PARTIALLY FUNCTIONAL LINEAR QUANTILE REGRESSION WITH MEASUREMENT ERRORS

Mengli Zhang, Lan Xue, Carmen D. Tekwe, Yang Bai and Annie Qu

Shanghai University of Finance and Economics, Oregon State University,

Indiana University School of Public Health and University of California, Irvine

Supplementary Material

This supplement material contains necessary lemmas, detailed proofs of the main theorems and additional simulation results.

S1 Technical Lemmas and Proofs

Lemma 1. *Under assumptions (A1)-(A3), (A6)* $m^{2\alpha_x+2} \leq cn$, $nm^{2\alpha_x}\Delta^{\nu_0}$*=O(1) as* $n \to \infty$*, we have, for* $i = 1, \ldots, n$*,* $j = 1, \ldots, m$*,*

$$
\begin{aligned}
|||\hat{K}_x - K_x|||^2 &= O_p(\Delta_n^{\nu_0}), \\
\|\hat{\phi}_j - \phi_j\|^2 &= O_p(j^{2\alpha_x+2}\Delta_n^{\nu_0} + j^2 n^{-1}), \\
(\hat{\mathbf{W}}_i - \mathbf{W}_i)^T \mathbf{b}_0 &= O_p(\frac{\sqrt{m}}{\sqrt{n}}), \\
\mathbf{b}_0^T(\hat{\Sigma}_u - \Sigma_u)\mathbf{b}_0 &= O_p(\frac{\sqrt{m}}{\sqrt{n}}).
\end{aligned}
$$

Proof. Here we assume $m^{2\alpha_x+2} \leq cn$, which plays the same role as $m = n^{1/(\alpha_x+2\beta)}$ in (Kato, 2012). Thus, following the same calculations as in supplement of (Kato, 2012), the first two equations in our lemma hold. The difference between $\hat{\mathbf{W}}_i$ and $\mathbf{W}_i$ consists of two parts. One is caused by smoothing observed functions $W_i(t)$. The other is cause by projecting $W_i(t)$ on estimated $\hat{\phi}_j(t)$ instead of $\phi_j(t)$. After simple calculations, the third equation is proved. Similarly, we can get the fourth equation. □

Lemma 2. *Let* $\rho_h^*(\mathbf{W}_i, \boldsymbol{\gamma})$=$\rho_h^*(Y_i - \mathbf{W}_i^T\mathbf{b}(\tau) - \mathbf{Z}_i^T\boldsymbol{\theta}(\tau), \mathbf{b}^T\Sigma_u\mathbf{b})$ *and* $l_i^*(\gamma) = \rho_h^*(\mathbf{W}_i, \boldsymbol{\gamma}) - E\rho_h^*(\mathbf{W}_i, \boldsymbol{\gamma})$*, where* $\boldsymbol{\gamma} = (\mathbf{b}^T(\tau), \boldsymbol{\theta}^T(\tau))^T$*. Under assumptions (A1)-(A2), (A4)-(A5), (A7)-(A8) and* $m^{2\alpha_x+2} \leq cn$*, we have, for any constant* $c > 0$*,*

$$\sup_{\|\mathbf{b}\|\leq c, \boldsymbol{\theta}\in\Theta} \frac{1}{n}\left|\sum_{i=1}^n l_i^*(\boldsymbol{\gamma})\right| = O_p\left\{h^{-1}exp(ch^{-2})n^{-1/2}\sqrt{mlog(m+n)}\right\}.$$

Proof. Let $\Gamma = \{\boldsymbol{\gamma} : \|\mathbf{b}\| \leq c, \boldsymbol{\theta} \in \Theta\}$. Assumptions required entail that, for any $\boldsymbol{\gamma} \in \Gamma$,

$$\begin{aligned}
\mathrm{E}\left[\rho_h^*(\mathbf{W}_i, \boldsymbol{\gamma})\right]^2 &\leq c\mathrm{E}\left\{\int_0^{1/h}\left(e_i^2 + \sigma^2\right) exp(t^2\sigma^2/2)dt\right\}^2 \\
&\leq ch^{-2}exp(\sigma^2h^{-2})\mathrm{E}[e_i^2 + \sigma^2]^2 \leq ch^{-2}exp(ch^{-2}),
\end{aligned}$$

where $\sigma^2 = \mathbf{b}^T\Sigma_u\mathbf{b}$, $e_i = Y_i - \mathbf{W}_i^T\mathbf{b} - \mathbf{Z}_i^T\boldsymbol{\theta}$. Let $V_n = \sqrt{\sum_{i=1}^n \mathrm{E}[l_i^*(\gamma)]^2} +$

$\sqrt{\sum_{i=1}^{n}[l_i^*(\gamma)]^2}$. Then $V_n = O_p(h^{-1}exp(ch^{-2})n^{-1/2})$. To prove Lemma 2, it suffices to show that, for sufficiently large $c > 0$,

$$P\left(\sup_{\boldsymbol{\gamma}\in\Gamma}\frac{|\sum_{i=1}^{n} l_i^*(\gamma)|}{V_n\sqrt{mlog(m+n)}} \geq c\right) \to 0, as\ n\ \to +\infty. \tag{S1.1}$$

