FUNCTIONAL LINEAR OPERATOR QUANTILE REGRESSION FOR SPARSE LONGITUDINAL DATA

Xingcai Zhou[1], Tingyu Lai[2] and Linglong Kong[3]

[1] *School of Statistics and Data Science, Nanjing Audit University, Nanjing 211815, China*

[2] *School of Mathematics and Statistics, Guangxi Normal University, Guilin 541006, China*

[3] *Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta T6G 2G1, Canada*

Supplementary Material

The supplementary material is organized as follows. Some estimation procedures are provided in Section S1. The choice of hyper-parameters is presented in Section S2. SLSE algorithm for functional linear operator QR model (3.14) is given in Section S3. Section S4 provides the proof of Theorem 2.1, while Some technical conditions and proofs of the main results for specific functional QR are presented in Section S5. Some additional experimental results are listed in Sections S6 and S7.

S1 Estimation procedure

All estimation for mean and covariance functions we used in the paper have been studied well by Yao (2007); Yao et al. (2005a,b); Şentürk and Müller (2010); Li and Hsing (2010), via the standard local linear smoothing procedure. For the readers to read smoothly, we summarize some estimations as follows. See Yao (2007); Yao et al. (2005a,b); Şentürk and Müller (2010); Li and Hsing (2010) for further details.

S1.1 For functional varying coefficient QR model

In the functional varying coefficient QR model for sparse longitudinal data, we aggregate data $\{(U_{ij}, V_{ij}, T_{ij}), i = 1, \cdots, n, j = 1, \cdots, N_i\}$.

Step 1: The mean estimates $\hat{\mu}_X$ and $\hat{\mu}_Y$ are obtained by smoothing the aggregated data (T_{ij}, U_{ij}) and (T_{ij}, V_{ij}), $i = 1, \cdots, n$, $j = 1, \cdots, N_i$. Define the local linear scatterplot smoothers for μ_X, based on the data (U_{ij}, T_{ij}), through minimizing

$$\sum_{i=1}^{n} \sum_{j=1}^{N_i} K_1 \left(\frac{T_{ij} - t}{b_X} \right) [U_{ij} - a_0 - a_1(t - T_{ij})]^2, \tag{S1.1}$$

with respect to a_0 and a_1, leading to $\hat{\mu}_X(t) = \hat{a}_0(t)$. Analogously for the mean function μ_Y of Y, based on the data (V_{ij}, T_{ij}).

Step 2: Compute the "raw" covariances of X and Y, based on

$$R_{X,i}(T_{ij}, T_{ik}) = (U_{ij} - \hat{\mu}_X(T_{ij}))\,(U_{ik} - \hat{\mu}_X(T_{ik}))$$

and $R_{Y,i}(T_{ij}, T_{ik}) = (V_{ij} - \hat{\mu}_Y(T_{ij}))\,(V_{ik} - \hat{\mu}_Y(T_{ik}))$, $i = 1, \cdots, n$, $j, k = 1, \cdots, N_i$, respectively. The smooth estimate $\hat{r}_{XX}$ (resp. $\hat{r}_{XY}$) of the covariance function r_{XX} (resp. r_{XY}) are got by scatterplot smoothing. Then, a nonparametric FPCA step yields the eigenfunction estimates $\hat{\phi}_k$ and $\hat{\psi}_k$, and the corresponding eigenvalues $\hat{\rho}_k$ and $\hat{\lambda}_k$ for the predictor and response trajectories.

Compute the "raw" covariances and the "raw" cross-covariances between X and Y, based on all observations from the same subject by

$$R_{X,i}(T_{ij},T_{ik}) = (U_{ij} - \hat{\mu}_X(T_{ij}))\,(U_{ik} - \hat{\mu}_X(T_{ik}))$$

and

$$R_{XY,i}(T_{ij},T_{ik}) = (U_{ij} - \hat{\mu}_X(T_{ij}))\,(V_{ik} - \hat{\mu}_Y(T_{ik}))\,,$$

$i = 1,\cdots,n,\ j,k = 1,\cdots,N_i$, by the local linear surface smoothers for r_{XX} and r_{XY}, respectively through minimizing

$$\sum_{i=1}^{n}\sum_{1\le j\ne k\le N_i} K_2\left(\frac{T_{ij}-s}{h_X},\frac{T_{ik}-t}{h_X}\right)[R_{X,i}(T_{ij},T_{ik}) - b_0 - b_{11}(s-T_{ij}) - b_{12}(t-T_{ik})]^2, \quad \text{(S1.2)}$$

with respect to b_0, b_{11} and b_{12}, and setting $\hat{r}_{XX}(s,t) = \hat{b}_0(s,t)$; and

$$\sum_{i=1}^{n}\sum_{j=1}^{N_i}\sum_{k=1}^{N_i} K_2\left(\frac{T_{ij}-s}{h_1},\frac{T_{ik}-t}{h_2}\right)[R_{XY,i}(T_{ij},T_{ik}) - b_0 - b_{11}(s-T_{ij}) - b_{12}(t-T_{ik})]^2, \quad \text{(S1.3)}$$

with respect to b_0, b_{11} and b_{12}, and setting $\hat{r}_{XY}(s,t) = \hat{b}_0(s,t)$.

Step 3: Give the initial estimators

$$\hat{\beta}_\tau^{(0)}(t) = \hat{\Gamma}_{XX}^{-1}\hat{\mathrm{E}}[\mathcal{L}_X^* Y] = \frac{\hat{r}_{XY}(t,t)}{\hat{r}_{XX}(t,t)}, \qquad \hat{\alpha}_\tau^{(0)}(t) = \hat{\mu}_Y(t) - \hat{\beta}_\tau^{(0)}(t)\hat{\mu}_X(t)$$

by least squares representation (He et al., 2000).

Step 4: Thus, we estimate

$$\widetilde{V}_{ij} = \hat{\alpha}_\tau^{(0)}(T_{ij}) + \hat{\beta}_\tau^{(0)}(T_{ij})U_{ij} - \left(\int_{\mathcal{T}} \hat{f}_t(0)dt\right)^{-1}\int_{\mathcal{T}} n^{-1}\sum_{i=1}^{n}\hat{\omega}_i(t)dt,$$

where $\hat{f}_t(0) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{N_i}\sum_{j=1}^{N_i} K_h\left(V_{ij} - \hat{\alpha}_\tau^{(0)}(t) - \hat{\beta}_\tau^{(0)}(t)U_{ij}\right)$,
$\hat{\omega}_i(t) = \mathbb{I}\left[\hat{V}_i^K(t) - \hat{\alpha}_\tau^{(0)}(t) - \hat{\beta}_\tau^{(0)}(t)\hat{U}_i^M(t) \le 0\right] - \tau$, and $\hat{U}_i^M(t) = \hat{\mu}_X(t) + \sum_{m=1}^{M}\hat{\zeta}_{im}\hat{\phi}_m(t)$,
$\hat{V}_i^K(t) = \hat{\mu}_Y(t) + \sum_{k=1}^{K}\hat{\xi}_{ik}\hat{\psi}_k(t)$.

For estimates $\hat{\zeta}_{im}$ and $\hat{\xi}_{ik}$, we only consider $\hat{\zeta}_{im}$, analogously for $\hat{\xi}_{ik}$. By functional principle analysis and numerical integration, the functional principle scores $\zeta_{im} = \int_{\mathcal{T}}(X_i(t) - \mu_X(t))\phi_m(t)dt$ can be estimated as

$$\hat{\zeta}_{im}^D = \sum_{l=1}^{N_i}(U_{il} - \hat{\mu}_X(T_{il}))\hat{\phi}_m(T_{il})(T_{il} - T_{i,l-1}),$$

which will works well when the grid of measurements is dense. However, for sparse functional data, $\hat{\zeta}_{im}^D$ will not provide reasonable approximations to ζ_{im}. Using the procedure of PACE in Yao et al. (2005a), we get

$$\hat{\zeta}_{im} = \widehat{E}[\zeta_{im}|\tilde{U}_i] = \hat{\rho}_m \hat{\boldsymbol{\phi}}_{im}^T \hat{\boldsymbol{\Sigma}}_{\boldsymbol{U}_i}^{-1}(\tilde{\boldsymbol{U}}_i - \hat{\boldsymbol{\mu}}_X), \tag{S1.4}$$

where $\hat{\boldsymbol{\phi}}_{im} = \left(\hat{\phi}_{im}(T_{i1}), \cdots, \hat{\phi}_{im}(T_{iN_i})\right)^T$, $\tilde{\boldsymbol{U}}_i = (U_{i1}, \cdots, U_{iN_i})^T$, $\hat{\boldsymbol{\mu}}_X = (\mu_X(T_{i1}), \cdots, \mu_X(T_{iN_i}))^T$, and the (j,l)th component of $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{U}_i}$ is

$$(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{U}_i})_{j,l} = \hat{r}_{XX}(T_{ij}, T_{il}) + \hat{\sigma}_X^2 \delta_{jl}$$

with $\delta_{jl} = 1$ if $j = l$ and 0 if $j \neq l$. For $\hat{\sigma}_X^2$, we can use the procedure of (S1.8).

Step 5: Calculate the mean estimate $\mu_{\widetilde{Y}}$ via local linear fitting, based on data $(T_{ij}, \widetilde{V}_{ij})$, $i = 1, \cdots, n$, $j = 1, \cdots, N_i$. Analogously, we can obtain the mean estimator $\hat{\mu}_{\tilde{Y}}$ by fitting data $\{(T_{ij}, \tilde{V}_{ij}), i = 1, \cdots, n_i, j = 1, \cdots, N_i\}$ to (S1.1).

Step 6: Compute the "raw" cross-covariances between X and $\widetilde{Y}$ based on

$$R_{i,X\widetilde{Y}} = (U_{ij} - \mu_X(T_{ij}))\left(\widetilde{V}_{ik} - \mu_{\widetilde{Y}}(T_{ik})\right),$$

$i = 1, \cdots, n$, $j, k = 1, \cdots, N_i$, which serve as input for the two-dimensional smoothing step to obtain $\hat{r}_{\widetilde{Y}\widetilde{Y}}$ and $\hat{r}_{X\widetilde{Y}}$, respectively.

Analogously, we can obtain the cross-covariances estimator $\hat{r}_{X\widetilde{Y}}$ between X and $\widetilde{Y}$ by fitting $R_{i,X\widetilde{Y}} = (U_{ij} - \hat{\mu}_X(T_{ij}))\left(\widetilde{V}_{ik} - \hat{\mu}_{\widetilde{Y}}(T_{ik})\right)$, $i = 1, \cdots, n$, $j, k = 1, \cdots, N_i$, to (S1.3).

Step 7: From (3.7), we have the first step estimate of the iterative algorithm for the functional varying coefficient QR model (3.3) as

$$\hat{\beta}_\tau^{(1)}(t) = \frac{\hat{r}_{X\widetilde{Y}}(t,t)}{\hat{r}_{XX}(t,t)}, \qquad \hat{\alpha}_\tau^{(1)}(t) = \hat{\mu}_{\widetilde{Y}}(t) - \hat{\beta}_\tau^{(1)}(t)\hat{\mu}_X(t). \tag{S1.5}$$

We sketch the functional estimation approach for functional varying coefficient QR model (3.3) in Algorithm 1 by combining the above steps. Thus, the final estimators $\hat{\alpha}_\tau^{(\mathrm{K})}$ and $\hat{\beta}_\tau^{(\mathrm{K})}$ are obtained via Algorithm 1.

Algorithm 1: SLSE algorithm for functional linear operator QR model (3.3).

Input: Kernel function $K(\cdot)$, bandwidth h, quantile level τ and the number of iterations K.

Calculate mean function $\hat{\mu}_X$ and $\hat{\mu}_Y$, covariance surface $\hat{r}_{XX}$, cross-covariance surface $\hat{r}_{XY}$, eigenfunctions $\hat{\phi}_k$ and$\hat{\psi}_k$, and eigenvalues $\hat{\rho}_k$ and $\hat{\lambda}_k$ by Steps 1-2.

