr	0.541	0.254	0.268	0.395	0.730	0.335	0.356	0.493	1.158	0.562	0.560	0.786	1.965	1.069	0.961	1.417	3.279	1.852	1.618	2.505
swelQr	0.537	0.291	0.249	0.351	0.799	0.453	0.358	0.474	1.539	0.945	0.697	0.842	3.044	1.940	1.351	1.983	4.277	2.966	1.948	3.081
proIpwQr	0.561	0.268	0.282	0.422	0.676	0.320	0.333	0.494	0.930	0.451	0.457	0.672	1.432	0.692	0.715	1.054	2.356	1.158	1.204	1.821

Table 10: (Setting 2) The mean biases, root of mean square errors (RMS), coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) and confidence interval(CI) lengths of different estimators for quantile regression coefficient with τ = $(0.1, 0.25, 0.5, 0.75, 0.9)$, n = 1000 and 40% missing in y_i

	τ = 0.1				τ = 0.25				τ = 0.5				τ = 0.75				τ = 0.9			
Method	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
$\beta_{j,\tau}^{true}$	0.657	-2.343	1.863	0.500	0.742	-2.258	1.897	0.500	0.889	-2.111	1.956	0.500	1.127	-1.873	2.051	0.500	1.464	-1.536	2.185	0.500
									Mean Biases											
fullQr	0.001	0.001	0.000	0.000	0.003	0.000	-0.001	-0.001	0.005	-0.001	-0.003	0.001	0.008	0.000	-0.004	0.001	0.017	-0.008	-0.009	0.007
NaiveQr	0.044	0.034	-0.012	-0.005	0.085	0.058	-0.025	-0.010	0.170	0.103	-0.053	-0.012	0.327	0.182	-0.103	-0.024	0.594	0.282	-0.192	-0.037
elmIpwQr	-0.041	0.456	-0.169	0.131	0.089	0.444	-0.105	0.067	0.297	0.438	-0.067	-0.013	0.585	0.553	-0.105	-0.039	0.961	0.678	-0.174	-0.085
elsIpwQr	0.020	0.007	-0.010	0.001	0.018	0.009	-0.010	0.003	0.030	0.018	-0.017	0.009	0.051	0.055	-0.027	0.019	0.095	0.122	-0.037	0.027
swelQr	-0.001	0.008	-0.005	0.002	0.008	0.003	-0.008	0.005	0.015	0.006	-0.012	0.012	0.016	0.031	-0.017	0.027	0.021	0.078	-0.025	0.057
proIpwQr	0.002	-0.008	-0.003	0.000	-0.017	-0.020	0.002	0.003	-0.045	-0.037	0.009	0.011	-0.102	-0.056	0.028	0.019	-0.192	-0.075	0.056	0.043
									RMS											
fullQr	0.058	0.029	0.029	0.046	0.069	0.035	0.035	0.055	0.102	0.050	0.053	0.082	0.176	0.088	0.091	0.142	0.338	0.173	0.176	0.277
NaiveQr	0.100	0.054	0.045	0.065	0.136	0.079	0.058	0.078	0.238	0.129	0.097	0.116	0.433	0.226	0.171	0.213	0.811	0.387	0.327	0.395
elmIpwQr	5.121	2.600	3.011	3.939	4.108	2.235	2.385	3.102	3.415	1.882	1.662	2.566	3.453	1.744	1.432	2.258	3.694	1.642	1.291	2.460
elsIpwQr	0.102	0.055	0.048	0.069	0.135	0.079	0.062	0.080	0.224	0.133	0.100	0.112	0.425	0.258	0.180	0.194	0.777	0.548	0.332	0.390
swelQr	0.092	0.050	0.043	0.061	0.120	0.071	0.056	0.072	0.195	0.119	0.089	0.103	0.362	0.224	0.164	0.183	0.637	0.501	0.305	0.357
proIpwQr	0.095	0.046	0.047	0.069	0.111	0.055	0.053	0.080	0.159	0.083	0.079	0.110	0.266	0.123	0.126	0.174	0.435	0.207	0.208	0.306
									95ECP											
fullQr	0.938	0.947	0.955	0.945	0.950	0.952	0.940	0.954	0.952	0.949	0.955	0.950	0.963	0.950	0.960	0.948	0.953	0.954	0.941	0.954
NaiveQr	0.915	0.866	0.930	0.936	0.901	0.815	0.935	0.960	0.848	0.765	0.900	0.952	0.840	0.766	0.897	0.950	0.846	0.850	0.890	0.953
elmIpwQr	0.839	0.785	0.855	0.895	0.771	0.689	0.839	0.889	0.771	0.652	0.851	0.892	0.816	0.697	0.887	0.908	0.848	0.758	0.887	0.908
elsIpwQr	0.928	0.876	0.924	0.930	0.920	0.834	0.926	0.955	0.900	0.809	0.918	0.966	0.860	0.831	0.912	0.965	0.879	0.886	0.921	0.951
swelQr	0.917	0.935	0.921	0.934	0.934	0.939	0.944	0.957	0.941	0.935	0.939	0.977	0.903	0.942	0.921	0.964	0.867	0.926	0.897	0.929
proIpwQr	0.941	0.929	0.922	0.934	0.946	0.928	0.950	0.945	0.929	0.908	0.929	0.959	0.902	0.916	0.933	0.953	0.918	0.925	0.946	0.951
									CI lengths											
fullQr	0.229	0.113	0.119	0.183	0.277	0.139	0.144	0.225	0.416	0.206	0.216	0.333	0.721	0.361	0.375	0.581	1.386	0.701	0.720	1.119
NaiveQr	0.357	0.169	0.173	0.257	0.447	0.212	0.217	0.321	0.676	0.320	0.324	0.482	1.201	0.560	0.574	0.850	2.302	1.084	1.090	1.636
elmIpwQr	2.201	0.875	1.018	1.599	1.967	0.780	0.860	1.421	2.027	0.838	0.897	1.416	2.569	1.172	1.148	1.788	3.823	1.832	1.695	2.758
elsIpwQr	0.370	0.176	0.181	0.272	0.452	0.214	0.225	0.334	0.673	0.313	0.330	0.484	1.130	0.555	0.563	0.828	2.108	1.080	1.040	1.551
swelQr	0.357	0.201	0.165	0.240	0.502	0.301	0.225	0.314	0.878	0.538	0.372	0.478	1.609	1.049	0.696	0.901	2.467	1.608	1.162	1.642
proIpwQr	0.367	0.174	0.180	0.271	0.441	0.213	0.219	0.323	0.625	0.293	0.308	0.448	0.964	0.458	0.482	0.709	1.659	0.803	0.850	1.261