The proof of equation (S1.1) makes use of the continuity property of $l_i^*(\boldsymbol{\gamma})$ that there exists $r_1 > 0$ such that for every $\boldsymbol{\gamma}_1 \in \Gamma$, $0 < d \leq 1$, $1 \leq i \leq n$,

$$\mathrm{E}\sup_{\|\boldsymbol{\gamma}_2-\boldsymbol{\gamma}_1\|\leq d}|l_i^*(\boldsymbol{\gamma}_2) - l_i^*(\boldsymbol{\gamma}_1)| \leq n^{r_1}d,$$

which is easily derived by calculating upper bound of $\|\frac{\partial\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma})}{\partial\boldsymbol{\gamma}}\|$. Now consider a partition of Γ with cubes $\mathrm{S}_1, ..., \mathrm{S}_M$ such that $\cup_{j=1}^{M}\mathrm{S}_j = \Gamma$ and the length of edge of each cube, δ_n, satisfies $\delta_n = n^{-(10+r_1)}/(m+p)$ or $M \leq (\frac{2c}{\delta_n})^{m+p}$. Then

$$\begin{aligned} & P\left(\sup_{\boldsymbol{\gamma}\in\Gamma}\frac{|\sum_{i=1}^{n} l_i^*(\boldsymbol{\gamma})|}{V_n\sqrt{mlog(m+n)}} \geq c\right) \\ \leq\ & \sum_{j=1}^{M} P\left(\sup_{\boldsymbol{\gamma}\in\mathrm{S}_j}\frac{|\sum_{i=1}^{n} l_i^*(\boldsymbol{\gamma})|}{V_n\sqrt{mlog(m+n)}} \geq c\right). \end{aligned} \tag{S1.2}$$

Let $\mathbf{c}_j$ be the center of S_j. Consider $\eth_i(j) = \{\sup_{\boldsymbol{\gamma}\in\mathrm{S}_j}|l_i^*(\boldsymbol{\gamma}) - l_i^*(\mathbf{c}_j)| \leq n^{-4}\}$, $l_i^*(j) = l_i^*(\mathbf{c}_j)I\{\eth_i(j)\}$, $l_i^{**}(j) = l_i^*(j) - \mathrm{E}l_i^*(j)$. For $\boldsymbol{\gamma} \in \mathrm{S}_j$, Hölder inequality

and the continuity of $l_i^*(\boldsymbol{\gamma})$ entail that

$$|\mathrm{E}[l_i^*(\boldsymbol{\gamma})I\{\eth_i^{\mathrm{C}}(j)\}]| \leq \{\mathrm{E}[l_i^*(\boldsymbol{\gamma})]^2\}^{1/2}\{P[\eth_i^{\mathrm{C}}(j)]\}^{1/2} \leq \{n^{-6}\mathrm{E}[l_i^*(\gamma)]^2\}^{1/2}, \quad \text{(S1.3)}$$

and observe that

$$\begin{aligned}
& \frac{|\sum_{i=1}^n l_i^*(\boldsymbol{\gamma})|}{V_n} \\
\leq\ & \frac{\left|\sum_{i=1}^n l_i^*(\boldsymbol{\gamma})I\{\eth_i^{\mathrm{C}}(j)\}\right| + |\sum_{i=1}^n l_i^*(\boldsymbol{\gamma})I\{\eth_i(j)\}|}{V_n} \\
\leq\ & \sqrt{\sum_{i=1}^n I\{\eth_i^{\mathrm{C}}(j)\}} + \frac{|\sum_{i=1}^n l_i^*(\boldsymbol{\gamma})I\{\eth_i(j)\} - \mathrm{E}[l_i^*(\boldsymbol{\gamma})I\{\eth_i(j)\}]| + \sum_{i=1}^n \left|\mathrm{E}[l_i^*(\boldsymbol{\gamma})I\{\eth_i^{\mathrm{C}}(j)\}]\right|}{V_n} \\
\leq\ & \sqrt{\sum_{i=1}^n I\{\eth_i^{\mathrm{C}}(j)\}} + n^{-5/2} + \frac{|\sum_{i=1}^n l_i^*(\boldsymbol{\gamma})I\{\eth_i(j)\} - \mathrm{E}[l_i^*(\boldsymbol{\gamma})I\{\eth_i(j)\}]|}{V_n}. \qquad \text{(S1.4)}
\end{aligned}$$

Moreover, notice that $l_i^*(\boldsymbol{\gamma}) = l_i^*(\boldsymbol{\gamma}) - l_i^*(\mathbf{c}_j) + l_i^*(\mathbf{c}_j)$ and we have

$$\begin{aligned}
& \left|\sum_{i=1}^n l_i^*(\boldsymbol{\gamma})I\{\eth_i(j)\} - \mathrm{E}[l_i^*(\boldsymbol{\gamma})I\{\eth_i(j)\}]\right| \\
\leq\ & \left|\sum_{i=1}^n [l_i^*(\boldsymbol{\gamma}) - l_i^*(\mathbf{c}_j)]I\{\eth_i(j)\} - \mathrm{E}[l_i^*(\boldsymbol{\gamma}) - l_i^*(\mathbf{c}_j)]I\{\eth_i(j)\}\right| \\
+\ & \left|\sum_{i=1}^n l_i^*(\mathbf{c}_j)I\{\eth_i(j)\} - \mathrm{E}[l_i^*(\mathbf{c}_j)I\{\eth_i(j)\}]\right| \\
\leq\ & 2n^{-3} + \left|\sum_{i=1}^n l_i^{**}(j)\right|. \qquad \text{(S1.5)}
\end{aligned}$$

By (S1.3), (S1.4) and (S1.5), we have

$$\frac{|\sum_{i=1}^n l_i^*(\boldsymbol{\gamma})|}{V_n} \leq \frac{2n^{-3} + |\sum_{i=1}^n l_i^{**}(j)|}{V_n} + \sqrt{\sum_{i=1}^n I\{\eth_i^{\mathrm{C}}(j)\}} + n^{-5/2}.$$

Similarly for V_n in the denominator, one has $V_n^*[1+o_p(1)] \leq 4\sqrt{3}V_n$, where $V_n^* = \sqrt{\sum_i \mathrm{E}[l_i^{**}(j)]^2} + \sqrt{\sum_i [l_i^{**}(j)]^2}$ and $V_n^* = O_p(V_n)$. Finally, we have

$$\frac{|\sum_{i=1}^n l_i^*(\boldsymbol{\gamma})|}{V_n} \leq \frac{|\sum_{i=1}^n l_i^{**}(j)| + 2n^{-3}}{(4\sqrt{3})^{-1}V_n^*[1+o_p(1)]} + \sqrt{\sum\nolimits_{i=1}^n I\{\eth_i^{\mathrm{C}}(j)\}} + n^{-5/2}. \quad \text{(S1.6)}$$