Initialize estimators $\hat{\alpha}_\tau^{(0)}$ and $\hat{\beta}_\tau^{(0)}$ by Step 3.

for $k = 1, 2, \cdots, \mathrm{K}$ **do**

Estimate $\hat{f}_t^{(k)}(0)$, $\hat{\omega}_i$, $\hat{U}_i$ and $\hat{V}_i$ for obtaining $\widetilde{V}_{ij}$ by Step 4.

Compute mean function $\hat{\mu}_{\widetilde{Y}}$ via Step 5, cross-covariance surface $\hat{r}_{X\widetilde{Y}}$ by Step 6.

Obtain $\hat{\alpha}_\tau^{(k)}$ and $\hat{\beta}_\tau^{(k)}$ by Step 7.

end

Output: The final estimators $\hat{\alpha}_\tau^{(\mathrm{K})}$ and $\hat{\beta}_\tau^{(\mathrm{K})}$.

For $\hat{\sigma}_X^2$ in (3.12): First, we estimate $V_X(t) := C_X(t) + \hat{\sigma}_X^2$, where $C_X(s,t) = E(X(s)E(t))$

and $C_X(t) := C_X(t,t)$, by the local linear smooth through minimizing

$$\sum_{i=1}^{n}\sum_{j=1}^{N_i} K_1\left(\frac{T_{ij}-t}{h_X}\right)[U_{ij}^2 - a_0 - a_1(t-T_{ij})]^2 \tag{S1.6}$$

with respective to a_0 and a_1, and setting $\hat{V}_X(t) = \hat{a}_0$; Second, we estimate $\hat{C}_X(t)$ by the local linear smooth via minimizing

$$\sum_{i=1}^{n}\sum_{1\le j\ne k\le N_i} K_2\left(\frac{T_{ij}-t}{h_X}, \frac{T_{ik}-t}{h_X}\right)[U_{ij}U_{ik} - b_0 - b_1(t-T_{ij})]^2 \tag{S1.7}$$

with respective to b_0 and b_1, and setting $\hat{C}_X(t) = \hat{b}_0$. Finally, the estimator $\hat{\sigma}_X^2$ is

$$\hat{\sigma}_X^2 = \frac{1}{|\mathcal{T}|}\int_{\mathcal{T}}\{\hat{V}_X(t) - \hat{C}_X(t)\}dt. \tag{S1.8}$$

S1.2 For functional linear QR model

S1.2.1 Functional Approach

From the estimation strategy in Section 2.2, the functional linear QR model (3.14) can be translated into the following functional linear operator regression model

$$\mathrm{E}[\widetilde{Y}(t)|X(t)] = \alpha_\tau(t) + (\mathcal{L}_X\beta)(t), \tag{S1.9}$$

where $(\mathcal{L}_X\beta)(t) = \int_{\mathcal{S}}\beta_\tau(s,t)X(s)ds$

$$\widetilde{Y}(t) = Q_{Y|X}^{(0)}(t;\tau) - \left(\int_{\mathcal{T}} f_t(0)dt\right)^{-1}\int_{\mathcal{T}}\{\mathbb{I}[Y(t) - Q_{Y|X}^{(0)}(t;\tau)\le 0] - \tau\}dt,$$

$$Q_{Y|X}^{(0)}(t;\tau) = \alpha_\tau^{(0)}(t) + \int_{\mathcal{S}}\beta_\tau^{(0)}(s,t)X(s)ds.$$

The model (S1.9) can be rewritten as

$$\mathrm{E}[\widetilde{Y}(t)|X(t)] = \mu_{\widetilde{Y}}(t) + \int_{\mathcal{S}}\beta_\tau(s,t)X^c(s)ds,$$

where $\mathrm{E}[\widetilde{Y}(t)] = \mu_{\widetilde{Y}}(t) = \alpha_\tau(t) + \int_{\mathcal{S}} \beta_\tau(s,t)\mu_X(s)ds$. By our result (2.12), one gets

$$\beta_\tau^{(1)}(s,t) = \Gamma_{X^cX^c}^{-1} E\left(\mathcal{L}_{X^c}^{*} \widetilde{Y}^c\right) = \sum_{m=1}^{\infty}\sum_{l=1}^{\infty} \frac{\mathrm{E}[\zeta_m \varsigma_l]}{\mathrm{E}[\zeta_m^2]}\phi_m(s)\varphi_l(t), \tag{S1.10}$$

where ς_l and φ_l are defined in (3.5), and ζ_m and ϕ_m are given in Section 2.1. $r_{\tilde{Y}\tilde{Y}}$ and $r_{X\tilde{Y}}$ are defined in (3.5) and (3.8), respectively. Next, we sketch estimation steps of $\beta_\tau^{(1)}(s,t)$ and $\alpha_\tau^{(1)}(s,t)$ as in Subsection 3.1, which is provided Appendix S1.2. The algorithm is similar to Algorithm 1, which is present in Appendix S3. We obtain the final estimators $\hat{\alpha}_\tau^{(\mathrm{K})}$ and $\hat{\beta}_\tau^{(\mathrm{K})}$.

S1.2.2 Steps of Estimation

In the model, we have data $\{(S_{il,U_{il}}), (T_{ij}, V_{ij}), i = 1, \cdots, n, l = 1, \cdots, L_i, j = 1, \cdots, N_i\}$. We give some necessary estimates in our algorithm.

Step 1: Smooth the aggregated data (T_{il}, U_{il}) and (S_{ij}, V_{ij}), $i = 1, \cdots, n$, $l = 1, \cdots, L_i$, $j = 1, \cdots, N_i$ to obtain the estimated mean functions μ_X and μ_Y, respectively.

Step 2: Compute the raw covariances based on all observations

$$R_{X,i}(S_{il_1}, S_{il_2}) = (U_{il_1} - \hat{\mu}_X(S_{il_1}))\,(U_{il_2} - \hat{\mu}_X(S_{il_2}))$$

and $R_{Y,i}(T_{ij_1}, T_{ij_2}) = (V_{ij_1} - \hat{\mu}_Y(T_{ij_1}))\,(V_{ij_2} - \hat{\mu}_Y(T_{ij_2}))$, $i = 1, \cdots, n$, $l_1, l_2 = 1, \cdots, L_i$, $j_1, j_2 = 1, \cdots, N_i$, smooth them to get the estimated covariance functions $\hat{r}_{XX}$ and $\hat{r}_{XY}$, then yield estimates $(\hat{\rho}_m, \hat{\phi}_m)$ for predictor and $(\hat{\lambda}_k, \hat{\psi}_k)$ for response, respectively.

Similar to (S1.2) for $\hat{r}_{XX}$, we obtain $\hat{r}_{XX} = \hat{b}_0$ by the local linear surface smoother via minimizing

$$\sum_{i=1}^{n}\sum_{1\le l_1\ne l_2\le L_i} K_2\left(\frac{T_{il_1}-s}{h_X}, \frac{T_{il_2}-t}{h_X}\right)\left[R_{X,i}(T_{il_1}, T_{il_2}) - b_0 - b_{11}(s - T_{il_1}) - b_{12}(t - T_{il_2})\right]^2,$$

where

$$R_{X,i}(T_{il_1}, T_{il_2}) = (U_{il_1} - \hat{\mu}_X(T_{il_1}))\,(U_{il_2} - \hat{\mu}_X(T_{il_2}))\,;$$

Analogously for the covariance function r_{XY} based on $\{R_{Y,i}(S_{ij_1}, S_{ij_2}), i = 1, \cdots, n, j_1, j_2 = 1, \cdots, N_i\}$, where

$$R_{Y,i}(S_{ij_1}, S_{ij_2}) = (V_{ij_1} - \hat{\mu}_Y(T_{ij_1}))\,(V_{ij_2} - \hat{\mu}_Y(T_{ij_2}))\,.$$

The estimates $(\hat{\rho}_m, \hat{\phi}_m)$ for predictor and $(\hat{\lambda}_k, \hat{\psi}_k)$ for response are the solutions of the following eigenequations

$$\int_{\mathcal{S}} \hat{r}_{XX}(s_1, s_2)\hat{\phi}_m(s_1)ds_1 = \hat{\rho}_m\hat{\phi}(s_2) \tag{S1.11}$$

and

$$\int_{\mathcal{T}} \hat{r}_{YY}(t_1, t_2)\hat{\psi}_m(t_1)ds_1 = \hat{\lambda}_m\hat{\psi}(t_2) \tag{S1.12}$$

with orthonormal constraints on $\{\hat{\phi}_m\}_{m\geq 1}$ and $\{\hat{\psi}_m\}_{m\geq 1}$, respectively. From (S1.11) and (S1.12), we know their eigenfunctions and eigenvalues are calculated as solutions of the above eigenequations. In practice, they are numerically obtained by discretization.

Step 3: Give the estimators of the mean regression model $\mathrm{E}[Y|X] = \alpha(t)+\int_{\mathcal{S}}\beta(s,t)X(s)ds$ as initial estimators of α_τ and β_τ,

$$\hat{\beta}_\tau^{(0)}(s,t) = \hat{\Gamma}_{XX}^{-1}\hat{\mathrm{E}}[\mathcal{L}_X^* Y] = \sum_{m=1}^{M}\sum_{k=1}^{K}\frac{\hat{\sigma}_{mk}}{\hat{\rho}_m}\hat{\phi}_m(s)\hat{\psi}_k(t),$$

$$\hat{\alpha}_\tau^{(0)}(t) = \hat{\mu}_Y(t) - \int_{\mathcal{S}}\hat{\beta}_\tau^{(0)}(s,t)\hat{\mu}_X(s)ds.$$

We obtain estimates for $\sigma_{mk} = \mathrm{E}[\zeta_m\xi_k]$ by

$$\hat{\sigma}_{mk} = \int_{\mathcal{T}}\int_{\mathcal{S}}\hat{\phi}_m(s)\hat{r}_{XY}(s,t)\psi_k(t)dsdt, m = 1, \cdots, M, k = 1, \cdots, K.$$

where $\hat{r}_{XY}$ is an estimate of the cross-covariance surface r_{XY}, which can be obtained by smoothing the raw cross-covariances $R_{XY,i}(S_{il}, T_{ik}) = (U_{il} - \hat{\mu}_X(S_{il}))(V_{ik} - \hat{\mu}_Y(T_{ik}))$, $i = 1, \cdots, n$, $l = 1, \cdots, L_i$, $j = 1, \cdots, N_i$.

We adopt two-dimensional scatterplot smoothing to estimate r_{XY} via minimizing

$$\sum_{i=1}^{n}\sum_{j=1}^{L_i}\sum_{k=1}^{N_i} K_2\left(\frac{T_{ij}-s}{h_1}, \frac{T_{ik}-t}{h_2}\right)[R_{XY,i}(T_{ij}, T_{ik}) - b_0 - b_{11}(s - T_{ij}) - b_{12}(t - T_{ik})]^2,$$

with respect to b_0, b_{11} and b_{12}, and setting $\hat{r}_{XY}(s,t) = \hat{b}_0(s,t)$, where

$$R_{XY,i}(T_{il}, S_{ik}) = (U_{il} - \hat{\mu}_X(T_{il}))(V_{ik} - \hat{\mu}_Y(S_{ik})),$$

for $i = 1, \cdots, n, l = 1, \cdots, L_i, j = 1, \cdots, N_i$.

Step 4: Estimate the surrogate response trajectory $\widetilde{Y}$ in (S1.9),

$$\widetilde{V}_{ij} = \hat{Q}^{(0)}_{Y|X}(T_{ij};\tau) - \left(\int_{\mathcal{T}} \hat{f}_t(0)dt\right)^{-1} \int_{\mathcal{T}} \frac{1}{n}\sum_{i=1}^{n} \hat{\omega}_i(t)dt,$$

where $\hat{Q}^{(0)}_{Y|X}(t;\tau) = \hat{\alpha}^{(0)}_\tau(t) + \int_{\mathcal{S}} \hat{\beta}^{(0)}_\tau(s,t)\hat{U}^M_i(s)ds$,

$$\hat{f}_t(0) = \frac{1}{n}\sum_{i=1}^{n} K_h\left(\hat{V}^K_i(t) - \hat{Q}^{(0)}_{Y|X}(t;\tau)\right),$$

$$\hat{\omega}_i(t) = \mathbb{I}\left[\hat{V}^K_i(t) - \hat{Q}^{(0)}_{Y|X}(t;\tau) \le 0\right] - \tau,$$

$$\hat{U}^M_i(t) = \hat{\mu}_X(t) + \sum_{m=1}^{M} \hat{\zeta}_{im}\hat{\phi}_m(t), \quad \hat{V}^K_i(t) = \hat{\mu}_Y(t) + \sum_{k=1}^{K} \hat{\xi}_{ik}\hat{\psi}_k(t).$$

The numbers M and K can be chosen by one-curve-leave-out cross-validation or by an AIC criterion (see Appendix S2). Also, see Yao et al. (2005b).