Table 11: (Setting 3) The mean biases, root of mean square errors (RMS), coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) and confidence interval(CI) lengths of different estimators for quantile regression coefficient with $\tau = (0.1, 0.25, 0.5, 0.75, 0.9)$, $n = 1000$ and 40% missing in y_i

	$\tau = 0.1$				$\tau = 0.25$				$\tau = 0.5$				$\tau = 0.75$				$\tau = 0.9$			
Method	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
$\beta_{j,\tau}^{true}$	0.507	-2.493	1.803	0.500	0.635	-2.365	1.854	0.500	0.871	-2.129	1.948	0.500	1.226	-1.774	2.090	0.500	1.664	-1.336	2.265	0.500
										Mean Biases										
fullQr	-0.001	0.003	0.001	0.000	0.001	0.002	0.001	0.000	0.005	0.006	-0.002	-0.001	0.011	0.008	-0.006	0.001	0.009	0.009	-0.005	0.004
NaiveQr	0.095	0.074	-0.028	-0.008	0.189	0.137	-0.057	-0.013	0.352	0.234	-0.110	-0.022	0.575	0.329	-0.184	-0.037	0.745	0.389	-0.241	-0.040
elmIpwQr	0.087	0.082	-0.026	-0.003	0.078	0.125	-0.010	0.003	0.109	0.196	-0.009	0.021	0.377	0.347	-0.106	0.023	0.643	0.478	-0.184	-0.007
elsIpwQr	0.038	0.022	-0.023	0.011	0.051	0.037	-0.029	0.013	0.069	0.061	-0.039	0.025	0.094	0.121	-0.050	0.041	0.140	0.183	-0.053	0.039
swelQr	0.020	0.015	-0.017	0.011	0.028	0.020	-0.022	0.017	0.029	0.031	-0.027	0.029	0.035	0.076	-0.036	0.056	0.058	0.133	-0.044	0.074
proIpwQr	0.009	-0.001	-0.009	0.006	-0.009	-0.015	-0.004	0.010	-0.039	-0.028	0.005	0.014	-0.088	-0.037	0.019	0.028	-0.122	-0.029	0.033	0.032
										RMS										
fullQr	0.083	0.041	0.043	0.066	0.112	0.057	0.056	0.087	0.165	0.081	0.083	0.133	0.258	0.124	0.129	0.203	0.414	0.201	0.214	0.335
NaiveQr	0.169	0.100	0.072	0.097	0.261	0.164	0.102	0.130	0.444	0.265	0.167	0.193	0.709	0.379	0.265	0.289	0.977	0.489	0.389	0.450
elmIpwQr	0.215	0.122	0.093	0.111	0.393	0.209	0.167	0.176	0.715	0.354	0.297	0.305	0.924	0.522	0.370	0.426	1.275	0.744	0.512	0.651
elsIpwQr	0.162	0.099	0.075	0.099	0.265	0.171	0.111	0.129	0.453	0.287	0.183	0.185	0.725	0.444	0.294	0.305	0.998	0.554	0.413	0.477
swelQr	0.146	0.087	0.067	0.089	0.235	0.153	0.099	0.120	0.402	0.250	0.166	0.176	0.663	0.377	0.279	0.294	0.896	0.474	0.401	0.450
proIpwQr	0.141	0.067	0.070	0.101	0.185	0.094	0.088	0.127	0.257	0.126	0.122	0.174	0.370	0.171	0.180	0.251	0.541	0.235	0.267	0.383
										95ECP										
fullQr	0.932	0.943	0.927	0.934	0.938	0.943	0.952	0.951	0.946	0.950	0.958	0.939	0.937	0.947	0.949	0.956	0.949	0.950	0.951	0.948
NaiveQr	0.896	0.819	0.905	0.941	0.838	0.677	0.908	0.951	0.744	0.539	0.853	0.934	0.727	0.607	0.845	0.955	0.802	0.768	0.882	0.953
elmIpwQr	0.820	0.714	0.832	0.926	0.704	0.597	0.807	0.905	0.679	0.525	0.823	0.887	0.778	0.610	0.880	0.942	0.824	0.747	0.914	0.937
elsIpwQr	0.892	0.813	0.907	0.938	0.856	0.733	0.921	0.952	0.802	0.687	0.886	0.952	0.798	0.740	0.883	0.956	0.829	0.861	0.904	0.951
swelQr	0.917	0.922	0.909	0.937	0.912	0.912	0.929	0.959	0.903	0.926	0.934	0.959	0.897	0.937	0.901	0.956	0.846	0.926	0.868	0.924
proIpwQr	0.921	0.926	0.931	0.927	0.944	0.931	0.932	0.943	0.936	0.937	0.946	0.952	0.926	0.941	0.944	0.964	0.923	0.961	0.933	0.952
										CI lengths										
fullQr	0.320	0.160	0.167	0.257	0.440	0.219	0.229	0.354	0.648	0.324	0.337	0.521	1.012	0.508	0.528	0.820	1.671	0.836	0.866	1.349
NaiveQr	0.544	0.265	0.261	0.388	0.745	0.357	0.358	0.532	1.064	0.502	0.511	0.755	1.650	0.762	0.785	1.175	2.602	1.211	1.249	1.861
elmIpwQr	0.570	0.272	0.276	0.408	0.853	0.386	0.415	0.580	1.482	0.631	0.706	0.900	2.375	1.016	1.110	1.450	3.335	1.524	1.579	2.248
elsIpwQr	0.530	0.260	0.264	0.393	0.745	0.358	0.369	0.545	1.076	0.515	0.528	0.774	1.636	0.781	0.810	1.183	2.455	1.207	1.207	1.842
swelQr	0.590	0.364	0.268	0.370	1.002	0.662	0.433	0.575	1.766	1.098	0.710	0.905	2.796	1.695	1.167	1.446	3.159	2.098	1.499	2.152
proIpwQr	0.532	0.259	0.263	0.389	0.751	0.356	0.364	0.530	1.038	0.502	0.504	0.724	1.457	0.682	0.717	1.041	2.091	0.981	1.040	1.566