Thus, the inequality (S1.2) follows

$$\begin{aligned}
& P\left(\sup_{\boldsymbol{\gamma}\in\Gamma} \frac{|\sum_{i=1}^n l_i^*(\boldsymbol{\gamma})|}{V_n\sqrt{mlog(m+n)}} \geq c\right) \\
\leq\ & \sum_{j=1}^M P\left(\frac{|\sum_{i=1}^n l_i^{**}(j)| + 2n^{-3}}{(4\sqrt{3})^{-1}V_n^*[1+o_p(1)]} + n^{-5/2} \geq c\sqrt{mlog(m+n)}\right) \\
+\ & \sum_{j=1}^M P\left(\sqrt{\sum\nolimits_{i=1}^n I\{\eth_i^{\mathrm{C}}(j)\}} \geq c\sqrt{mlog(m+n)}\right) \\
\leq\ & \sum_{j=1}^M P\left(\frac{|\sum_{i=1}^n l_i^{**}(j)|}{(4\sqrt{3})^{-1}V_n^*} \geq c\sqrt{mlog(m+n)}\right) \\
+\ & \sum_{j=1}^M P\left(\sqrt{\sum\nolimits_{i=1}^n I\{\eth_i^{\mathrm{C}}(j)\}} \geq c\sqrt{mlog(m+n)}\right),
\end{aligned}$$

where the last equation holds for large n. Following Lemma 3.1 in (He and Shao, 2000) and exponential inequality for binomial distribution ((Ledoux and Talagrand, 1991), p.51), we have

$$\begin{aligned}
& \sum_{j=1}^M P\left(\frac{|\sum_{i=1}^n l_i^{**}(j)|}{(4\sqrt{3})^{-1}V_n^*} \geq c\sqrt{mlog(m+n)}\right) \\
\leq\ & \sum_{j=1}^M Cexp(-16C^{-1}c^2 mlog(m+n)), \qquad \text{(S1.7)}
\end{aligned}$$

where C is a constant, and

$$\begin{aligned} & \sum_{j=1}^{M} P\left(\sqrt{\sum\nolimits_{i=1}^{n} I\{\eth_i^{\mathrm{C}}(j)\}} \geq c\sqrt{mlog(m+n)}\right) \\ \leq & \sum_{j=1}^{M} \left[\frac{3\sum_{i=1}^{n} P(\eth_i^{\mathrm{C}}(j))}{c^2 mlog(m+n)}\right]^{c^2 mlog(m+n)}. \end{aligned} \tag{S1.8}$$

The upper bounds of inequality (S1.7) and (S1.8) go to zero after some calculations when c is sufficiently large and n goes to infinity, which proves equation (S1.1) and completes our proof. □

Lemma 3. *Under the same assumptions of Lemma 2, similarly denote* $l_i(\boldsymbol{\gamma}) = [\rho_h(\mathbf{X}_i, \boldsymbol{\gamma}) - E\rho_h(\mathbf{X}_i, \boldsymbol{\gamma})]/n$*, where* $\rho_h(\mathbf{X}_i, \boldsymbol{\gamma}) = \rho_h(Y_i - \mathbf{X}_i^T \mathbf{b}(\tau) - \mathbf{Z}_i^T \boldsymbol{\theta}(\tau))$*. As* $n \to \infty$*, we have*

$$\sup_{\boldsymbol{\gamma} \in B \otimes \Theta} |\sum_{i=1}^{n} l_i(\boldsymbol{\gamma})| = O_p(n^{-1/2}\sqrt{mlog(m+n)}).$$

Proof. The proof of Lemma 3 is very similar to proof of Lemma 2. So we omit it here. □

S1.1 Proof of Theorem 1

One notes that

$$\begin{aligned}
&\int_0^1 \{\hat{\beta}(t,\tau) - \beta_0(t,\tau)\}^2 dt \\
= &\int_0^1 \left\{ \sum_{j=1}^{m} (\hat{b}_j - b_{0j})(\tau)\hat{\phi}_j(t) + \sum_{j=1}^{m} b_{0j}(\tau)(\hat{\phi}_j(t) - \phi_j(t)) - \sum_{j=m+1}^{+\infty} b_{0j}(\tau)\phi_j(t) \right\}^2 dt \\
\leq &\ 3\|\hat{\mathbf{b}}(\tau) - \mathbf{b}_0(\tau)\|^2 + 3m \sum_{j=1}^{m} b_{0j}^2(\tau)\|\hat{\phi}_j - \phi_j\|^2 + 3 \sum_{j=m+1}^{\infty} b_{0j}^2(\tau) \\
= &\ \mathrm{I} + \mathrm{II} + \mathrm{III}. \qquad \text{(S1.9)}
\end{aligned}$$

Suppose $m^{2\alpha_x+2}n^{-1} = O(1)$ and $nm^{2\alpha_x}\Delta^{\nu_0} = O(1)$. Then assumption (A5) and Lemma 1 entail that

$$\begin{aligned}
\mathrm{II} &= O_p(m^{-2\beta+2\alpha_x+4}\Delta^{\nu_0} + mn^{-1}) = O_p(mn^{-1}), \\
\mathrm{III} &= O(m^{-2\beta+1}).
\end{aligned}$$