For estimates $\hat{\zeta}_{im}$ and $\hat{\xi}_{ik}$, we only consider $\hat{\zeta}_{im}$, analogously for $\hat{\xi}_{ik}$. By functional principle analysis and numerical integration, the functional principle scores $\zeta_{im} = \int_{\mathcal{T}}(X_i(t) -$

$\mu_X(t))\phi_m(t)dt$ can be estimated as

$$\hat{\zeta}_{im}^D = \sum_{l=1}^{L_i}(U_{il} - \hat{\mu}_X(T_{il}))\hat{\phi}_m(T_{il})(T_{il} - T_{i,l-1}),$$

which will works well when the grid of measurements is dense. However, for sparse functional data, $\hat{\zeta}_{im}^D$ will not provide reasonable approximations to ζ_{im}. Using the procedure of PACE in Yao et al. (2005a), we get

$$\hat{\zeta}_{im} = \widehat{E}[\zeta_{im}|\tilde{U}_i] = \hat{\rho}_m\hat{\boldsymbol{\phi}}_{im}^T\hat{\boldsymbol{\Sigma}}_{\boldsymbol{U}_i}^{-1}(\tilde{\boldsymbol{U}}_i - \hat{\boldsymbol{\mu}}_X),$$

where $\hat{\boldsymbol{\phi}}_{im} = \left(\hat{\phi}_{im}(T_{i1}), \cdots, \hat{\phi}_{im}(T_{iL_i})\right)^T$, $\tilde{\boldsymbol{U}}_i = (U_{i1}, \cdots, U_{iL_i})^T$, $\hat{\boldsymbol{\mu}}_X = (\mu_X(T_{i1}), \cdots, \mu_X(T_{iL_i}))^T$, and the (j,l)th component of $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{U}_i}$ is

$$(\hat{\boldsymbol{\Sigma}}_{\boldsymbol{U}_i})_{j,l} = \hat{r}_{XX}(T_{ij}, T_{il}) + \hat{\sigma}_X^2\delta_{jl}$$

with $\delta_{jl} = 1$ if $j = l$ and 0 if $j \neq l$. For $\hat{\sigma}_X^2$, we can use the procedure of (S1.8).

Step 5: Calculate the mean function of $\widetilde{Y}$,

$$\hat{\mu}_{\widetilde{Y}}(t) = \hat{\alpha}_\tau^{(0)}(t) + \int_{\mathcal{S}}\hat{\beta}_\tau^{(0)}(s,t)\left[\frac{1}{n}\sum_{i=1}^n \hat{U}_i(s)\right]ds - \left(\int_{\mathcal{T}}\hat{f}_t(0)dt\right)^{-1}\int_{\mathcal{T}}\frac{1}{n}\sum_{i=1}^n \hat{\omega}_i(t)dt.$$

Step 6: Compute the "raw" covariance of $\widetilde{Y}$ and cross-covariances between X and $\widetilde{Y}$ based on $R_{i,\widetilde{Y}\widetilde{Y}}(T_{ij}, T_{ik}) = \left(\widetilde{V}_{ij} - \hat{\mu}_{\widetilde{Y}}(T_{ij})\right)\left(\widetilde{V}_{ik} - \hat{\mu}_{\widetilde{Y}}(T_{ik})\right)$ and

$$R_{i,X\widetilde{Y}}(S_{il}, T_{ik}) = (U_{il} - \hat{\mu}_X(S_{il}))\left(\widetilde{V}_{ik} - \hat{\mu}_{\widetilde{Y}}(T_{ik})\right),$$

$i = 1, \cdots, n$, $j, k = 1, \cdots, N_i$, and $l = 1, \cdots, L_i$., which serve as input for the two-dimensional smoothing step to obtain $\hat{r}_{\widetilde{Y}\widetilde{Y}}$ and $\hat{r}_{X\widetilde{Y}}$, respectively. Similar to Step 3, estimate $\tilde{\sigma}_{mk} = E[\zeta_m\varphi_k]$ by

$$\hat{\tilde{\sigma}}_{mk} = \int_{\mathcal{T}}\int_{\mathcal{S}}\hat{\phi}_m(s)\hat{r}_{X\widetilde{Y}}(s,t)\varphi_k(t)dsdt.$$

Similar to (S1.2) and (S1.3), we obtain $\hat{r}_{\tilde{Y}\tilde{Y}}$ and $\hat{r}_{X\tilde{Y}}$ based on

$$R_{i,\widetilde{Y}\widetilde{Y}} = \left(\widetilde{V}_{ij} - \hat{\mu}_{\widetilde{Y}}(S_{ij})\right)\left(\widetilde{V}_{ik} - \hat{\mu}_{\widetilde{Y}}(S_{ik})\right)$$

and $R_{i,X\widetilde{Y}} = (U_{il} - \hat{\mu}_X(T_{il}))\left(\widetilde{V}_{ik} - \hat{\mu}_{\widetilde{Y}}(S_{ik})\right)$, $i = 1, \cdots, n$, $j, k = 1, \cdots, N_i$, and $l = 1, \cdots, L_i$.

Step 7: From (S1.10), we have the first step estimate of the iterative algorithm for the functional linear QR model (3.14) as

$$\begin{aligned} \hat{\beta}_{\tau}^{(1)}(s,t) &= \sum_{m=1}^{\tilde{M}}\sum_{k=1}^{\tilde{K}} \frac{\hat{\tilde{\sigma}}_{mk}}{\hat{\rho}_m}\hat{\phi}_m(s)\hat{\varphi}_k(t), \\ \hat{\alpha}_{\tau}^{(1)}(t) &= \hat{\mu}_{\widetilde{Y}}(t) - \int_{\mathcal{S}} \hat{\beta}_{\tau}^{(1)}(s,t)\hat{\mu}_X(s)ds. \end{aligned} \tag{S1.13}$$

S1.2.3 The τth quantile target trajectory in Theorem 3.6

For sparse and irregular measurements of the new predictor trajectory X^*, the prediction of the τth quantile response trajectory would be gained via

$$Q_{Y^*|X^*}(t;\tau) = \alpha_{\tau}(t) + \sum_{m=1}^{\infty}\sum_{k=1}^{\infty}\frac{\tilde{\sigma}_{mk}}{\rho_k}\zeta_m^*\varphi_k(t), \tag{S1.14}$$

where $\zeta_m^* = \int_{\mathcal{S}}(X^*(s) - \mu_X(s))\phi_m(s)ds$ is the mth FPC score of the X^*, $(\tilde{\sigma}_{mk}, \varphi_k(t))$ is obtained based on the X and $\widetilde{Y}^*$ with

$$\widetilde{Y}^*(t) = Q_{Y|X}(t;\tau) - \left(\int_{\mathcal{T}} f_t(0)dt\right)^{-1}\int_{\mathcal{T}}\{\mathbb{I}[Y(t) - Q_{Y|X}^*(t;\tau) \le 0] - \tau\}dt,$$

$$Q_{Y|X}(t;\tau) = \alpha_{\tau}(t) + \int_{\mathcal{S}}\beta_{\tau}(s,t)X(s)ds.$$

Next, we give the best linear prediction for ζ_m^* available to the sparsity of the data.

Let $X^*(S_l^*)$ be the value of the predictor function X^* at location S_l^*, $U_l^* = X(S_l^*) + \epsilon_{X,l}^*$, $l = 1, \cdots, L_i$, with L^* a random number. Denote the observations $\boldsymbol{U}^* = (U_1^*, \cdots, U_{L^*}^*)^T$, the locations $\boldsymbol{S}^* = (S_1^*, \cdots, S_{L^*}^*)^T$, and let $\boldsymbol{X}^* = (X^*(S_1^*), \cdots, X^*(S_{L^*}^*))^T$, $\boldsymbol{\mu}_X^* = (\mu_X(S_1^*), \cdots, \mu_X(S_{L^*}^*))^T$ and $\boldsymbol{\phi}_m^* = (\phi_m(S_1^*), \cdots, \phi_m(S_{L^*}^*))^T$. Assume that the FPC scores ζ_m^* and the errors $\epsilon_{X,l}^*$ are jointly Gaussian. Following Yao et al. (2005b,a), the best linear prediction for ζ_m^* is

$$\tilde{\zeta}_m^* = \rho_m \boldsymbol{\phi}_m^{*T} \boldsymbol{\Sigma}_{U^*}^{-1} (\boldsymbol{U}^* - \boldsymbol{\mu}_X^*), \tag{S1.15}$$

where $\boldsymbol{\Sigma}_{U^*} = \mathrm{Cov}(\boldsymbol{U}^* | \boldsymbol{S}^*, L^*) = \mathrm{Cov}(\boldsymbol{X}^* | \boldsymbol{S}^*, L^*) + \sigma_X^2 I_{L^*}$, its (j,l)th entry $(\boldsymbol{\Sigma}_{U^*})_{j,l} = r_{XX}(S_j, S_l) + \sigma_X^2 \delta_{jl}$. The α_τ, ψ_k, σ_{mk} and ρ_k can be estimated from the data based on Algorithm S in Appendix S3. Together with (S1.15), we have the estimate

$$\hat{\zeta}_m^* = \hat{\rho}_m \hat{\boldsymbol{\phi}}_m^{*T} \hat{\boldsymbol{\Sigma}}_{U^*}^{-1} (\boldsymbol{U}^* - \hat{\boldsymbol{\mu}}_X^*), \tag{S1.16}$$

where $(\hat{\boldsymbol{\Sigma}}_{U^*})_{jl} = \hat{r}_{XX}(S_j, S_l) + \hat{\sigma}_X^2 \delta_{jl}$. Thus, the τth quantile predicted trajectory is obtained by

$$\widehat{Q}_{Y^*|X^*}^{\mathrm{K},M,K}(t;\tau) = \hat{\alpha}_\tau^{\mathrm{K}}(t) + \sum_{m=1}^{M} \sum_{k=1}^{K} \frac{\hat{\tilde{\sigma}}_{mk}^{\mathrm{K}}}{\hat{\rho}_k} \hat{\zeta}_m^* \hat{\varphi}_k^{\mathrm{K}}(t), \tag{S1.17}$$

where $(\hat{\tilde{\sigma}}_{mk}^{\mathrm{K}}, \hat{\varphi}_k^{\mathrm{K}}(t))$ is obtained based on X and $\widetilde{Y}^{\mathrm{K}}$ with

$$\widetilde{Y}^{\mathrm{K}}(t) = Q_{Y|X}^{\mathrm{K}}(t;\tau) - \left(\int_{\mathcal{T}} f_t(0) dt \right)^{-1} \int_{\mathcal{T}} \{ \mathbb{I}[Y(t) - Q_{Y|X}^{\mathrm{K}}(t;\tau) \leq 0] - \tau \} dt,$$

$$Q_{Y|X}^{\mathrm{K}}(t;\tau) = \hat{\alpha}_\tau^{\mathrm{K}}(t) + \int_{\mathcal{S}} \hat{\beta}_\tau^{\mathrm{K}}(s,t) X(s) ds.$$

In addition, for asymptotic pointwise confidence bands for quantile response trajectories, we have the following some narration.