Table 12: (Setting 3) The root of mean square errors (RMS) of the proposed estimators with different numbers of k_n (k_n = $20, 30, 40, 50, 100$) and m (m = $10, 20, 50, 100$) for quantile regression coefficients with differen sample sizes and 40% missing in y_i

		τ = 0.25				τ = 0.5				τ = 0.75			
k_n	m	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
						n = 500							
20	10	0.268	0.125	0.126	0.189	0.388	0.171	0.174	0.256	0.537	0.239	0.244	0.364
	20	0.264	0.125	0.124	0.189	0.383	0.167	0.173	0.252	0.537	0.238	0.245	0.362
	50	0.261	0.126	0.124	0.188	0.384	0.172	0.174	0.250	0.541	0.242	0.247	0.363
	100	0.265	0.125	0.125	0.188	0.382	0.172	0.171	0.249	0.544	0.242	0.247	0.361
30	10	0.292	0.136	0.131	0.191	0.410	0.183	0.183	0.256	0.545	0.261	0.250	0.361
	20	0.301	0.137	0.132	0.192	0.406	0.184	0.184	0.254	0.557	0.264	0.250	0.364
	50	0.303	0.140	0.134	0.190	0.402	0.187	0.184	0.254	0.555	0.267	0.253	0.365
	100	0.302	0.140	0.135	0.190	0.401	0.190	0.183	0.253	0.567	0.266	0.257	0.360
40	10	0.311	0.144	0.135	0.192	0.420	0.200	0.187	0.255	0.578	0.285	0.260	0.371
	20	0.318	0.146	0.140	0.191	0.416	0.203	0.186	0.260	0.591	0.286	0.271	0.373
	50	0.313	0.146	0.139	0.188	0.424	0.208	0.192	0.256	0.602	0.288	0.272	0.364
	100	0.320	0.155	0.142	0.189	0.449	0.240	0.196	0.258	0.649	0.305	0.279	0.382
50	10	0.444	0.210	0.211	0.192	0.433	0.387	0.193	0.266	0.599	0.416	0.268	0.382
	20	0.324	0.154	0.143	0.190	0.443	0.227	0.198	0.260	0.619	0.306	0.276	0.381
	50	0.375	0.181	0.158	0.193	0.525	0.257	0.201	0.286	0.643	0.315	0.279	0.383
	100	0.391	0.198	0.161	0.196	0.533	0.258	0.200	0.294	0.640	0.316	0.285	0.383
100	10	1.004	0.386	0.435	0.355	0.855	0.529	0.320	0.386	0.888	0.530	0.343	0.473
	20	0.962	0.357	0.379	0.485	0.877	0.539	0.292	0.493	0.963	0.587	0.341	0.583
	50	1.009	0.385	0.390	0.510	0.865	0.571	0.297	0.504	0.952	0.591	0.375	0.616
	100	1.011	0.375	0.382	0.511	0.879	0.539	0.300	0.502	0.930	0.540	0.372	0.615
						n = 1000							
20	10	0.196	0.090	0.095	0.133	0.269	0.121	0.131	0.181	0.385	0.162	0.182	0.265
	20	0.197	0.090	0.096	0.134	0.271	0.125	0.130	0.182	0.387	0.168	0.184	0.261
	50	0.197	0.091	0.096	0.134	0.273	0.126	0.129	0.182	0.397	0.172	0.188	0.256
	100	0.198	0.091	0.096	0.134	0.274	0.127	0.129	0.183	0.399	0.173	0.188	0.255
30	10	0.198	0.092	0.096	0.134	0.277	0.126	0.134	0.183	0.395	0.167	0.186	0.272
	20	0.197	0.091	0.095	0.134	0.274	0.126	0.132	0.182	0.396	0.166	0.186	0.266
	50	0.199	0.092	0.096	0.134	0.275	0.126	0.132	0.182	0.400	0.170	0.188	0.261
	100	0.199	0.092	0.096	0.133	0.276	0.128	0.131	0.181	0.406	0.171	0.190	0.258
40	10	0.202	0.094	0.097	0.135	0.279	0.133	0.136	0.183	0.410	0.177	0.189	0.275
	20	0.200	0.092	0.097	0.135	0.280	0.127	0.134	0.182	0.409	0.169	0.189	0.269
	50	0.199	0.093	0.096	0.133	0.285	0.128	0.135	0.182	0.407	0.169	0.188	0.265
	100	0.200	0.093	0.096	0.133	0.284	0.129	0.134	0.182	0.412	0.174	0.191	0.262
50	10	0.208	0.098	0.097	0.136	0.288	0.139	0.138	0.186	0.416	0.185	0.190	0.277
	20	0.202	0.094	0.097	0.136	0.283	0.130	0.135	0.182	0.410	0.175	0.189	0.271
	50	0.201	0.094	0.096	0.134	0.286	0.131	0.135	0.180	0.416	0.173	0.192	0.265
	100	0.201	0.093	0.096	0.134	0.290	0.132	0.135	0.181	0.421	0.174	0.192	0.264
100	10	0.223	0.109	0.101	0.138	0.310	0.162	0.147	0.187	0.461	0.213	0.203	0.278
	20	0.215	0.102	0.099	0.136	0.302	0.150	0.142	0.185	0.453	0.198	0.201	0.276
	50	0.212	0.102	0.099	0.135	0.306	0.146	0.141	0.184	0.457	0.193	0.202	0.272
	100	0.211	0.100	0.098	0.134	0.308	0.145	0.141	0.183	0.460	0.192	0.201	0.272