Let $\boldsymbol{\gamma} = (\mathbf{b}^T(\tau), \boldsymbol{\theta}^T(\tau))^T$ be the vector of unknown parameters at τth quantile. Correspondingly, denote the vector of true coefficients as $\boldsymbol{\gamma}_0 = (\mathbf{b}_0^T(\tau), \boldsymbol{\theta}_0^T(\tau))^T$. Theorem 1 follows if we can show that $\|\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|^2 = O_p(a_n^2)$ with

$$a_n^2 = m^{-2\beta+1} + h^{-1}exp(ch^{-2})m^{\alpha_x+1/2}[log(m+n)]^{1/2}n^{-1/2} + hm^{\alpha_x}. \quad \text{(S1.10)}$$

It is sufficient to show for some constant $C > 0$,

$$P\left(\inf_{\|\boldsymbol{\gamma}-\boldsymbol{\gamma}_0\|=Ca_n} \mathrm{E}_n\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma}) > \mathrm{E}_n\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma}_0)\right) \to 1, \tag{S1.11}$$

where $\mathrm{E}_n\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma})$ is the objective function. Here with some abuse of notation $\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma})$ is used to emphasize the dependence of the function ρ_h^* on $\mathbf{W}_i$ and $\boldsymbol{\gamma}$. We have following decomposition,

$$\begin{aligned}
& \inf_{\|\boldsymbol{\gamma}-\boldsymbol{\gamma}_0\|=Ca_n} \mathrm{E}_n\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma}) - \mathrm{E}_n\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma}_0) \\
= & \inf_{\|\boldsymbol{\gamma}-\boldsymbol{\gamma}_0\|=Ca_n} \{\mathrm{E}_n\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma}) - \mathrm{E}_n\rho_\tau(\mathbf{X}_i,\boldsymbol{\gamma})\} + \inf_{\|\boldsymbol{\gamma}-\boldsymbol{\gamma}_0\|=Ca_n} \mathrm{E}_n\rho_\tau(\mathbf{X}_i,\boldsymbol{\gamma}) - \mathrm{E}_n\rho_\tau(\mathbf{X}_i,\boldsymbol{\gamma}_0) \\
& + \mathrm{E}_n\rho_\tau(\mathbf{X}_i,\boldsymbol{\gamma}_0) - \mathrm{E}_n\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma}_0) \\
= & \ \mathrm{IV} + \mathrm{V} + \mathrm{VI}.
\end{aligned} \tag{S1.12}$$

In addition, one notes that

$$\begin{aligned}
& \sup_{\|\boldsymbol{\gamma}-\boldsymbol{\gamma}_0\|\le Ca_n} |\mathrm{E}_n\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma}) - \mathrm{E}_n\rho_\tau(\mathbf{X}_i,\boldsymbol{\gamma})| \\
\le & \sup_{\|\boldsymbol{\gamma}-\boldsymbol{\gamma}_0\|\le Ca_n} |\mathrm{E}_n\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma}) - \mathrm{E}_n\rho_h(\mathbf{X}_i,\boldsymbol{\gamma})| + \sup_{\|\boldsymbol{\gamma}-\boldsymbol{\gamma}_0\|\le Ca_n} |\mathrm{E}_n\rho_h(\mathbf{X}_i,\boldsymbol{\gamma}) - \mathrm{E}_n\rho_\tau(\mathbf{X}_i,\boldsymbol{\gamma})| \\
= & \ O_p(h^{-1}exp(ch^{-2})\sqrt{mlog(m+n)}n^{-1/2}) + O_p(h),
\end{aligned} \tag{S1.13}$$

by Lemmas 2 and 3, and equations (A9) and (A10) in (Wang et al., 2012). Thus, the terms IV and VI in (S1.12) are of order $O_p\left\{h^{-1}exp(ch^{-2})\sqrt{mlog(m+n)}n^{-1/2} + h\right\}$.

Let $\Delta\varepsilon_i = (\mathbf{X}_i^T, \mathbf{Z}_i^T)(\boldsymbol{\gamma} - \boldsymbol{\gamma}_0)$. Now for term V in (S1.12), Knight's identity gives that,

$$\begin{aligned}
\mathrm{V} &= -\mathrm{E}_n\{\Delta\varepsilon_i[\tau - I(\varepsilon_{i0} < 0)]\} + \mathrm{E}_n\left\{\int_0^{\Delta\varepsilon_i}[I(\varepsilon_{i0} < s) - I(\varepsilon_{i0} < 0)]ds\right\} \\
&= \left\{-\mathrm{E}[\Delta\varepsilon_i(\tau - I(\varepsilon_{i0} < 0))] + \mathrm{E}\int_0^{\Delta\varepsilon_i}[I(\varepsilon_{i0} < s) - I(\varepsilon_{i0} < 0)]ds\right\}[1 + o_p(1)] \\
&= (\mathrm{E}_1 + \mathrm{E}_2)[1 + o_p(1)].
\end{aligned}$$

Let $Q_i = Q_\tau(Y_i|X_i(t), Z_i)$ and $T_i = \sum_{j=m+1}^{\infty} b_{0j}X_{ij}$. Assumptions (A2) and (A5) entail that $\mathrm{E}(T_i)^2 \leq cm^{-2\beta-\alpha_x+1}$. We have

$$\begin{aligned}
\mathrm{E}_1 &= -\mathrm{E}\left\{\Delta\varepsilon_i[F_{Y|X(t),Z}(Q_i) - F_{Y|X(t),Z}(Q_i - T_i)]\right\} \\
&= -\mathrm{E}\left[\Delta\varepsilon_i f_{Y|X(t),Z}(Q_i)T_i\right](1 + o(1)) \\
&\leq c\sqrt{\mathrm{E}(\Delta\varepsilon_i)^2}\sqrt{\mathrm{E}T_i^2} \\
&\leq c\sqrt{m^{-2\beta-\alpha_x+1}}\sqrt{\mathrm{E}(\Delta\varepsilon_i)^2}. \qquad \text{(S1.14)}
\end{aligned}$$