Let $\boldsymbol{\zeta}_M^* = (\zeta_1^*, \cdots, \zeta_M^*)$, $\tilde{\boldsymbol{\zeta}}_M^* = \left(\tilde{\zeta}_1^*, \cdots, \tilde{\zeta}_M^*\right)$, where $\tilde{\zeta}_m^*$ is as in (S1.15). Define $\boldsymbol{H} = \text{Cov}(\boldsymbol{\zeta}_M^*, \boldsymbol{U}^* | L^*, \boldsymbol{S}^*) = (\rho_1 \phi_1^*, \cdots, \rho_M \phi_M^*)^T$, which a $M \times L^*$ matrix. Further, the covariance matrix of $\tilde{\boldsymbol{\zeta}}_M^*$ is $\text{Cov}(\tilde{\boldsymbol{\zeta}}_M^* | L^*, \boldsymbol{S}^*) = \boldsymbol{H}\boldsymbol{\Sigma}_{\boldsymbol{U}^*}^{-1}\boldsymbol{H}^T$. Since $\tilde{\boldsymbol{\zeta}}_m^* = \text{E}\left[\boldsymbol{\zeta}_M^* | \boldsymbol{U}^*, L^*, \boldsymbol{S}^*\right]$ is the project of $\boldsymbol{\zeta}_M^*$ on the space spanned by the linear functions of $\boldsymbol{U}^*$ given L^* and $\boldsymbol{S}^*$,

$$\text{Cov}\left(\tilde{\boldsymbol{\zeta}}_M^* - \boldsymbol{\zeta}_M^* | L^*, \boldsymbol{S}^*\right) = \boldsymbol{D} - \boldsymbol{H}\boldsymbol{\Sigma}_{\boldsymbol{U}^*}^{-1}\boldsymbol{H}^T \equiv \boldsymbol{\Omega}_M,$$

where $\boldsymbol{D} = \text{diag}(\rho_1, \cdots, \rho_M)$. Again, we have

$$\tilde{\boldsymbol{\zeta}}_M^* - \boldsymbol{\zeta}_M^* \sim N(0, \boldsymbol{\Omega}_M),$$

under Gaussian assumptions and conditional on L^* and $\boldsymbol{S}^*$.

First, we have estimates $\hat{\boldsymbol{\Omega}}_M = \hat{\boldsymbol{D}} - \hat{\boldsymbol{H}}\hat{\Sigma}_{\boldsymbol{U}^*}^{-1}\hat{\boldsymbol{H}}^T$, where

$$\hat{\boldsymbol{D}} = \text{diag}(\hat{\rho}_1, \cdots, \hat{\rho}_M)$$

and $\hat{\boldsymbol{H}} = (\hat{\rho}_1 \hat{\phi}_1^*, \cdots, \hat{\rho}_M \hat{\phi}_M^*)^T$. Then, obtain the estimates

$$\hat{\boldsymbol{\varphi}}_t^{\text{K},K} = (\varphi_1^{\text{K}}(t), \cdots, \varphi_K^{\text{K}}(t))^T$$

of $\boldsymbol{\varphi}_t^K = (\varphi_1(t), \cdots, \varphi_K(t))^T$, and $\hat{\boldsymbol{P}}_{M,K}^{\text{K}} = (\hat{\tilde{\sigma}}_{mk}^{\text{K}} / \hat{\rho}_m)_{1 \le m \le M, 1 \le k \le K}$ of the matrix $M \times K$ $\boldsymbol{P}_{M,K} = (\tilde{\sigma}_{mk} / \rho_m)_{1 \le m \le M, 1 \le k \le K}$ based on the data. Last, we write the prediction (S1.17) as $\widehat{Q}_{Y^*|X^*}^{\text{K},M,K}(t;\tau) = \hat{\alpha}_\tau^{\text{K}}(t) + \left(\hat{\boldsymbol{\zeta}}_M^*\right)^T \boldsymbol{P}_{M,K}^{\text{K}} \hat{\boldsymbol{\varphi}}_t^{\text{K},K}$.

S1.3 For functional varying coefficient QR model with history index

In the subsection, we consider the scenario $(\mathcal{L}_X\beta)(t) = \beta_\tau(t)\int_0^\Delta \gamma_\tau(s)X(t-s)ds$ for $t \in \mathcal{T} = [\Delta, T]$, where $\Delta > 0$ is the length of a sliding window and $T > \Delta$. Our functional varying coefficient QR model with history index is

$$Q_{Y|X}(t;\tau) = \alpha_\tau(t) + \beta_\tau(t)\int_0^\Delta \gamma_\tau(s)X(t-s)ds. \tag{S1.18}$$

The model provides a parsimonious and intuitive balance by introducing a history index function γ_τ, which serves to convey the effects of the recent past of the predictor on current response Şentürk and Müller (2010) at the τth quantile level. The varying coefficient function β_τ represents the magnitude of this influence as a function of time. Suppose that α_τ, β_τ and γ are smooth. For identifiability, we assume that γ_τ is normalized via $\int_0^\Delta \gamma_\tau^2(u)du = 1$ and $\gamma_\tau(0) > 0$ at each τth quantile level. In the model (S1.18), we assume that the history index function γ does not change over time. Thus, the time effects encoded in β_1 and history effects encoded in γ are separate, which are two easily interpretable one-dimensional component functions for the functional regression model. Once γ has been estimated, (S1.18) reduce to a functional varying coefficient QR model. Here, we also present some main results of the model. Its function approach is similar to subsection 3.1.1.

S1.3.1 Functional approach

Note that even if γ is known, the predictors of the reduced varying coefficient model, that is, $\int_0^\Delta \gamma_\tau(u)X(t-u)du$, may be infeasible by numerical integration, due to the sparsity of the observations for the predictor trajectory in the history window $[t-\Delta, t]$. We propose an estimation algorithm for the functional varying coefficient QR model to meet the challenges of data sparsity and non-smooth QR loss, which builds on our surrogate least square estimation for QR and functional principal component analysis.

By the estimation strategy in Section 2.2, given initial estimators $\alpha_\tau^{(0)}$, $\beta_\tau^{(0)}$ and $\gamma_\tau^{(0)}$, we can translate the model (S1.18) into the following functional linear operator regression

$$\mathrm{E}\left\{\widetilde{Y}(t)|X(s), s\in[t-\Delta, t]\right\} = \alpha_\tau(t) + (\mathcal{L}_X\beta)(t), \tag{S1.19}$$

where $(\mathcal{L}_X\beta)(t) = \beta_\tau(t)\int_0^\Delta \gamma_\tau(s)X(t-s)ds$, and

$$\widetilde{Y}(t) = Q^{(0)}_{Y|X}(t;\tau) - \left(\int_{\mathcal{T}} f_t(0)dt\right)^{-1}\int_{\mathcal{T}}\left(\mathbb{I}\left[Y(t) - Q^{(0)}_{Y|X}(t;\tau)\le 0\right] - \tau\right)dt$$

with $Q^{(0)}_{Y|X}(t;\tau) = \alpha_\tau^{(0)}(t) + \beta_\tau^{(0)}(t)\int_0^\Delta \gamma_\tau^{(0)}(s)X(t-s)ds$. We again write the model (S1.19) as

$$\begin{aligned}\mathrm{E}\left(\widetilde{Y}^c(t)|X^c(s), s\in[t-\Delta,t]\right) &= \beta_\tau(t)\int_0^\Delta \gamma_\tau(s)X^c(t-s)ds \\ &= \int_0^\Delta \varrho_\tau(s;t)X^c(t-s)ds\end{aligned} \tag{S1.20}$$

with $\alpha_\tau(t) = \mu_{\widetilde{Y}}(t) - \int_0^\Delta \varrho_\tau(s;t)\mu_X(t-s)ds$ and $\varrho_\tau(s;t) = \beta_\tau(t)\gamma_\tau(s)$. The functions $\varrho_\tau(s;t)$ include the factor $\gamma(s)$ for each t. Due to $\int_0^\Delta \gamma_\tau^2(s)ds = 1$, for each fixed time point t,

$$\gamma_\tau(s) = \frac{\varrho_\tau(s;t)}{\left[\int_0^\Delta \varrho_\tau^2(s;t)ds\right]^{1/2}}. \tag{S1.21}$$

Once the estimator of $\varrho_\tau(s;t)$ is obtained at a single time point t, it is sufficient to get $\gamma_\tau(s)$ by (S1.21). But, for improving the finite sample behavior and stability of the resulting estimators, we average the representation (S1.21) over an equidistant grid of time $t_1, \cdots, t_R$ in $\mathcal{T}$, that is,

$$\gamma_\tau(s) = \frac{\sum_{r=1}^{R} \varrho_\tau(s;t_r)}{\left[\int_0^\Delta \left(\sum_{r=1}^{R} \varrho_\tau(s;t_r)\right)^2 ds\right]^{1/2}}. \tag{S1.22}$$

Here the number of time points, R, typically would be small. From (S1.20), we know that once the history index function γ_τ is estimated, the model (S1.20) reduces to a functional varying coefficient mean regression model. The first task below is to get an estimate of $\varrho_\tau(\cdot, t)$.

Let $Z_t(s) = X^c(t-s)$ for $s \in [0, \Delta]$ and its auto-covariance function $r_t(s_1, s_2) = \text{Cov}(Z_t(s_1), Z_t(s_2)) = r_{XX}(t-s_1, t-s_2)$ for $s_1, s_2 \in [0, \Delta]$. We have the covariance expansion $r_t(s_1, s_2) = \sum_{m=1}^{\infty} \rho_{tm}\phi_{tm}(s_1)\phi_{tm}(s_2)$ with eigenfunctions ϕ_{tm} and eigenvalues ρ_{tm}. Expanding $\varrho_\tau(s;t) = \sum_{m=1}^{\infty} \varrho_{\tau,m}(t)\phi_{tm}(s)$, $s \in [0, \Delta]$, with suitable expansion coefficients $\varrho_{\tau,m}(t)$ for each $t \in \mathcal{T}$, and having the Karhunen-Loève expansion $Z_t(s) = \sum_{m=1}^{\infty} \zeta_{tm}\phi_{tm}(s)$, with random coefficients $\zeta_{tm} = \int_0^\Delta Z_i(s)\phi_{tm}(s)ds$. Letting a functional linear operator $(\mathcal{L}_{X^c}\varrho_\tau)(t) = \int_0^\Delta \varrho_\tau(s;t)X^c(t-s)ds$, we have the expected SLSE as follows

$$\varrho_\tau^{(1)} = \text{argmin}_{\varrho_\tau} \text{E} \left\| \widetilde{Y}^c - \mathcal{L}_{X^c}\varrho_\tau \right\|_2^2.$$

That is, find the corresponding values $\varrho_{\tau,m}$, $m = 1, 2, \cdots$, satisfying

$$\frac{d}{d\varrho_{\tau,m}(t)} \left[E\left(\widetilde{Y}^c(t) - \sum_{m=1}^{\infty} \varrho_{\tau,m}(t)\zeta_{tm} \right)^2 \right] = 0, \quad m = 1, 2, \cdots.$$

By a straightforward calculation, we have

$$\varrho_\tau^{(1)}(s;t) = \sum_{m=1}^{\infty} \varrho_{\tau,m}^{(1)}(t)\phi_{tm}(s), \qquad \varrho_{\tau,m}^{(1)}(t) = \frac{1}{\rho_{tm}} \int_0^\Delta r_{X\widetilde{Y}}(t-s,t)\phi_{tm}(s)ds. \tag{S1.23}$$

Let $\widetilde{X}(t)=\int_0^{\Delta}\gamma_\tau(s)X^c(t-s)ds$. Once $\gamma_\tau(s)$ is known, the model (S1.19) or (S1.20) reduces to functional linear operator model

$$\mathrm{E}\left(\widetilde{Y}^c(t)|X(s),s\in[t-\Delta,t]\right)=\left(\mathcal{L}_{\widetilde{X}}\beta_\tau\right)(t), \tag{S1.24}$$

where $\left(\mathcal{L}_{\widetilde{X}}\beta_\tau\right)(t)=\beta_\tau(t)\widetilde{X}(t)$, as the model (3.6) of 1st scenario in Section 3.1. Intuitively, we can obtain β_τ by applying the procedure developed in Subsection 3.1.1 via replacing X^c with $\widetilde{X}$. However, in sparse longitudinal settings, the numerical integration involved in estimating $\widetilde{X}$ often does not yield good approximations. For the sparse case, we give a simpler approach that avoids the estimation of $\widetilde{X}$ separately for each subject as Şentürk and Müller (2010) has done. By population least squares for functional linear operator regression, one gets

$$\beta_\tau^{(1)}(t)=\mathrm{argmin}_{\beta_\tau}\mathrm{E}\left\|\widetilde{Y}^c(t)-\beta_\tau(t)\widetilde{X}(t)\right\|_2^2=\mathcal{L}_{X\widetilde{X}}^{-1}E\left(\mathcal{L}_X^*\widetilde{Y}\right)=\frac{\mathrm{Cov}(X(t),\widetilde{Y}(t))}{\mathrm{Cov}(X(t),\widetilde{X}(t))}.$$

From (S1.24), we have

$$\mathrm{Cov}(X(t),\widetilde{X}(t))=\int_0^{\Delta}\gamma_\tau(s)\mathrm{Cov}(X(t),X(t-s)ds=\int_0^{\Delta}\gamma_\tau(s)r_{XX}(t-s,t)ds.$$

Therefore

$$\begin{aligned}\beta_\tau^{(1)}(t)&=\frac{r_{X\widetilde{Y}}(t,t)}{\int_0^{\Delta}\gamma_\tau(s)r_{XX}(t-s,t)ds},\\ \alpha_\tau^{(1)}(t)&=\mu_{\widetilde{Y}}(t)-\beta_\tau^{(1)}(t)\int_0^{\Delta}\gamma_\tau(s)\mu_X(t-s)ds.\end{aligned} \tag{S1.25}$$

Thus, we sketch a one-step iteration of the SLSE algorithm for the functional varying coefficient QR model with history index (S1.18). By combining the above steps, the final estimators $\hat{\alpha}_\tau^{\mathrm{K}}$, $\hat{\beta}_\tau^{\mathrm{K}}$ and $\hat{\gamma}_\tau^{\mathrm{K}}$ via the following Algorithm 2.