Table 13: (Setting 3) The coverage probabilities of bootstrap confidence intervals with a nominal level of 0.95 (95ECP) of the proposed estimators with different numbers of k_n ($k_n = 20, 30, 40, 50, 100$) and m ($m = 10, 20, 50, 100$) for quantile regression coefficients with differen sample sizes and 40% missing in y_i

		$\tau = 0.25$				$\tau = 0.5$				$\tau = 0.75$			
k_n	m	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
							$n = 500$						
20	10	0.936	0.948	0.952	0.964	0.934	0.954	0.948	0.960	0.942	0.944	0.960	0.962
	20	0.934	0.940	0.958	0.952	0.926	0.956	0.946	0.952	0.942	0.940	0.956	0.958
	50	0.940	0.928	0.956	0.954	0.924	0.942	0.950	0.950	0.936	0.930	0.956	0.962
	100	0.940	0.940	0.952	0.958	0.932	0.948	0.948	0.954	0.938	0.926	0.948	0.958
30	10	0.934	0.950	0.960	0.966	0.934	0.964	0.946	0.960	0.942	0.944	0.960	0.968
	20	0.944	0.956	0.950	0.958	0.932	0.960	0.950	0.958	0.946	0.948	0.960	0.966
	50	0.942	0.952	0.960	0.968	0.926	0.956	0.942	0.956	0.944	0.938	0.956	0.964
	100	0.946	0.958	0.946	0.956	0.932	0.954	0.946	0.958	0.936	0.944	0.954	0.968
40	10	0.942	0.956	0.958	0.964	0.946	0.968	0.950	0.964	0.956	0.952	0.958	0.974
	20	0.944	0.958	0.954	0.964	0.942	0.962	0.954	0.962	0.954	0.960	0.958	0.972
	50	0.946	0.948	0.956	0.960	0.940	0.960	0.950	0.958	0.950	0.952	0.962	0.974
	100	0.954	0.946	0.954	0.962	0.938	0.960	0.946	0.958	0.942	0.946	0.958	0.966
50	10	0.940	0.944	0.952	0.958	0.940	0.958	0.950	0.962	0.962	0.960	0.960	0.980
	20	0.942	0.954	0.958	0.960	0.948	0.960	0.960	0.964	0.958	0.950	0.960	0.976
	50	0.950	0.956	0.956	0.968	0.942	0.966	0.958	0.964	0.958	0.956	0.970	0.974
	100	0.950	0.956	0.958	0.962	0.942	0.962	0.948	0.962	0.950	0.946	0.966	0.968
100	10	0.942	0.930	0.944	0.962	0.940	0.940	0.950	0.964	0.958	0.952	0.968	0.962
	20	0.942	0.960	0.952	0.962	0.946	0.958	0.960	0.966	0.960	0.960	0.966	0.968
	50	0.954	0.956	0.968	0.968	0.958	0.952	0.962	0.964	0.962	0.952	0.964	0.972
	100	0.954	0.952	0.956	0.964	0.944	0.952	0.944	0.958	0.966	0.956	0.964	0.974
							$n = 1000$						
20	10	0.930	0.934	0.942	0.952	0.946	0.950	0.932	0.948	0.940	0.954	0.956	0.948
	20	0.938	0.930	0.940	0.946	0.946	0.946	0.934	0.956	0.952	0.944	0.936	0.938
	50	0.936	0.920	0.920	0.950	0.940	0.942	0.942	0.952	0.942	0.942	0.946	0.942
	100	0.928	0.928	0.930	0.944	0.948	0.932	0.934	0.946	0.936	0.938	0.944	0.940
30	10	0.932	0.936	0.930	0.950	0.958	0.944	0.936	0.952	0.956	0.952	0.944	0.936
	20	0.940	0.950	0.944	0.952	0.946	0.946	0.942	0.956	0.950	0.950	0.952	0.938
	50	0.928	0.928	0.934	0.950	0.946	0.950	0.936	0.950	0.942	0.956	0.938	0.938
	100	0.932	0.930	0.930	0.950	0.942	0.940	0.938	0.948	0.938	0.950	0.944	0.948
40	10	0.938	0.930	0.928	0.946	0.952	0.934	0.930	0.940	0.944	0.948	0.944	0.950
	20	0.942	0.946	0.942	0.950	0.954	0.952	0.946	0.950	0.946	0.958	0.962	0.932
	50	0.924	0.932	0.932	0.954	0.942	0.946	0.940	0.952	0.946	0.954	0.942	0.936
	100	0.934	0.932	0.926	0.952	0.946	0.946	0.938	0.944	0.940	0.956	0.948	0.944
50	10	0.934	0.936	0.930	0.948	0.952	0.944	0.936	0.940	0.950	0.950	0.952	0.934
	20	0.938	0.928	0.928	0.954	0.954	0.938	0.940	0.948	0.944	0.946	0.948	0.940
	50	0.940	0.946	0.932	0.944	0.954	0.946	0.944	0.954	0.946	0.958	0.954	0.930
	100	0.930	0.932	0.928	0.956	0.946	0.952	0.954	0.956	0.946	0.966	0.944	0.944
100	10	0.942	0.920	0.932	0.944	0.958	0.924	0.940	0.946	0.948	0.934	0.940	0.942
	20	0.946	0.936	0.950	0.950	0.956	0.960	0.940	0.960	0.944	0.950	0.952	0.952
	50	0.938	0.934	0.950	0.964	0.940	0.950	0.952	0.958	0.952	0.960	0.958	0.948
	100	0.948	0.950	0.946	0.952	0.956	0.956	0.958	0.956	0.954	0.970	0.952	0.940