For E_2, we have

$$\begin{aligned}
\mathrm{E}_2 &= \mathrm{E}\int_0^{\Delta\varepsilon_i}[F_{Y|X(t),Z}(s + Q_i - T_i) - F_{Y|X(t),Z}(Q_i - T_i)]ds \\
&= \left\{\mathrm{E}\int_0^{\Delta\varepsilon_i} f_{Y|X(t),Z}(Q_i - T_i)sds\right\}[1 + o(1)]
\end{aligned}$$

$$\begin{aligned} &\geq c\mathrm{E}\left[f_{Y|X(t),Z}(Q_i - T_i)(\Delta\varepsilon_i)^2\right] \\ &\geq c\mathrm{E}(\Delta\varepsilon_i)^2. \end{aligned} \tag{S1.15}$$

Combining results in (S1.13), (S1.14) and (S1.15) into error decomposition in (S1.12), one has

$$\begin{aligned} &\inf_{\|\boldsymbol{\gamma}-\boldsymbol{\gamma}_0\|=Ca_n} \mathrm{E}_n\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma}) - \mathrm{E}_n\rho_h^*(\mathbf{W}_i,\boldsymbol{\gamma}_0) \\ \geq\ & c\sqrt{\mathrm{E}(\Delta\varepsilon_i)^2}\left(\sqrt{\mathrm{E}(\Delta\varepsilon_i)^2} - c\sqrt{m^{-2\beta-\alpha_x+1}}\right) \\ +\ & O_p(h^{-1}exp(ch^{-2})\sqrt{mlog(m+n)}n^{-1/2} + h). \end{aligned}$$

Here $\mathrm{E}(\Delta\varepsilon_i)^2 = (\boldsymbol{\gamma}-\boldsymbol{\gamma}_0)^T\Lambda\mathbf{A}\Lambda(\boldsymbol{\gamma}-\boldsymbol{\gamma}_0)$, with $\Lambda = \mathrm{diag}\left(\kappa_{x1}^{1/2},\ldots,\kappa_{xm}^{1/2},1,\ldots,1\right)$ is a diagonal matrix with dimension $m+p$. Therefore assumptions (A2) and (A3) give that $\mathrm{E}(\Delta\varepsilon_i)^2 \geq m^{-\alpha_x}\|\boldsymbol{\gamma}-\boldsymbol{\gamma}_0\|^2\lambda_{min}(A)$. Therefore, (S1.11) follows by noting that the first term is the dominating term uniformly in $\|\boldsymbol{\gamma}-\boldsymbol{\gamma}_0\| = Ca_n$ as long as C is large enough. □

S1.2 Proof of Theorem 2

For simplicity, we replaced $\hat{\mathbf{W}}_i$ and $\hat{\Sigma}_u$ with their true values $\mathbf{W}_i$ and Σ_u respectively in the objective function. However, Theorems 1 and 2 still hold if the estimated values are used by $(\hat{\mathbf{W}}_i - \mathbf{W}_i)^T\mathbf{b}_0 = O_p(\frac{\sqrt{m}}{\sqrt{n}})$ and $\mathbf{b}_0^T(\hat{\Sigma}_u - \Sigma_u)\mathbf{b}_0 = O_p(\frac{\sqrt{m}}{\sqrt{n}})$ in Lemma 1. Denote $\hat{\boldsymbol{\theta}}_1^* = \sqrt{n}\{\mathrm{E}_n[\mathbf{Z}_i^*\mathbf{Z}_i^{*T}\frac{\partial^2\rho_h^*}{\partial\varepsilon^2}(\varepsilon_{i0}-\mathbf{U}_i^T\mathbf{b}_0, \mathbf{b}_0^T\Sigma_u\mathbf{b}_0)]\}^{-1}$

$\mathrm{E}_n[\tilde{\mathbf{Z}}_i \frac{\partial \rho_h^*}{\partial \varepsilon}(\varepsilon_{i0} - \mathbf{U}_i^T \mathbf{b}_0, \mathbf{b}_0^T \Sigma_u \mathbf{b}_0)]$. Obviously, it converges to $\mathrm{N}(\mathbf{0}, \mathbf{D^{-1}BD^{-1}})$ according to Theorem 4. in (Wang et al., 2012). Thus we only need to prove $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = \hat{\boldsymbol{\theta}}_1^* + o_p(1)$. One notes that the objective function can be rewritten as

$$\mathrm{E}_n[\rho_h^*(\varepsilon_{i0} + T_i - \frac{1}{\sqrt{n}} Z_i^{*T} \boldsymbol{\theta}_1 - \mathbf{X}_i^T(\mathbf{b} - \mathbf{b}_0) - (\mathbf{Z}_i - \mathbf{Z}_i^*)^T(\boldsymbol{\theta} - \boldsymbol{\theta}_0) - \mathbf{U}_i^T \mathbf{b}, \mathbf{b}^T \Sigma_u \mathbf{b})],$$

where $\boldsymbol{\theta}_1 = \sqrt{n}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$, $T_i = \sum_{j=m+1}^{\infty} b_{0j} X_{ij}$. Note that $\hat{\boldsymbol{\theta}}_1 = \sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ is the minimum of $\mathrm{E}_n[\rho_h^*(\varepsilon_{i0} + T_i - \frac{1}{\sqrt{n}} Z_i^{*T} \boldsymbol{\theta}_1 - \mathbf{X}_i^T(\hat{\mathbf{b}} - \mathbf{b}_0) - (\mathbf{Z}_i - \mathbf{Z}_i^*)^T(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) - \mathbf{U}_i^T \hat{\mathbf{b}}, \hat{\mathbf{b}}^T \Sigma_u \hat{\mathbf{b}})]$. For simplicity, we define $\mathbf{D}_i = (\mathbf{X}_i^T, (\mathbf{Z}_i - \mathbf{Z}_i^*)^T)^T$ and $\mathbf{d}(\mathbf{b}, \boldsymbol{\theta}) = ((\mathbf{b} - \mathbf{b}_0)^T, (\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T)^T$. To prove our theorem, it suffices to prove that, for any $\delta > 0$,