Algorithm 2: SLSE algorithm for functional VCQR model with history index (S1.18).

Input: Kernel function $K(\cdot)$, bandwidth h, quantile level τ and the number of iterations K.

Calculate mean function $\hat{\mu}_X$ and $\hat{\mu}_Y$, covariance surface $\hat{r}_{XX}$, cross-covariance surface $\hat{r}_{XY}$, eigenfunctions $\hat{\phi}_k$ and$\hat{\psi}_k$, and eigenvalues $\hat{\rho}_k$ and $\hat{\lambda}_k$ by Steps 1-2.

Initialize estimators $\hat{\gamma}_\tau^{(0)}$ $\hat{\alpha}_\tau^{(0)}$ and $\hat{\beta}_\tau^{(0)}$ by Step 2.

for $k = 1, 2, \cdots, \mathrm{K}$ **do**

Estimate $\hat{Q}^{(k)}_{Y_i|X_i}(T_{ij};\tau)$ $\hat{f}_t^{(k)}(0)$, $\hat{U}_i$ and $\hat{V}_i$ for obtaining $\widetilde{V}_{ij}$ by Step 3.

Compute mean function $\hat{\mu}_{\widetilde{Y}}$ and cross-covariance surface $\hat{r}_{X\widetilde{Y}}$ by Step 4.

Obtain $\hat{\gamma}_\tau^{(k)}$, $\hat{\alpha}_\tau^{(k)}$ and $\hat{\beta}_\tau^{(k)}$ by Step 5.

end

Output: The final estimators $\hat{\gamma}_\tau^{(\mathrm{K})}$, $\hat{\alpha}_\tau^{(\mathrm{K})}$ and $\hat{\beta}_\tau^{(\mathrm{K})}$.

S1.3.2 Steps of Estimation

In the model, dataset is $\{(T_{ij}, U_{ij}, V_{ij}), i = 1, \cdots, n, j = 1, \cdots, N_i\}$. We give some estimations used in Algorithm 2. Now, we present the estimation procedure of $\alpha_\tau^{(1)}$, $\beta_\tau^{(1)}$ and $\gamma_\tau^{(1)}$ as follows. Our data is (T_{ij}, U_{ij}, V_{ij}), $i = 1, \cdots, n$, $j = 1, \cdots, N_i$. Some details are presented in Appendix S1.3.

Step 1: Obtain estimates $\hat{r}_{XX}$, $\hat{r}_{XY}$, $\hat{\mu}_X$ and $\hat{\mu}_Y$ as Steps 1-2 in the 1st scenario. Obtain estimates $\hat{r}_{XX}$, $\hat{r}_{XY}$, $\hat{\mu}_X$ and $\hat{\mu}_Y$ as Steps 1-2 in the 1st scenario. Also see (S1.1)-(S1.3).

Step 2: (1) Given a fixed time t, reversing the time order of the data for all subjects that are observed in the window $[t - \Delta, t]$, estimate $\hat{r}_t$ of covariance surface r_t and obtain estimates $\hat{\phi}_{tm}$, $\hat{\rho}_{tm}$ of the eigenfunctions and eigenvalues ϕ_{tm}, ρ_{tm} of processes Z_t. As (S1.23), and applying numerical integration, we have

$$\hat{\varrho}^{(0)}_{\tau,m}(t) = \frac{1}{\hat{\rho}_{tm}} \int_0^{\Delta} \hat{r}_{XY}(t-s,t)\hat{\phi}_{tm}(s)ds$$
$$\hat{\varrho}^{(0)}_{\tau}(s;t) = \sum_{m=1}^{M_t} \hat{\varrho}^{(0)}_{\tau,m}(t)\hat{\phi}_{tm}(s).$$

(2) Applying (S1.22), and identifiability conditions $\int_0^{\Delta} \gamma_\tau^2(u)du = 1$ and $\gamma_\tau(0) > 0$, one gets the estimated history index function

$$\hat{\gamma}^{(0)}_{\tau}(s) = \frac{\sum_{r=1}^{R} \hat{\varrho}^{(0)}_{\tau}(s;t_r)}{\left[\int_0^{\Delta} \left(\sum_{r=1}^{R} \hat{\varrho}^{(0)}_{\tau}(s;t_r)\right)^2 ds\right]^{1/2}} (-1)^{\mathcal{I}^{(0)}},$$

where $\mathcal{I}^{(0)}$ is the indicator function for $\sum_{r=1}^{R} \hat{\varrho}^{(1)}_{\tau}(0;t_r) < 0$.

(3) As (S1.25), one gets

$$\hat{\beta}^{(0)}_{\tau}(t) = \frac{\hat{r}_{XY}(t,t)}{\int_0^{\Delta} \hat{\gamma}_\tau(s)\hat{r}_{XX}(t-s,t)ds},$$
$$\hat{\alpha}^{(0)}_{\tau}(t) = \mu_Y(t) - \hat{\beta}^{(0)}_{\tau}(t) \int_0^{\Delta} \hat{\gamma}^{(0)}_{\tau}(s)\hat{\mu}_X(t-s)ds.$$

Recall that $Z_t(s) = X^c(t-s)$ for $s \in [0,\Delta]$ and its auto-covariance function $r_t(s_1,s_2) = \mathrm{Cov}(Z_t(s_1), Z_t(s_2)) = r_{XX}(t-s_1, t-s_2)$ for $s_1, s_2 \in [0,\Delta]$. Now, we estimate r_t. Let $S_{ij} = t - T_{ij}$, $S_{il} = t - T_{il}$, $t \in [\Delta, T]$, T_{ij}, $T_{il} \in [t-\Delta, t]$ and $S_{ij}, S_{il} \in [0,\Delta]$. Give the local linear surface smoother for r_t via minimizing

$$\sum_{i=1}^{n} \sum_{1\le j\ne k\le L_i} K_2\left(\frac{S_{ij}-s_1}{h_X}, \frac{S_{ik}-s_2}{h_X}\right) [R_{t,i}(S_{ij}, S_{ik}) - b_0 - b_{11}(s_1 - S_{ij}) - b_{12}(s_2 - S_{ik})]^2,$$

with respect to b_0, b_{11} and b_{12}, where $R_{t,i}(S_{ij}, S_{ik}) = R_{X,i}(T_{ij}, T_{ik})$; then get $\hat{r}_t(s_1,s_2) = \hat{b}_0(s_1,s_2)$.

Based on $\hat{r}_t$, by the eigenequations

$$\int \hat{r}_t(s_1, s_2)\hat{\phi}_{tm}(s_1)ds_1 = \hat{\rho}_{tm}\hat{\phi}_{tm}(s_2)$$

with orthonormal constraints on $\{\hat{\phi}_{tm}\}_{m\geq}$, to obtain $(\hat{\rho}_{tm}, \hat{\phi}_{tm})$ for each $t \in [\Delta, T]$.

Step 3: Calculate surrogate response individual trajectories $\widetilde{V}_{ij} = \widetilde{V}_i(T_{ij})$. During $\widetilde{Y}_{ij}$ involving numerical integration, they often don't yield good approximations in sparse longitudinal settings. We use functional approach as follows:

$$\widetilde{V}_{ij} = \hat{Q}^{(0)}_{Y_i|X_i}(T_{ij};\tau) - \left(\int_{\mathcal{T}} \hat{f}_t(0)dt\right)^{-1} \int_{\mathcal{T}} \left(\mathbb{I}\left[\hat{V}_i^K(t) - \hat{Q}^{(0)}_{Y_i|X_i}(t;\tau) \leq 0\right] - \tau\right) dt,$$

where $\hat{Q}^{(0)}_{Y_i|X_i}(t;\tau) = \hat{\alpha}^{(0)}_\tau(t) + \hat{\beta}^{(0)}_\tau(t)\int_0^\Delta \hat{\gamma}^{(0)}_\tau(s)\hat{U}_i^M(t-s)ds$,

$$\hat{f}_t(0) = \frac{1}{n}\sum_{i=1}^n \frac{1}{N_i}\sum_{j=1}^{N_i} K_h\left(V_{ij} - \hat{\alpha}^{(0)}_\tau(t) - \hat{\beta}^{(0)}_\tau(t)U_{ij}\right)$$

$$\hat{U}_i^M(t) = \hat{\mu}_X(t) + \sum_{m=1}^M \hat{\zeta}_{im}\hat{\phi}_m(t), \quad \hat{V}_i^K(t) = \hat{\mu}_Y(t) + \sum_{k=1}^K \hat{\xi}_{ik}\hat{\psi}_k(t).$$

The estimates $\hat{\zeta}_{im}$ and $\hat{\xi}_{ik}$ can obtained with similar arguments of (S1.4).

Step 4: Based on data $(T_{ij}, U_{ij}, \widetilde{V}_{ij})$, $i = 1, \cdots, n$ and $j = 1, \cdots, N_i$, estimate $\mu_{\widetilde{Y}}$ and $\hat{r}_{X\widetilde{Y}}$, as Step 1.

Based on data $(T_{ij}, U_{ij}, \widetilde{V}_{ij})$, $i = 1, \cdots, n$ and $j = 1, \cdots, N_i$, estimate $\mu_{\widetilde{Y}}$ and $\hat{r}_{X\widetilde{Y}}$ with similar arguments to Step 1.

Step 5: As in Step 2, we obtain successively, for a fixed $t \in \mathcal{T}$,

$$\hat{\varrho}^{(1)}_{\tau,m}(t) = \frac{1}{\hat{\rho}_{tm}}\int_0^\Delta \hat{r}_{X\widetilde{Y}}(t-s,t)\hat{\phi}_{tm}(s)ds$$

$$\hat{\varrho}^{(1)}_\tau(s;t) = \sum_{m=1}^{M_t} \hat{\varrho}^{(1)}_{\tau,m}(t)\hat{\phi}_{tm}(s);$$

the estimated history index function

$$\hat{\gamma}_\tau^{(1)}(s) = \frac{\sum_{r=1}^R \hat{\varrho}_\tau^{(1)}(s;t_r)}{\left[\int_0^\Delta \left(\sum_{r=1}^R \hat{\varrho}_\tau^{(1)}(s;t_r)\right)^2 ds\right]^{1/2}} (-1)^{\mathcal{I}^{(1)}},$$

where $\mathcal{I}^{(1)}$ is the indicator function for $\sum_{r=1}^R \hat{\varrho}_\tau^{(1)}(0;t_r) < 0$; and

$$\hat{\beta}_\tau^{(1)}(t) = \frac{\hat{r}_{X\tilde{Y}}(t,t)}{\int_0^\Delta \hat{\gamma}_\tau^{(1)}(s)\hat{r}_{XX}(t-s,t)ds},$$

$$\hat{\alpha}_\tau^{(1)}(t) = \mu_{\tilde{Y}}(t) - \hat{\beta}_\tau^{(1)}(t)\int_0^\Delta \hat{\gamma}_\tau^{(1)}(s)\hat{\mu}_X(t-s)ds.$$

The local error variance σ_{tX}^2 and the local eigenfunction and eigenvalue estimators are obtained analogously to the global estimates, and estimates $\hat{\sigma}_{tX}^2$ and $\hat{r}_t$ yields estimates of the noise contaminated local covariance surface $\hat{\Sigma}_{t\tilde{Z}}$. The estimate σ_{tX}^2 is similar to the procedure of (S1.8).