Table 14: (Setting 3) The confidence interval lengths (CI lengths) of bootstrap confidence intervals with a nominal level of 0.95 of the proposed estimators with different numbers of k_n $(k_n = 20, 30, 40, 50, 100)$ and m $(m = 10, 20, 50, 100)$ for quantile regression coefficients with differen sample sizes and 40% missing in y_i

		$\tau = 0.25$				$\tau = 0.5$				$\tau = 0.75$			
k_n	m	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3	β_0	β_1	β_2	β_3
						$n = 500$							
20	10	1.162	0.561	0.568	0.805	1.592	0.787	0.760	1.105	2.152	1.065	1.056	1.539
	20	1.147	0.558	0.558	0.792	1.575	0.785	0.751	1.105	2.131	1.052	1.045	1.539
	50	1.149	0.566	0.579	0.814	1.589	0.779	0.756	1.102	2.122	1.043	1.043	1.532
	100	1.159	0.567	0.576	0.805	1.578	0.788	0.752	1.099	2.129	1.041	1.044	1.540
30	10	1.339	0.641	0.662	0.875	1.748	0.894	0.836	1.175	2.308	1.182	1.120	1.653
	20	1.351	0.676	0.685	0.902	1.749	0.921	0.837	1.202	2.291	1.198	1.118	1.643
	50	1.312	0.645	0.651	0.858	1.736	0.911	0.814	1.180	2.267	1.178	1.103	1.632
	100	1.344	0.664	0.675	0.888	1.748	0.909	0.819	1.181	2.280	1.174	1.106	1.651
40	10	1.551	0.735	0.776	0.969	1.958	1.014	0.926	1.267	2.514	1.319	1.193	1.748
	20	1.545	0.745	0.780	0.972	1.950	1.031	0.918	1.279	2.506	1.299	1.187	1.736
	50	1.583	0.767	0.801	1.007	1.964	1.063	0.917	1.297	2.506	1.341	1.175	1.768
	100	1.507	0.747	0.760	0.975	1.929	1.050	0.899	1.276	2.482	1.326	1.177	1.734
50	10	1.643	0.772	0.836	1.007	2.039	1.082	0.963	1.318	2.606	1.383	1.249	1.830
	20	1.772	0.859	0.872	1.078	2.126	1.180	0.983	1.384	2.644	1.460	1.254	1.873
	50	1.766	0.878	0.895	1.094	2.118	1.194	0.994	1.374	2.630	1.475	1.263	1.867
	100	1.765	0.891	0.879	1.121	2.173	1.210	0.992	1.391	2.690	1.469	1.258	1.880
100	10	2.215	1.053	1.101	1.302	2.569	1.417	1.185	1.610	3.084	1.750	1.432	2.134
	20	2.354	1.149	1.167	1.392	2.693	1.556	1.239	1.676	3.168	1.856	1.478	2.211
	50	2.476	1.259	1.268	1.523	2.802	1.702	1.300	1.802	3.295	1.984	1.510	2.338
	100	2.498	1.270	1.246	1.531	2.800	1.701	1.280	1.780	3.298	1.994	1.498	2.317
						$n = 1000$							
20	10	0.744	0.351	0.360	0.528	1.055	0.499	0.511	0.733	1.487	0.684	0.720	1.039
	20	0.735	0.350	0.356	0.526	1.052	0.493	0.507	0.723	1.491	0.671	0.716	1.033
	50	0.733	0.349	0.357	0.528	1.046	0.494	0.506	0.728	1.474	0.672	0.714	1.026
	100	0.728	0.348	0.355	0.525	1.050	0.494	0.506	0.725	1.478	0.670	0.713	1.029
30	10	0.761	0.361	0.370	0.538	1.096	0.517	0.528	0.742	1.544	0.710	0.740	1.061
	20	0.751	0.359	0.364	0.534	1.081	0.512	0.522	0.741	1.532	0.698	0.735	1.049