$$\begin{aligned} P\Bigg(\inf_{\|\boldsymbol{\theta}_1 - \hat{\boldsymbol{\theta}}_1^*\| = \delta} & \mathrm{E}_n[\rho_h^*(\varepsilon_{i0} + T_i - \frac{1}{\sqrt{n}} Z_i^{*T} \boldsymbol{\theta}_1 - \mathbf{D}_i^T \mathbf{d}(\hat{\mathbf{b}}, \hat{\boldsymbol{\theta}}) - \mathbf{U}_i^T \hat{\mathbf{b}}, \hat{\mathbf{b}}^T \Sigma_u \hat{\mathbf{b}})] \\ & > \mathrm{E}_n[\rho_h^*(\varepsilon_{i0} + T_i - \frac{1}{\sqrt{n}} Z_i^{*T} \hat{\boldsymbol{\theta}}_1^* - \mathbf{D}_i^T \mathbf{d}(\hat{\mathbf{b}}, \hat{\boldsymbol{\theta}}) - \mathbf{U}_i^T \hat{\mathbf{b}}, \hat{\mathbf{b}}^T \Sigma_u \hat{\mathbf{b}})] \Bigg) \to 1. \end{aligned} \tag{S1.16}$$

As proved in Theorem 1, $\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\| = O_p(a_n)$ and $\|\hat{\mathbf{b}} - \mathbf{b}_0\| = O_p(a_n)$ with a_n defined in (S1.10) . Thus, by assumption (A10), equation (S1.16) follows from

following equation

$$\begin{aligned}
\sup_{\|\boldsymbol{\theta}_1-\hat{\boldsymbol{\theta}}_1^*\|=\delta} \quad & I(\hat{\boldsymbol{\theta}},\hat{\mathbf{b}})\Big|\mathrm{E}_n[\rho_h^*(\varepsilon_{i0}+T_i-\frac{1}{\sqrt{n}}Z_i^{*T}\boldsymbol{\theta}_1-\mathbf{D}_i^T\mathbf{d}(\hat{\mathbf{b}},\hat{\boldsymbol{\theta}})-\mathbf{U}_i^T\hat{\mathbf{b}},\hat{\mathbf{b}}^T\Sigma_u\hat{\mathbf{b}})] \\
- \quad & \mathrm{E}_n[\rho_h^*(\varepsilon_{i0}+T_i-\frac{1}{\sqrt{n}}Z_i^{*T}\hat{\boldsymbol{\theta}}_1^*-\mathbf{D}_i^T\mathbf{d}(\hat{\mathbf{b}},\hat{\boldsymbol{\theta}})-\mathbf{U}_i^T\hat{\mathbf{b}},\hat{\mathbf{b}}^T\Sigma_u\hat{\mathbf{b}})] \\
- \quad & \frac{1}{2n}(\boldsymbol{\theta}_1-\hat{\boldsymbol{\theta}}_1^*)^T\mathrm{E}_n[\mathbf{Z}_i^*\mathbf{Z}_i^{*T}\frac{\partial^2\rho_h^*}{\partial\varepsilon^2}(\varepsilon_{i0}-\mathbf{U}_i^T\mathbf{b}_0,\mathbf{b}_0^T\Sigma_u\mathbf{b}_0)](\boldsymbol{\theta}_1-\hat{\boldsymbol{\theta}}_1^*)\Big| \\
= \quad & o_p(\frac{1}{n}),
\end{aligned} \tag{S1.17}$$

where $I(\hat{\boldsymbol{\theta}},\hat{\mathbf{b}})=I(\|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\|\le c_1a_n,\|\hat{\mathbf{b}}-\mathbf{b}_0\|\le c_2a_n)$. By the definition of $\hat{\boldsymbol{\theta}}_1^*$, the left side of equation (S1.17) is less than or equal to

$$\begin{aligned}
2\sup_{\substack{\|\boldsymbol{\theta}_1\|\le c\\ \|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\le c_1a_n\\ \|\mathbf{b}-\mathbf{b}_0\|\le c_2a_n}} \quad & \Big|\mathrm{E}_n[\rho_h^*(\varepsilon_{i0}+T_i-\frac{1}{\sqrt{n}}Z_i^{*T}\boldsymbol{\theta}_1-\mathbf{D}_i^T\mathbf{d}(\mathbf{b},\boldsymbol{\theta})-\mathbf{U}_\mathbf{i}^\mathbf{T}\mathbf{b},\mathbf{b}^\mathbf{T}\boldsymbol{\Sigma}_\mathbf{u}\mathbf{b})] \\
- \quad & \mathrm{E}_n[\rho_h^*(\varepsilon_{i0}+T_i-\mathbf{D}_i^T\mathbf{d}(\mathbf{b},\boldsymbol{\theta})-\mathbf{U}_\mathbf{i}^\mathbf{T}\mathbf{b},\mathbf{b}^\mathbf{T}\boldsymbol{\Sigma}_\mathbf{u}\mathbf{b})] \\
+ \quad & \frac{1}{\sqrt{n}}\boldsymbol{\theta}_1^T\mathrm{E}_n[\tilde{\mathbf{Z}}_i\frac{\partial\rho_h^*}{\partial\varepsilon}(\varepsilon_{i0}-\mathbf{U}_i^T\mathbf{b}_0,\mathbf{b}_0^T\Sigma_u\mathbf{b}_0)] \\
- \quad & \frac{1}{2n}\boldsymbol{\theta}_1^T\{\mathrm{E}_n[\mathbf{Z}_i^*\mathbf{Z}_i^{*T}\frac{\partial^2\rho_h^*}{\partial\varepsilon^2}(\varepsilon_{i0}-\mathbf{U}_i^T\mathbf{b}_0,\mathbf{b}_0^T\Sigma_u\mathbf{b}_0)]\}\boldsymbol{\theta}_1\Big|
\end{aligned}$$