S1.3.3 The τth quantile target trajectory in Theorem S1.3.2

Based on our functional VCQR with history index model (S1.18), we give the prediction of the τth quantile response trajectory Y^* for a new subject with the predictor process $\{X^*(s), s \in [t-\Delta, t]\}$ via the following form

$$Q_{Y^*|X^*}(t;\tau) = \alpha_\tau(t) + \beta_\tau(t)\int_0^\Delta \gamma_\tau(s)X^*(t-s)ds, \tag{S1.26}$$

with

$$\gamma_\tau(s) == \frac{\sum_{r=1}^R \sum_{m'=1}^\infty \varrho_{\tau,m'}(t_r)\phi_{t_r m'}(s)}{\left[\int_0^\Delta \left(\sum_{r=1}^R \varrho_\tau(s;t_r)\right)^2 ds\right]^{1/2}}$$

$$X^*(t-s) = \mu_X(t-s) + \sum_{m=1}^\infty \zeta_{tm}^*\phi_{tm}(s).$$

Let $Z^*_{tj} = Z^*_t(T_{tj})$ be the jth measurement for the predictor trajectory $Z^*_t(s) = X^{*c}(t-s)$, $s \in [0, \Delta]$, at time T_{tj}, $j = 1, \cdots, N^*_t$, N^*_t is the random number of measurements, and $\widetilde{\boldsymbol{Z}}^*_t = (\widetilde{Z}^*_{t1}, \cdots, \widetilde{Z}^*_{tN^*_t})$ with the noise contaminated $\widetilde{Z}^*_{tj}$ of Z^*_{tj}. Under the local FPCs ζ^*_{tm} and the measurement errors are jointly Gaussian, we have the best prediction estimates of the scores ζ^*_{tm}, conditional on $\widetilde{\boldsymbol{Z}}^*_t$, N^*_t and $\boldsymbol{T}^*_t = (T^*_{t1}, \cdots, T^*_{tN^*_t})$,

$$\hat{\zeta}^*_{tm} = \hat{\rho}_{tm} \hat{\boldsymbol{\phi}}^{*T}_{tm} \widehat{\boldsymbol{\Sigma}}^{-1}_{t\widetilde{\boldsymbol{Z}}^*_t} \widetilde{\boldsymbol{Z}}^*_t, \tag{S1.27}$$

where $\hat{\boldsymbol{\phi}}^*_{tm}$ and $\widehat{\boldsymbol{\Sigma}}_{t\widetilde{\boldsymbol{Z}}^*_t}$ are the estimates of $\boldsymbol{\phi}^*_{tm} = \left(\phi_{tm}(T^*_{t1}), \cdots, \phi_{tm}(T^*_{tN^*_t})\right)^T$ and $\boldsymbol{\Sigma}_{t\widetilde{\boldsymbol{Z}}^*_t} = \mathrm{Cov}(\widetilde{\boldsymbol{Z}}^*_t | N^*_t, \boldsymbol{T}^*_t)$, respectively; define $\tilde{\zeta}^*_{tm} = \rho_{tm} \boldsymbol{\phi}^{*T}_{tm} \boldsymbol{\Sigma}^{-1}_{t\widetilde{\boldsymbol{Z}}^*_t} \widetilde{\boldsymbol{Z}}^*_t$ analogously. Based on (S1.22) and (S1.26), we obtain the τth predicted quantile trajectories

$$\begin{aligned} \widehat{Q}^{(\mathrm{K},\mathcal{M},M_t)}_{Y^*|X^*}(t;\tau) \;=\;& \hat{\alpha}^{(\mathrm{K})}_\tau(t) + \int_0^\Delta \hat{\varrho}^{(\mathrm{K})}(s;\tau)\hat{\mu}_X(t-s)ds \\ &+ \frac{\hat{\beta}^{(\mathrm{K})}_\tau(t) \sum_{m=1}^{M_t} \hat{\zeta}^*_{tm} \sum_{r=1}^{R} \sum_{m'=1}^{M_r} \hat{\varrho}^{(\mathrm{K})}_{\tau,m'}(t_r) \int_0^\Delta \hat{\phi}_{tm}(s)\hat{\phi}_{t_r m'}(s)ds}{\left[\int_0^\Delta \left(\sum_{r=1}^{R} \hat{\varrho}^{(\mathrm{K})}_\tau(s;t_r)\right)^2 ds\right]^{1/2}}, \end{aligned} \tag{S1.28}$$

where $\mathcal{M} = \sum_{r=1}^{R} M_r$. Define the τth quantile target trajectory

$$\begin{aligned} \widetilde{Q}_{Y^*|X^*}(t;\tau) \;=\;& \alpha_\tau(t) + \int_0^\Delta \varrho(s;\tau)\mu_X(t-s)ds \\ &+ \frac{\beta_\tau(t) \sum_{m=1}^{\infty} \tilde{\zeta}^*_{tm} \sum_{r=1}^{R} \sum_{m'=1}^{\infty} \varrho_{\tau,m'}(t_r) \int_0^\Delta \phi_{tm}(s)\phi_{t_r m'}(s)ds}{\left[\int_0^\Delta \left(\sum_{r=1}^{R} \varrho_\tau(s;t_r)\right)^2 ds\right]^{1/2}}. \end{aligned} \tag{S1.29}$$

In addition, for asymptotic pointwise confidence bands for quantile response trajectories, we have the following some narration.

We here construct asymptotic pointwise confidence bands for the quantile response trajectory, let $\hat{\boldsymbol{\zeta}}_{*,t}^{M_t} = (\hat{\zeta}_{t1}^*, \cdots, \hat{\zeta}_{tM_t}^*)^T$, and define $\tilde{\boldsymbol{\zeta}}_{*,t}^{M_t}$ and $\boldsymbol{\zeta}_{*,t}^{M_t}$ analogously. Call that $\hat{\zeta}_{t1}^* = \hat{\rho}_{tm}\hat{\boldsymbol{\phi}}_{tm}^{*T}\widehat{\boldsymbol{\Sigma}}_{t\widetilde{\boldsymbol{Z}}_t^*}^{-1}\widetilde{\boldsymbol{Z}}_t^*$, $\tilde{\zeta}_{tm}^* = \rho_{tm}\boldsymbol{\phi}_{tm}^{*T}\boldsymbol{\Sigma}_{t\widetilde{\boldsymbol{Z}}_t^*}^{-1}\widetilde{\boldsymbol{Z}}_t^*$ and $\zeta_{tm}^* = \int_0^\Delta Z_i^*(s)\phi_{tm}^*(s)ds$, $m = 1, \cdots, M_t$. One gets a $M_t \times N_t^*$ matrix $\boldsymbol{H}_t = \text{Cov}(\boldsymbol{\zeta}_{*,t}^{M_t}, \widetilde{\boldsymbol{Z}}_t^* | \boldsymbol{T}_t^*, N_t^*) = (\rho_{t1}\boldsymbol{\phi}_{t1}^*, \cdots, \rho_{tM_t}\boldsymbol{\phi}_{tM_t}^*)^T$, $\tilde{\boldsymbol{\zeta}}_{*,t}^{M_t} = \boldsymbol{H}_t\boldsymbol{\Sigma}_{t\widetilde{\boldsymbol{Z}}_t^*}^{-1}\widetilde{\boldsymbol{Z}}_t^*$, and $\text{Cov}(\tilde{\boldsymbol{\zeta}}_{*,t}^{M_t}|\boldsymbol{T}_t^*, N_t^*) = \text{Cov}(\tilde{\boldsymbol{\zeta}}_{*,t}^{M_t}, \boldsymbol{\zeta}_{*,t}^{M_t}|\boldsymbol{T}_t^*, N_t^*) = \boldsymbol{H}_t\boldsymbol{\Sigma}_{t\widetilde{\boldsymbol{Z}}_t^*}^{-1}\boldsymbol{H}_t^T$. With similar arguments as Subsections 3.1.3 and 3.2.3, given $\boldsymbol{T}_t^*$ and N_t^*, we have

$$\tilde{\boldsymbol{\zeta}}_{*,t}^{M_t} - \boldsymbol{\zeta}_{*,t}^{M_t} \sim N(\mathbf{0}, \boldsymbol{\Omega}_{tM_t}),$$

where $\boldsymbol{\Omega}_{tM_t} = \boldsymbol{D}_t - \boldsymbol{H}_t\boldsymbol{\Sigma}_{t\widetilde{\boldsymbol{Z}}_t^*}^{-1}\boldsymbol{H}_t^T$ and $\boldsymbol{D}_t = \text{diag}(\rho_{t1}, \cdots, \rho_{tM_t})$.

Further, define $\hat{\boldsymbol{\phi}}_{t,\mathcal{M}}^{(\text{K})} = \hat{\beta}_\tau^{(\text{K})}(t)\left(\widehat{\widetilde{X}}_{t1}^{(c,\text{K})}, \cdots, \widehat{\widetilde{X}}_{tM_t}^{(c,\text{K})}\right)^T$, where

$$\widehat{\widetilde{X}}_{tm}^{(c,\text{K})} = \frac{\sum_{r=1}^R \sum_{m'=1}^{M_r} \hat{\varrho}_{\tau,m'}^{(\text{K})}(t) \int_0^\Delta \hat{\phi}_{tm}(s)\hat{\phi}_{tm'}(s)ds}{\left[\int_0^\Delta \left(\sum_{r=1}^R \hat{\varrho}_\tau^{(\text{K})}(s;t_r)\right)^2 ds\right]^{1/2}}, \quad m = 1, \cdots, M_t;$$

denote $\hat{\boldsymbol{\Omega}}_{tM_t} = \hat{\boldsymbol{D}}_t - \hat{\boldsymbol{H}}_t\widehat{\boldsymbol{\Sigma}}_{t\widetilde{\boldsymbol{Z}}_t^*}^{-1}\hat{\boldsymbol{H}}_t^T$ with $\hat{\boldsymbol{D}}_t = \text{diag}(\hat{\rho}_{t1}, \cdots, \hat{\rho}_{tM_t})$ and $\hat{\boldsymbol{H}}_t = (\hat{\rho}_{t1}\hat{\boldsymbol{\phi}}_{t1}^*, \cdots, \hat{\rho}_{tM_t}\hat{\boldsymbol{\phi}}_{tM_t}^*)^T$.

The τth quantile predicted trajectories are estimated by

$$\widehat{Q}_{Y^*|X^*}^{\mathcal{M},M_t,\text{K}}(t;\tau) = \hat{\alpha}_\tau^{(\text{K})}(t) + \int_0^\Delta \hat{\varrho}^{(\text{K})}(s;\tau)\hat{\mu}_X(t-s)ds + \left(\hat{\boldsymbol{\zeta}}_{*,t}^{M_t}\right)^T \hat{\boldsymbol{\phi}}_{t,\mathcal{M}}^{(\text{K})},$$

S1.3.4 Asymptotic properties

We first provide uniform consistency for history index γ_τ^{K}, varying coefficients α_τ^{K} and β_τ^{K} in the functional varying coefficient QR model (S1.18).