	50	0.748	0.357	0.358	0.529	1.069	0.509	0.514	0.734	1.525	0.689	0.724	1.040
	100	0.754	0.361	0.367	0.535	1.082	0.513	0.518	0.734	1.522	0.690	0.730	1.037
40	10	0.797	0.375	0.386	0.542	1.135	0.538	0.544	0.754	1.591	0.742	0.759	1.072
	20	0.795	0.375	0.380	0.540	1.131	0.538	0.538	0.753	1.586	0.725	0.753	1.071
	50	0.782	0.374	0.372	0.546	1.118	0.537	0.529	0.750	1.590	0.718	0.748	1.058
	100	0.783	0.380	0.377	0.539	1.120	0.543	0.531	0.745	1.586	0.721	0.751	1.053
50	10	0.869	0.407	0.412	0.561	1.201	0.580	0.566	0.791	1.667	0.779	0.788	1.114
	20	0.841	0.404	0.409	0.572	1.189	0.580	0.558	0.779	1.646	0.766	0.777	1.093
	50	0.864	0.412	0.413	0.572	1.187	0.594	0.559	0.778	1.649	0.774	0.774	1.093
	100	0.853	0.415	0.409	0.562	1.182	0.600	0.552	0.773	1.639	0.774	0.774	1.092
100	10	1.107	0.513	0.534	0.669	1.434	0.725	0.664	0.890	1.881	0.952	0.870	1.247
	20	1.184	0.579	0.573	0.725	1.519	0.823	0.677	0.953	1.938	1.004	0.875	1.291
	50	1.261	0.601	0.608	0.736	1.549	0.849	0.704	0.967	1.982	1.046	0.902	1.293
	100	1.268	0.626	0.602	0.746	1.577	0.894	0.695	0.990	2.013	1.058	0.899	1.320

Table 15: The average computing time (in seconds) for different methods to estimate the coefficients $\boldsymbol{\beta}_\tau$ at the quantile levels $\tau = 0.25, 0.5$ and 0.75 in three considered settings with various sample sizes and missing rates

Setting	missing rate	fullQr	NaiveQr	elmIpwQr	elsIpwQr	swelQr	(k_n, m)	proIpwQr (20, 10)	(20, 20)	(40, 10)	(40, 20)
						sample size n = 500					
1	20%	0.002	0.002	35.232	4.911	7.573		0.840	0.977	1.153	1.295
1	40%	0.003	0.002	51.342	6.262	7.169		0.855	1.105	1.105	1.358
2	20%	0.002	0.002	29.956	5.009	7.679		0.853	0.993	1.061	1.189
2	40%	0.003	0.002	58.871	6.700	7.320		0.872	1.117	1.063	1.375
3	20%	0.002	0.002	31.244	4.931	6.900		0.799	0.978	1.130	1.247
3	40%	0.002	0.002	56.053	5.526	6.400		0.832	1.092	1.070	1.347
						sample size n = 1000					
1	20%	0.004	0.004	73.301	6.112	9.399		1.561	1.855	2.232	2.509
1	40%	0.003	0.004	174.934	7.623	9.180		1.501	2.067	2.012	2.527
2	20%	0.004	0.002	65.714	6.595	9.580		1.512	1.878	2.221	2.492
2	40%	0.004	0.002	163.688	7.827	9.187		1.490	2.098	1.904	2.512
3	20%	0.004	0.003	77.813	6.475	8.456		1.507	1.842	2.105	2.412
3	40%	0.004	0.002	198.241	7.544	7.357		1.474	2.026	1.932	2.454

Table 16: The estimated coefficients (Est) and the corresponding standard errors (SE) and p-values (p) from different methods with $\tau = (0.1, 0.25, 0.5, 0.75, 0.9)$, $n = 7148$ and 40% missing in response y_i