$$\begin{aligned}
=2\sup_{\substack{\|\boldsymbol{\theta}_1\|\le c\\ \|\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\|\le c_1a_n\\ \|\hat{\mathbf{b}}-\mathbf{b}_0\|\le c_2a_n}} \quad & \Big|\mathrm{E}_n[\rho_h(\varepsilon_{i0}+T_i-\frac{1}{\sqrt{n}}Z_i^{*T}\boldsymbol{\theta}_1-\mathbf{D}_i^T\mathbf{d}(\mathbf{b},\boldsymbol{\theta}))] \\
& -\mathrm{E}_n[\rho_h(\varepsilon_{i0}+T_i-\mathbf{D}_i^T\mathbf{d}(\mathbf{b},\boldsymbol{\theta}))]+\frac{\mathbf{1}}{\sqrt{\mathbf{n}}}\boldsymbol{\theta}_\mathbf{1}^\mathbf{T}\mathrm{E}_\mathbf{n}[\mathbf{Z}_\mathbf{i}^*\frac{\partial\rho_\mathbf{h}}{\partial\varepsilon}(\varepsilon_{\mathbf{i0}})] \\
& -\frac{1}{2n}\boldsymbol{\theta}_1^T\{\mathrm{E}_n[\mathbf{Z}_i^*\mathbf{Z}_i^{*T}\frac{\partial^2\rho_h}{\partial\varepsilon^2}(\varepsilon_{i0})]\}\boldsymbol{\theta}_1\Big|\,[1+o_p(1)].
\end{aligned} \tag{S1.18}$$

The equation (S1.18) holds by (A10), $\mathrm{E}(\tilde{\mathbf{Z}}_i - \mathbf{Z}_i^*) = \mathbf{0}$ and the fact that ρ_h is the conditional expectation of ρ_h^* at any points with taking expectation of measurement errors, $\mathbf{U}_i^T\mathbf{b}$. Furthermore, applying Taylor expansion to $\rho_h(\cdot)$ in formula (S1.18) at ε_{i0} and ignoring terms higher than 2-order, we can get the dominant term of (S1.18) as

$$
\begin{aligned}
& \sup_{\substack{\|\boldsymbol{\theta}_1\| \le c \\ \|\boldsymbol{\theta}-\boldsymbol{\theta}_0\| \le c_1 a_n \\ \|\mathbf{b}-\mathbf{b}_0\| \le c_2 a_n}} \frac{2}{\sqrt{n}} \left| \boldsymbol{\theta}_1^T \mathrm{E}_n\{\mathbf{Z}_i^*[T_i - \mathbf{D}_i^T \mathbf{d}(\mathbf{b}, \boldsymbol{\theta})] \frac{\partial^{\mathbf{2}} \rho_{\mathbf{h}}}{\partial \varepsilon^{\mathbf{2}}}(\varepsilon_{\mathbf{i0}})\} \right| \\
= & \sup_{\substack{\|\boldsymbol{\theta}_1\| \le c \\ \|\boldsymbol{\theta}-\boldsymbol{\theta}_0\| \le c_1 a_n \\ \|\mathbf{b}-\mathbf{b}_0\| \le c_2 a_n}} \frac{2}{\sqrt{n}} \left| \boldsymbol{\theta}_1^T \mathrm{E}_n[\mathbf{Z}_i^* T_i \frac{\partial^2 \rho_h}{\partial \varepsilon^2}(\varepsilon_{i0})] + \boldsymbol{\theta}_1^T \mathrm{E}_n[\mathbf{Z}_i^* \mathbf{X}_i^T \frac{\partial^2 \rho_h}{\partial \varepsilon^2}(\varepsilon_{i0})](\mathbf{b} - \mathbf{b}_0) \right. \\
& + \left. \boldsymbol{\theta}_1^T \mathrm{E}_n[\mathbf{Z}_i^* (\mathbf{Z}_i - \mathbf{Z}_i^*)^T \frac{\partial^2 \rho_h}{\partial \varepsilon^2}(\varepsilon_{i0})](\boldsymbol{\theta} - \boldsymbol{\theta}_0) \right| \\
& \le \frac{c}{\sqrt{n}} \|\mathrm{E}_n[\mathbf{Z}_i^* T_i \frac{\partial^2 \rho_h}{\partial \varepsilon^2}(\varepsilon_{i0})]\| + a_n \frac{c}{\sqrt{n}} \|\mathrm{E}_n[\mathbf{Z}_i^* \mathbf{X}_i^T \frac{\partial^2 \rho_h}{\partial \varepsilon^2}(\varepsilon_{i0})]\| \\
& + a_n \frac{c}{\sqrt{n}} \|\mathrm{E}_n[\mathbf{Z}_i^* (\mathbf{Z}_i - \mathbf{Z}_i^*)^T \frac{\partial^2 \rho_h}{\partial \varepsilon^2}(\varepsilon_{i0})]\| \qquad \text{(S1.19)}
\end{aligned}
$$

Notice that $\mathrm{E}\|\mathbf{Z}_i^*\|^4$ is bounded and that each of the three summation terms in formula (S1.19) is zero mean by assumption (A9). The first term is of order $o_p(\frac{1}{n})$ plus assumptions (A2), (A5) and (A10). The second and third terms are also $o_p(\frac{1}{n})$ by $a_n\sqrt{m} = o(1)$. So far, we have completed the proof of $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = \hat{\boldsymbol{\theta}}_1^* + o_p(1)$ and the result follows straightly.

Table 1: Average number of basis selected by BIC criterion for data generated in Case 1. The values in the parentheses are Monte Carlo standard deviations.