Theorem S1.3.1. *Let* $\sup_{s\in[0,\Delta]} |\hat{\gamma}_\tau^{(0)}(s) - \gamma_\tau(s)| = O_p(a_n)$ $\sup_{t\in\mathcal{T}} |\hat{\alpha}_\tau^{(0)}(t) - \alpha_\tau(t)| = O_p(a_n)$ *and* $\sup_{t\in\mathcal{T}} |\hat{\beta}_\tau^{(0)}(t) - \beta_\tau(t)| = O_p(a_n)$. *Under Conditions 1-3, and Assumptions (A1)-(A6)* *and (C1)-(C2) in the Appendix S5, we have for k iterations of Algorithm 2,*

$$\sup_{s\in[0,\Delta]} |\hat{\gamma}_\tau^{(k)}(s) - \gamma_\tau(s)| = O_p\left\{\frac{1}{\sqrt{n}}\left[\frac{1}{b_X} + \frac{1}{b_Y} + \frac{1}{h_X^2} + \frac{1}{h_Y^2} + \frac{1}{h_1 h_2}\right] + \delta_{3n} + a_n^{k+1}\right\},$$
$$\sup_{t\in\mathcal{T}} |\hat{\alpha}_\tau^{(k)}(t) - \alpha_\tau(t)| = O_p\left\{\frac{1}{\sqrt{n}}\left[\frac{1}{b_X} + \frac{1}{b_Y} + \frac{1}{h_X^2} + \frac{1}{h_Y^2} + \frac{1}{h_1 h_2}\right] + \delta_{3n} + a_n^{k+1}\right\}, \quad \text{(S1.30)}$$
$$\sup_{t\in\mathcal{T}} |\hat{\beta}_\tau^{(0)}(t) - \beta_\tau(t)| = O_p\left\{\frac{1}{\sqrt{n}}\left[\frac{1}{b_X} + \frac{1}{b_Y} + \frac{1}{h_X^2} + \frac{1}{h_Y^2} + \frac{1}{h_1 h_2}\right] + \delta_{3n} + a_n^{k+1}\right\}.$$

The τth quantile target trajectory $\widetilde{Q}_{Y^*|X^*}(t;\tau)$ is defined in Subsection S1.3.3 of SM.

Theorem S1.3.2. *Let* $\sup_{s\in[0,\Delta]} |\hat{\gamma}_\tau^{(0)}(s) - \gamma_\tau(s)| = O_p(a_n)$ $\sup_{t\in\mathcal{T}} |\hat{\alpha}_\tau^{(0)}(t) - \alpha_\tau(t)| = O_p(a_n)$ *and* $\sup_{t\in\mathcal{T}} |\hat{\beta}_\tau^{(0)}(t) - \beta_\tau(t)| = O_p(a_n)$ *with* $0 \le a_n < 1$. *Under Conditions 1-3, and Assumptions (A1)-(A7) and (C1)-(C2) in the Appendix S5, given* N^* *and* $\boldsymbol{T}^*$, *for all* $t \in \mathcal{T}$, *the predicted* τ*th quantile response trajectories in the functional varying coefficient QR with history index model (S1.18) satisfy*

$$\lim_{n\to\infty} \widehat{Q}_{Y^*|X^*}^{(\mathrm{K},\mathcal{M},M_t)}(t;\tau) = \widetilde{Q}_{Y^*|X^*}(t;\tau), \quad \text{in proabaility},$$

with $M_t(n), M_1(n), \cdots, M_R(n) \to \infty$ *as* $n \to \infty$, *and the iteration number* K *enough large.*

The τth quantile predicted trajectories are estimated by

$$\widehat{Q}_{Y^*|X^*}^{\mathcal{M},M_t,\mathrm{K}}(t;\tau) = \hat{\alpha}_\tau^{(\mathrm{K})}(t) + \int_0^\Delta \hat{\varrho}^{(\mathrm{K})}(s;\tau)\hat{\mu}_X(t-s)ds + \left(\hat{\boldsymbol{\zeta}}_{*,t}^{M_t}\right)^T \hat{\boldsymbol{\phi}}_{t,\mathcal{M}}^{(\mathrm{K})},$$

The following result provides its asymptotic distribution.

Theorem S1.3.3. *Let* $\sup_{s\in[0,\Delta]} |\hat{\gamma}_\tau^{(0)}(s) - \gamma_\tau^*(s)| = O_p(a_n)$ $\sup_{t\in\mathcal{T}} |\hat{\alpha}_\tau^{(0)}(t) - \alpha_\tau^*(t)| = O_p(a_n)$ *and* $\sup_{t\in\mathcal{T}} |\hat{\beta}_\tau^{(0)}(t) - \beta_\tau^*(t)| = O_p(a_n)$ *with* $0 \le a_n < 1$. *Under Conditions 1-3, and Assumptions (A1)-(A7), (A8)(iii) and (C1)-(C2) in the Appendix S5, given* N^* *and* $\boldsymbol{T}^*$, *for a given*

$\tau \in (0,1)$, $\{X^*(s), s \in [t-\Delta, t]\}$, *all* $t \in \mathcal{T} = [\Delta, T]$, $x \in \mathbb{R}$,

$$\lim_{n\to\infty} P\left(\frac{\widehat{Q}^{\mathrm{K},\mathcal{M},M_t}_{Y^*|X^*}(t;\tau) - Q_{Y^*|X^*}(t;\tau)}{\sqrt{\hat{\omega}^{\mathrm{K},\mathcal{M},M_t}_{\tau}(t)}} \le x\right) = \Phi(x),$$

where $\hat{\omega}^{\mathrm{K},\mathcal{M},M_t}_{\tau}(t) = \left(\hat{\boldsymbol{\phi}}^{(\mathrm{K})}_{t,\mathcal{M}}\right)^T \hat{\boldsymbol{\Omega}}_{tM_t}\hat{\boldsymbol{\phi}}^{(\mathrm{K})}_{t,\mathcal{M}}$ *is a estimator of* $\omega^{\mathcal{M},M_t}_{\tau}(t) = \left(\boldsymbol{\phi}_{t,\mathcal{M}}\right)^T \boldsymbol{\Omega}_{tM_t}\boldsymbol{\phi}_{t,\mathcal{M}}$ *as* $M_t(n), M_1(n), \cdots, M_R(n) \to \infty$ *when* $n \to \infty$*, and the iteration number* K *enough large.*

As a consequence, the $(1-\alpha)100\%$ asymptotic pointwise confidence bands for $Q_{Y^*|X^*}(t;\tau)$, given $X^*(s)$, $s \in [t-\Delta, t]$, is constructed by

$$\widehat{Q}^{\mathrm{K},\mathcal{M},M_t}_{Y^*|X^*}(t;\tau) \pm \Phi\left(1 - \frac{\alpha}{2}\right)\sqrt{\hat{\omega}^{\mathrm{K},\mathcal{M},M_t}_{\tau}(t)}.$$

S2 Choice of hyper-parameter

The choice of hyper-parameter M (the number of eigenfunctions), which is used in Sections 3.1 and S1.3, has been considered in Yao et al. (2005b). We excerpt from Yao et al. (2005b) as follows, for the readers to read smoothly.

One-curve-leave-out cross-validation aims to minimize

$$\mathrm{CV}_X(M) = \sum_{i=1}^{n}\sum_{j=1}^{N_i}[U_{ij} - \hat{X}_i^{(-i)}(T_{ij})]^2$$

with respect to M, where $\hat{X}_i^{(-i)}(t) = \hat{\mu}_X^{(-i)}(t) + \sum_{i=1}^{M}\hat{\zeta}_{im}^{(-i)}\hat{\phi}_m^{(-i)}(t)$, and $\hat{\zeta}_{im}^{(-i)}$ is calculated by (S1.4), $\hat{\phi}_m^{(-i)}(t)$ and $\hat{\phi}_m^{(-i)}$ are the estimated mean and eigenfunctions after removing the data for X_i.

The ACI criterion as a function of M is as follows:

$$\begin{aligned}\mathrm{AIC}(M) &= \sum_{i=1}^{n}\left\{\frac{1}{2\hat{\sigma}_X^2}\left(\tilde{\boldsymbol{U}}_i - \hat{\boldsymbol{\mu}}_{X_i} - \sum_{m=1}^{M}\hat{\zeta}_{im}\hat{\phi}_{im}\right)^T\left(\tilde{\boldsymbol{U}}_i - \hat{\boldsymbol{\mu}}_{X_i} - \sum_{m=1}^{M}\hat{\zeta}_{im}\hat{\phi}_{im}\right)\right.\\ &\quad \left. + \frac{L_i}{2}\log(2\pi) + \frac{L_i}{2}\log\hat{\sigma}_X^2\right\} + M,\end{aligned}$$

where $\tilde{\boldsymbol{U}}_i = (U_{i1}, \cdots, U_{iN_i})^T$, $\hat{\boldsymbol{\mu}}_{X_i} = (\hat{\mu}_X(T_{i1}), \cdots, \hat{\mu}_X(T_{iN_i}))^T$, $\hat{\boldsymbol{\phi}}_{im} = (\hat{\phi}_m(T_{i1}), \cdots, \hat{\phi}_m(T_{iN_i}))^T$, and $\hat{\zeta}_{im}$ is calculated by (S1.4). For the response process Y, we proceed analogously for the corresponding estimates for the components of model (3.2).

The number of eigenfunctions included in the local expansions of Section S1.3, M_t or M_r when $t = t_r$ are chosen analogously by $\mathrm{AIC}(M_t)$, where in the above definition N_i, $\hat{\mu}_{X_i}$, $\tilde{\boldsymbol{U}}_i$, $\hat{\boldsymbol{\phi}}_{im}$, $\hat{\zeta}_{im}$ and $\hat{\sigma}_X^2$ are replaced by their local counterparts at t.

S3 Algorithm S

We present Algorithm S for functional linear QR model.

Algorithm S: SLSE algorithm for functional linear operator QR model (3.14).

Input: Kernel functions $K(\cdot)$, $K_1(\cdot)$ and $K_2(\cdot)$, bandwidths h, h_b, h_X, h_Y, h_1 and h_2, quantile level τ, and the numbers of eigenfunctions M and K, and of iterations K.

Calculate mean function $\hat{\mu}_X$ and $\hat{\mu}_Y$, covariance surface $\hat{r}_{XX}$, cross-covariance surface $\hat{r}_{XY}$, eigenfunctions $\hat{\phi}_k$ and$\hat{\psi}_k$, and eigenvalues $\hat{\rho}_k$ and $\hat{\lambda}_k$ by Steps 1-2.

Initialize estimators $\hat{\alpha}_\tau^{(0)}$ and $\hat{\beta}_\tau^{(0)}$ by Step 3.

for $k = 1, 2, \cdots, \mathrm{K}$ **do**

Estimate $\hat{f}_t^{(k)}(0)$, $\hat{\omega}_i$, $\hat{U}_i$ and $\hat{V}_i$ for obtaining $\widetilde{V}_{ij}$ by Step 4.

Compute mean function $\hat{\mu}_{\widetilde{Y}}$ via Step 5, cross-covariance surfaces $\hat{r}_{\widetilde{Y}\widetilde{Y}}$ and $\hat{r}_{X\widetilde{Y}}$ by Step 6.

Obtain $\hat{\alpha}_\tau^{(k)}$ and $\hat{\beta}_\tau^{(k)}$ by Step 7.

end

Output: The final estimators $\hat{\alpha}_\tau^{(\mathrm{K})}$ and $\hat{\beta}_\tau^{(\mathrm{K})}$.

S4 Proof of Theorem 2.1

Proof: For the surrogate least squares estimation (2.13), we can rewrite it as a linear operator regression model

$$\widetilde{Y}(t) = \mathcal{L}_X \beta_\tau^{(1)}(t) + \epsilon(t)$$

with $\epsilon \in L_2(\mathcal{T})$ is a random error process, with the assumption that X and ϵ are uncorrelated, and that $E[\epsilon(t)] = 0$ for all t. Applying $\mathcal{L}_X^*$ to both sides of the above linear operator regression model, and taking expectations, we have

$$E[\mathcal{L}_X^* \widetilde{Y}] = E[\mathcal{L}_X^* \mathcal{L}_X] \beta_\tau^{(1)} + E[\mathcal{L}_X^* \epsilon],$$

where $E[\mathcal{L}_X^* \mathcal{L}_X] = \Gamma_{XX}$ and $E[\mathcal{L}_X^* \epsilon] = 0$. Hence, one obtains the functional normal equation

$$\Gamma_{XX} \beta_\tau^{(1)} = E[\mathcal{L}_X^* \widetilde{Y}].$$

The proof for (a) follows from Conway (1985) and He et al. (2000).