	fullQr			NaiveQr			elsIpwQr			proIpwQr		
Covariate	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p
τ = 0.1												
Intercept	92.07	2.47	0.00	87.56	3.54	0.00	87.49	3.40	0.00	90.42	2.80	0.00
Race	-2.18	2.36	0.36	-4.46	3.18	0.16	-4.41	3.14	0.16	-1.53	2.99	0.61
Insurance type	-4.38	1.28	0.00	-6.03	1.48	0.00	-6.10	1.47	0.00	-5.04	1.49	0.00
Marital status	0.67	1.23	0.59	2.06	1.43	0.15	2.08	1.54	0.18	0.75	1.22	0.54
First language	-0.83	2.89	0.78	-0.72	5.19	0.89	-0.73	5.25	0.89	0.12	3.69	0.97
Age	0.16	0.04	0.00	0.15	0.05	0.00	0.15	0.05	0.00	0.18	0.05	0.00
Gender	1.69	1.19	0.16	-0.49	1.53	0.75	-0.49	1.52	0.75	0.61	1.32	0.65
τ = 0.25												
Intercept	109.32	3.00	0.00	98.44	3.61	0.00	98.46	3.20	0.00	103.65	3.53	0.00
Race	0.07	1.97	0.97	2.85	2.71	0.29	2.90	2.82	0.30	2.51	2.81	0.37
Insurance type	-2.46	1.43	0.09	-3.71	1.44	0.01	-3.86	1.37	0.01	-3.09	1.63	0.06
Marital status	1.07	1.35	0.43	2.05	1.32	0.12	2.12	1.52	0.16	0.93	1.63	0.57
First language	4.25	5.61	0.45	3.83	4.51	0.40	2.87	4.63	0.54	9.58	5.23	0.07
Age	0.15	0.05	0.00	0.20	0.05	0.00	0.20	0.05	0.00	0.21	0.05	0.00
Gender	4.66	1.12	0.00	3.84	1.42	0.01	3.64	1.48	0.01	3.66	1.35	0.01
τ = 0.5												
Intercept	138.92	4.15	0.00	127.96	6.31	0.00	128.30	7.06	0.00	134.51	6.32	0.00
Race	10.33	3.75	0.01	24.99	7.84	0.00	25.47	8.94	0.00	17.56	6.83	0.01
Insurance type	-0.55	1.92	0.77	-0.99	2.96	0.74	-1.25	2.91	0.67	0.53	3.46	0.88
Marital status	1.27	1.65	0.44	3.23	2.88	0.26	3.33	3.12	0.29	0.60	2.81	0.83
First language	10.94	4.54	0.02	13.26	7.53	0.08	12.68	8.04	0.12	13.05	7.12	0.07
Age	0.11	0.07	0.11	0.17	0.10	0.11	0.17	0.11	0.14	0.16	0.10	0.13
Gender	5.28	1.85	0.00	7.85	2.71	0.00	7.76	3.02	0.01	5.97	2.48	0.02
τ = 0.75												
Intercept	191.98	6.66	0.00	234.45	15.71	0.00	234.99	17.40	0.00	194.08	8.95	0.00
Race	27.00	5.92	0.00	44.09	8.53	0.00	44.63	9.80	0.00	30.20	5.60	0.00
Insurance type	5.87	3.11	0.06	10.71	7.15	0.13	10.12	7.35	0.17	6.09	4.35	0.16
Marital status	6.64	2.78	0.02	15.84	7.13	0.03	15.89	7.51	0.03	7.79	3.06	0.01
First language	12.18	6.58	0.06	18.62	13.31	0.16	19.16	14.88	0.20	10.19	5.39	0.06
Age	-0.21	0.11	0.05	-0.67	0.25	0.01	-0.67	0.28	0.02	-0.18	0.14	0.20
Gender	6.15	2.67	0.02	10.56	6.93	0.13	10.02	7.65	0.19	6.20	3.25	0.06
τ = 0.9												
Intercept	295.01	19.00	0.00	363.76	22.01	0.00	365.44	26.80	0.00	298.71	16.23	0.00
Race	62.98	17.41	0.00	92.03	16.02	0.00	87.72	18.31	0.00	59.05	19.32	0.00
Insurance type	10.13	6.99	0.15	6.17	9.22	0.50	4.86	9.08	0.59	7.41	5.36	0.17
Marital status	17.79	4.67	0.00	17.52	7.47	0.02	17.78	6.84	0.01	15.64	4.82	0.00
First language	10.24	14.90	0.49	19.41	15.98	0.23	23.73	18.04	0.19	14.41	13.36	0.28
Age	-0.96	0.28	0.00	-1.39	0.32	0.00	-1.39	0.39	0.00	-0.92	0.23	0.00
Gender	5.87	5.08	0.25	8.55	8.35	0.31	8.58	7.62	0.26	5.79	5.24	0.27

Table 17: The estimated coefficients (Est) and the corresponding standard errors (SE) and p-values (p) from different methods with $\tau = (0.1, 0.25, 0.5, 0.75, 0.9)$, $n = 7148$ and 60% missing in response y_i