τ	α	n	Oracle	Naive
0.5	1.1	200	1.730(0.468)	1.500(0.522)
		400	1.980(0.141)	1.830(0.403)
		600	2.030(0.171)	2.000(0.201)
0.5	2	200	1.530(0.521)	1.280(0.494)
		400	1.800(0.402)	1.520(0.502)
		600	1.920(0.273)	1.680(0.469)
0.75	1.1	200	1.810(0.506)	1.500(0.560)
		400	2.000(0.318)	1.880(0.433)
		600	2.070(0.256)	2.000(0.284)
0.75	2	200	1.520(0.611)	1.200(0.402)
		400	1.740(0.485)	1.520(0.522)
		600	1.920(0.273)	1.800(0.426)

Table 2: The value of selected $\hat{h}$ based on SIMEX for data generated in Case 1.

α	n	$\tau = 0.5$	$\tau = 0.75$
1.1	200	2.147	1.854
	400	1.714	1.526
	600	1.480	1.372
2	200	2.288	1.705
	400	1.808	1.569
	600	1.552	1.405

Table 3: Bias, Variance and MSE of parametric coefficient estimators in three different data generating models.

τ	Method	Bias($\hat{\theta_0}$)	Bias($\hat{\theta_1}$)	Bias($\hat{\theta_2}$)	Var($\hat{\theta_0}$)	Var($\hat{\theta_1}$)	Var($\hat{\theta_2}$)	MSE($\hat{\theta_0}$)	MSE($\hat{\theta_1}$)	MSE($\hat{\theta_2}$)
					Case 1					
0.5	CL	-0.0269	-0.0064	0.0361	0.0535	0.0088	0.0389	0.0536	0.0088	0.0398
	NAIVE	-0.1975	0.0013	0.2005	0.0209	0.0130	0.0100	0.0597	0.0129	0.0501
	ORACLE	-0.0230	-0.0048	0.0275	0.0272	0.0105	0.0177	0.0275	0.0105	0.0183
0.75	CL	-0.0050	0.0080	0.0287	0.0389	0.0125	0.0323	0.0385	0.0124	0.0328
	NAIVE	-0.1625	0.0016	0.1738	0.0245	0.0186	0.0086	0.0507	0.0184	0.0387
	ORACLE	-0.0300	0.0049	0.0281	0.0316	0.0182	0.0203	0.0322	0.0180	0.0209
					Case 2					
0.5	CL	-0.0283	0.0191	0.0226	0.0345	0.0105	0.0371	0.0350	0.0108	0.0372
	NAIVE	-0.2253	0.0184	0.2018	0.0114	0.0139	0.0080	0.0620	0.0141	0.0486
	ORACLE	-0.0238	0.0115	0.0162	0.0175	0.0120	0.0105	0.0179	0.0120	0.0107
0.75	CL	0.0334	0.0106	0.0242	0.0465	0.0126	0.0388	0.0472	0.0126	0.0390
	NAIVE	-0.2922	0.0014	0.2894	0.0172	0.0177	0.0065	0.1024	0.0176	0.0901
	ORACLE	-0.0284	0.0010	0.0224	0.0248	0.0138	0.0158	0.0254	0.0136	0.0162
					Case 3					

Table 3(continued table)

τ	Method	Bias($\hat{\theta_0}$)	Bias($\hat{\theta_1}$)	Bias($\hat{\theta_2}$)	Var($\hat{\theta_0}$)	Var($\hat{\theta_1}$)	Var($\hat{\theta_2}$)	MSE($\hat{\theta_0}$)	MSE($\hat{\theta_1}$)	MSE($\hat{\theta_2}$)
0.5	CL	-0.0064	-0.0185	0.0132	0.0337	0.0106	0.0242	0.0334	0.0109	0.0241
	NAIVE	-0.1761	-0.0177	0.1787	0.0148	0.0148	0.0050	0.0457	0.0150	0.0369
	ORACLE	-0.0211	-0.0116	0.0175	0.0218	0.0140	0.0111	0.0220	0.0140	0.0113
0.75	CL	-0.0234	-0.0077	0.0338	0.0317	0.0096	0.0290	0.0319	0.0096	0.0298
	NAIVE	-0.1682	-0.0065	0.1869	0.0173	0.0154	0.0074	0.0455	0.0153	0.0422
	ORACLE	0.0045	0.0012	-0.0050	0.0247	0.0137	0.0165	0.0244	0.0136	0.0164

Table 4: Bias, Variance and IMSE of functional coefficient estimators in three different data generating models.

	CL			NAIVE			ORACLE		
τ	$Bias^2$	Var	MSE	$Bias^2$	Var	MSE	$Bias^2$	Var	MSE
Case 1									
0.5	0.0141	0.0960	0.1101	0.0907	0.0295	0.1202	0.0134	0.0323	0.0457
0.75	0.0124	0.0812	0.0936	0.0862	0.0280	0.1142	0.0106	0.0486	0.0592
Case 2									
0.5	0.0123	0.0828	0.0950	0.0939	0.0221	0.1161	0.0093	0.0291	0.0384
0.75	0.0537	0.0942	0.1479	0.1873	0.0252	0.2125	0.0105	0.0349	0.0454
Case 3									
0.5	0.0111	0.0608	0.0718	0.0761	0.0176	0.0937	0.0084	0.0273	0.0357
0.75	0.0154	0.0831	0.0984	0.0885	0.0300	0.1185	0.0080	0.0344	0.0424

Bibliography

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

Kato, K. (2012) Estimation in functional linear quantile regression. *The Annals of Statistics*, **40**, 3108–3136.

Ledoux, M. and Talagrand, M. (1991) Probability in banach spaces. ergebnisse der mathematik, série 3, vol. 23.

Wang, H. J., Stefanski, L. A. and Zhu, Z. (2012) Corrected-loss estimation for quantile regression with covariate measurement errors. *Biometrika*, **99**, 405–421.