	fullQr			NaiveQr			elsIpwQr			proIpwQr		
Covariate	Est	SE	p	Est	SE	p	Est	SE	p	Est	SE	p
						τ = 0.1						
Intercept	92.07	2.35	0.00	83.59	3.79	0.00	82.22	4.24	0.00	91.89	3.04	0.00
Race	-2.18	2.23	0.33	-5.50	3.23	0.09	-5.95	3.78	0.12	-1.83	2.52	0.47
Insurance type	-4.38	1.37	0.00	-4.16	2.39	0.08	-4.68	2.29	0.04	-3.95	1.29	0.00
Marital status	0.67	1.20	0.58	4.24	2.17	0.05	4.35	2.50	0.08	0.55	1.40	0.69
First language	-0.83	2.71	0.76	-1.68	6.32	0.79	-1.85	6.80	0.79	-3.40	4.70	0.47
Age	0.16	0.04	0.00	0.11	0.06	0.07	0.14	0.07	0.05	0.16	0.04	0.00
Gender	1.69	1.36	0.21	-0.70	2.27	0.76	-0.68	1.90	0.72	3.36	1.79	0.06
						τ = 0.25						
Intercept	109.32	3.26	0.00	96.62	3.38	0.00	95.98	3.28	0.00	108.43	7.35	0.00
Race	0.07	2.11	0.97	2.27	4.50	0.62	1.51	4.69	0.75	2.69	5.00	0.59
Insurance type	-2.46	1.61	0.13	-4.60	2.17	0.03	-4.79	2.02	0.02	-1.63	2.27	0.47
Marital status	1.07	1.35	0.43	2.75	1.72	0.11	2.92	2.05	0.15	-0.10	2.30	0.97
First language	4.25	4.90	0.39	2.20	7.97	0.78	0.82	8.19	0.92	11.54	6.18	0.06
Age	0.15	0.05	0.00	0.16	0.05	0.00	0.17	0.05	0.00	0.19	0.10	0.07
Gender	4.66	1.22	0.00	5.28	1.80	0.00	6.20	2.26	0.01	6.56	2.38	0.01
						τ = 0.5						
Intercept	138.92	3.52	0.00	146.11	13.77	0.00	143.29	11.88	0.00	157.24	11.31	0.00
Race	10.33	3.73	0.01	38.64	12.85	0.00	36.48	11.85	0.00	18.31	6.03	0.00
Insurance type	-0.55	2.04	0.79	6.62	6.81	0.33	6.14	6.26	0.33	8.21	5.06	0.11
Marital status	1.27	1.73	0.46	5.93	5.14	0.25	7.24	6.12	0.24	0.47	4.43	0.92
First language	10.94	4.07	0.01	15.07	14.07	0.28	14.32	16.12	0.37	7.51	9.61	0.44
Age	0.11	0.06	0.08	-0.20	0.24	0.38	-0.17	0.20	0.40	-0.09	0.18	0.59
Gender	5.28	1.77	0.00	19.53	7.23	0.01	20.54	7.84	0.01	16.06	3.94	0.00
						τ = 0.75						
Intercept	191.98	6.70	0.00	272.12	18.20	0.00	278.29	18.65	0.00	225.83	10.18	0.00
Race	27.00	4.52	0.00	46.57	13.40	0.00	39.50	14.84	0.01	22.04	5.00	0.00
Insurance type	5.87	3.43	0.09	8.25	8.48	0.33	10.55	9.35	0.26	9.64	5.88	0.10
Marital status	6.64	2.59	0.01	14.54	6.17	0.02	12.76	7.57	0.09	0.32	5.74	0.96
First language	12.18	5.38	0.02	4.81	11.04	0.66	4.36	12.04	0.72	9.71	7.43	0.19
Age	-0.21	0.11	0.06	-0.63	0.27	0.02	-0.73	0.29	0.01	-0.32	0.16	0.05
Gender	6.15	2.50	0.01	8.56	6.37	0.18	8.39	6.66	0.21	9.55	5.43	0.08
						τ = 0.9						
Intercept	295.01	17.77	0.00	386.94	22.32	0.00	385.32	26.00	0.00	298.51	14.78	0.00
Race	62.98	15.29	0.00	95.88	20.35	0.00	94.62	16.55	0.00	49.20	15.57	0.00
Insurance type	10.13	6.05	0.09	5.26	12.33	0.67	4.56	13.43	0.73	11.44	6.25	0.07
Marital status	17.79	5.25	0.00	21.06	11.03	0.06	17.25	10.14	0.09	11.38	5.28	0.03
First language	10.24	13.37	0.44	3.41	24.66	0.89	2.70	21.58	0.90	3.11	12.57	0.80
Age	-0.96	0.26	0.00	-1.19	0.37	0.00	-1.16	0.40	0.00	-0.73	0.22	0.00
Gender	5.87	5.38	0.28	-1.66	10.15	0.87	1.30	10.03	0.90	0.68	5.80	0.91

Table 18: The estimates of $\boldsymbol{\theta}$ from different methods with n = 18744 and 61.8% missing in response y_i

Method	Intercept	Race	Insurance type	Marital status	First language	Age	Glucose
elsIpwQr	-1.255	0.120	0.115	-0.344	-0.723	0.014	0.000
proIpwQr	-2.499	0.070	0.021	-0.218	-0.768	0.014	0.008

Table 19: The estimates of $\boldsymbol{\theta}_0 = (-0.6, 0, 0, 0, 0, 0, 2.7)^\top$ from different methods with $n = 7148$ and 40% artificially missing in response y_i

Method	Intercept	Race	Insurance type	Marital status	First language	Age	$\tilde{Y}^2$
elsIpwQr	0.355	0.024	-0.140	0.246	-0.015	0.001	-0.002
proIpwQr	-0.508	-0.028	0.000	-0.051	0.120	0.002	2.307

Table 20: The estimates of $\boldsymbol{\theta}_0 = (-1.7, 0, 0, 0, 0, 0, 2.7)^\top$ from different methods with $n = 7148$ and 60% artificially missing in response y_i

Method	Intercept	Race	Insurance type	Marital status	First language	Age	$\tilde{Y}^2$
elsIpwQr	0.358	0.000	0.178	-0.519	0.099	-0.010	0.005
proIpwQr	-1.488	0.114	0.105	0.147	0.191	-0.003	3.142