For the proof of (b), we first have

$$\begin{aligned}
& E[\mathcal{L}_X^* \widetilde{Y}] - \Gamma_{XX} \beta_\tau \\
= \; & E[\mathcal{L}_X^* \mathcal{L}_X \beta_\tau^{(0)}] - \left(\int_{\mathcal{T}} f_t(0) dt \right)^{-1} E[\mathcal{L}_X^* \int_{\mathcal{T}} \{ \mathbb{I}[Y(t) - (\mathcal{L}_X \beta_\tau^{(0)})(t) \leq 0] - \tau \} dt] - \Gamma_{XX} \beta_\tau \\
= \; & \Gamma_{XX} (\beta_\tau^{(0)} - \beta_\tau) - \left(\int_{\mathcal{T}} f_t(0) dt \right)^{-1} E\left[\mathcal{L}_X^* \int_{\mathcal{T}} \{ F_t(\mathcal{L}_X(\beta_\tau^{(0)} - \beta_\tau)(t)) - F_t(0) \} dt \right].
\end{aligned}$$

By second order Taylor expansion, under Condition 3 we have for the second term,

$$\begin{aligned}
& \left(\int_{\mathcal{T}} f_t(0)dt\right)^{-1} E\left[\mathcal{L}_X^* \int_{\mathcal{T}} \{F_t(\mathcal{L}_X(\beta_\tau^{(0)} - \beta_\tau)(t)) - F_t(0)\} dt\right] \\
= & \left(\int_{\mathcal{T}} f_t(0)dt\right)^{-1} E\left[\mathcal{L}_X^* \int_{\mathcal{T}} \left\{ f_t(0)(\mathcal{L}_X(\beta_\tau^{(0)} - \beta_\tau)(t)) + C(\mathcal{L}_X(\beta_\tau^{(0)} - \beta_\tau)(t))^2 \right\} dt\right] \\
= & \Gamma_{XX}(\beta_\tau^{(0)} - \beta_\tau) + C\left(\int_{\mathcal{T}} f_t(0)dt\right)^{-1} E\left[\mathcal{L}_X^*(\mathcal{L}_X(\beta_\tau^{(0)} - \beta_\tau)(t))^2\right] \\
= & \Gamma_{XX}(\beta_\tau^{(0)} - \beta_\tau) + C\left(\int_{\mathcal{T}} f_t(0)dt\right)^{-1} E\left[\mathcal{L}_X^* \left\langle \mathcal{L}_X(\beta_\tau^{(0)} - \beta_\tau)(t), \mathcal{L}_X(\beta_\tau^{(0)} - \beta_\tau)(t) \right\rangle\right] \\
= & \Gamma_{XX}(\beta_\tau^{(0)} - \beta_\tau) + C\left(\int_{\mathcal{T}} f_t(0)dt\right)^{-1} E\left[\mathcal{L}_X^* \left\langle (\beta_\tau^{(0)} - \beta_\tau)(t), \mathcal{L}_X^*\mathcal{L}_X(\beta_\tau^{(0)} - \beta_\tau)(t) \right\rangle\right].
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
E[\mathcal{L}_X^* \widetilde{Y}] - \Gamma_{XX}\beta_\tau &= C\left(\int_{\mathcal{T}} f_t(0)dt\right)^{-1} E\left[\mathcal{L}_X^* \left\langle (\beta_\tau^{(0)} - \beta_\tau)(t), \mathcal{L}_X^*\mathcal{L}_X(\beta_\tau^{(0)} - \beta_\tau)(t) \right\rangle\right] \\
&= O\left(E\left[\mathcal{L}_X^* \left\langle (\beta_\tau^{(0)} - \beta_\tau)(t), \mathcal{L}_X^*\mathcal{L}_X(\beta_\tau^{(0)} - \beta_\tau)(t) \right\rangle\right]\right) \\
&= O\left(\left\langle (\beta_\tau^{(0)} - \beta_\tau)(t), E\left[\mathcal{L}_X\mathcal{L}_X^*\mathcal{L}_X\right](\beta_\tau^{(0)} - \beta_\tau)(t) \right\rangle\right) \\
&= O\left(\left\| \left[E^{1/2}(\mathcal{L}_X\mathcal{L}_X^*\mathcal{L}_X)\right](\beta_\tau^{(0)} - \beta_\tau) \right\|^2\right).
\end{aligned}$$

Thus, it completes the proof of (b). By the k rounds of iteration, we get (c). So, we complete the proof of Theorem 2.1.

S5 Proofs of the main resluts for specific functional QR for sparse longitudinal data

In the section, we respectively give the proofs the three FLQR models. Subsection S5.1: functional varying coefficient QR model; Subsection S5.2: functional linear QR model; Subsection S5.3: functional varying coefficient QR model with history index. First, we present a common set of assumptions needed for all FLQR models, which are listed under (A).

The data (S_{il}, U_{il}) and (T_{ij}, V_{ij}), $i = 1, \cdots, n$, $l = 1, \cdots, L_i$, $j = 1, \cdots, N_i$, as described in (3.1) and (3.2), are assumed to have the same distributed as (S, U) and (T, V), with joint densities $g_1(s, x)$ and $g_2(t, y)$. Assume also that the observation times/locations S_{il} are i.i.d. with marginal densities $f_S(s)$; T_{ij} are i.i.d. with marginal densities $f_T(t)$. Let S_1 and S_2 be i.i.d. as S, and U_1 and U_2 be repeated observations of X made on the same subject at times/locations S_1 and S_2 separately. The predictor and response measurements made on the same at different times/locations are allowed to be dependent. Assume $(S_{il_1}, S_{il_2}, U_{il_1}, U_{il_2})$, $1 \le l_1 \ne l_2 \le L_i$, is identically distributed as (S_1, S_2, U_1, U_2) with joint density function $g_X(s_1, s_2, u_1, u_2)$, and analogously for $(T_{ij_1}, T_{ij_2}, V_{ij_1}, V_{ij_2})$ with identical joint density function $g_Y(t_1, t_2, v_1, v_2)$. About the above (joint) density functions, we give some regularity assumptions.

(A1) Let p_1 and p_2 be integers with $0 \le p_1, p_2 \le p = p_1 + p_2 = 2$. The derivative $(d^p/ds^p)f_S(s)$ and $(d^p/dt^p)f_T(t)$ exist and are continuous on $s \in \mathcal{S}$ and $t \in \mathcal{T}$ with $f_S(s) > 0$ and $f_T(t) > 0$ on $s \in \mathcal{S}$ and $t \in \mathcal{T}$, respectively; $(d^p/ds^p)g_1(s, u)$ and $(d^p/dt^p)g_2(t, v)$ exist and are continuous on $\mathcal{S} \times \mathbb{R}$ and $\mathcal{T} \times \mathbb{R}$, respectively; $(d^p/ds_1^{p_1}ds_2^{p_2})g_X(s_1, s_2, u_1, u_2)$ and $(d^p/dt_1^{p_1}dt_2^{p_2})g_Y(t_1, t_2, v_1, v_2)$ exist and are continuous on $\mathcal{S}^2 \times \mathbb{R}^2$ and $\mathcal{T}^2 \times \mathbb{R}^2$, respectively.

(A2) The number of measurements L_i and N_i made on the ith subject are random variables such that $L_i \overset{i.i.d.}{\sim} L$, $N_i \overset{i.i.d.}{\sim} N$, where L and N are positive discrete random variables, with $P(L > 1) > 0$ and $P(N > 1) > 0$. The observation times/locations are assumed to be independent of the number of measurements, i.e., for any subsets $\mathscr{L}_i \subseteq \{1, \cdots, L_i\}$ and $\mathscr{N}_i \subseteq \{1, \cdots, N_i\}$, and for all $i = 1, \cdots, n$, $(\{S_{il}, U_{il} : l \in \mathscr{L}_i\})$ is independent of L_i, and $(\{T_{il}, V_{il} : l \in \mathscr{N}_i\})$ is independent of N_i.

Let $K(\cdot)$ be the nonnegative univariate kernel function that is used in the kernel density estimator of $f_t(0)$ for $\varepsilon(t)$ at zero, and $K_1(\cdot)$ and $K_2(\cdot, \cdot)$ be the nonnegative univariate and bivariate kernel functions that are applied to the smoothing for the mean function μ_X and μ_Y, covariance surface r_{XX}, r_{YY}, $r_{\widetilde{Y}\widetilde{Y}}$, cross-covariance surface r_{XY} and $r_{X\widetilde{Y}}$, and local covariance surfaces r_t (see Section S1). In addition, assume that K_1 and K_2 are compactly supported densities with zero means and finite variances. Let h be the bandwidth used for density estimator of $f_t(0)$, $b_X = b_X(n)$ and $b_Y = b_Y(n)$ be the bandwidths used for the mean functions such as μ_X, μ_Y, and so on; $h_X = h_X(n)$ and $h_Y = h_Y(n)$ be the bandwidths used for covariance surfaces such as r_{XX}, r_{XY}, and so on; and $h_1 = h_1(n)$ and $h_2 = h_2(n)$ be the bandwidths for obtaining cross-covariance surfaces such as r_{XY}, $r_{X\widetilde{Y}}$, and so on. Further, we define the Fourier transformations of $K_1(u)$ and $K_2(u, v)$ as $\kappa_1(t) = \int e^{-iut}K_1(u)du$ and $\kappa_2(t, s) = \int\int e^{-(iut+ivt)}K_2(u, v)dudv$, respectively. We give some assumptions about kernel functions and bandwidths as follows.

(A3) The kernel function $K(\cdot)$ is integrable with $\int_{-\infty}^{\infty} K(u)du = 1$, and $K(u) = 0$ if $|u| \geq 1$. Further, assume $K(\cdot)$ is differentiable and its derivative $K'(\cdot)$ is bounded; The Fourier transformation $\kappa_1(t)$ is absolutely integrable, i.e., $\int |\kappa_1(t)|dt < \infty$, and $\kappa_2(t,s)$ also is absolutely integrable, i.e., $\int\int |\kappa_2(t,s)|dtds < \infty$. As the number of subjects $n \to \infty$, $h \to 0$ and $nh \to \infty$; $b_X \to 0$, $b_Y \to 0$, $nb_X^4 \to \infty$, $nb_Y^4 \to \infty$, $nb_X^6 < \infty$, $nb_Y^6 < \infty$; $h_X \to 0$, $h_Y \to 0$, $nh_X^6 \to \infty$, $nh_Y^6 \to \infty$, $nh_X^8 < \infty$, $nh_Y^8 < \infty$.

(A4) Assume that U and V have finite fourth moments, that is, $E[(U - \mu_X(S))^4] < \infty$ and $E[(V - \mu_X(T))^4] < \infty$.

(A5) Assume that the FPC scores ζ_{il} and measurement errors $\varepsilon_{X,il}$ in (3.1) are jointly Gaussian. In addition, ξ_{ij} and measurement errors $\varepsilon_{Y,ij}$ in (3.2) are also jointly Gaussian. Special, for the functional varying coefficient QR model with history index, assume that the FPC scores ζ_{itl} $(1 \leq l \leq N_{it}, t \in \mathcal{T})$ and the measurement errors $\varepsilon_{X,itl}$ are jointly Gaussian.

(A6) Assume that the numbers $M = M(n)$ and $K = K(n)$ of included eigenfunctions depend on the sample size, such that $M(n) \to \infty$ and $K(n) \to \infty$ as $n \to \infty$. And they satisfy the rate conditions given in assumption (B5) of Yao et al. (2005b).

(A7) The number and locations of measurements for a subject or cluster remain unaltered as the sample size $n \to \infty$.

(A8) (i) There exists a continuous positive definite function $\omega_\tau^M(t)$ such that $\omega_\tau^M(t) \to \omega_\tau(t)$, as $M \to \infty$. (ii) There exists a continuous positive definite function $\omega_\tau^{M,K} \to \omega_\tau(t)$ as $M, K \to \infty$. (iii) For all $t \in \mathcal{T}$, there exists a continuous positive define function $\omega_\tau^{\mathcal{M},M_t}(t) \to \omega_\tau(t)$ as $M_t, M_1, \cdots, M_R \to \infty$.

Next, we list some special assumptions. Assumptions (B) is needed for the functional linear QR models; and Assumptions (C) for the functional varying coefficient QR models with history index.

(B1) Assume that $\tilde{M} = \tilde{M}(n)$ and $\tilde{K} = \tilde{K}(n)$ of included eigenfunctions depend on the sample size in (S1.13), such that as $n \to \infty$, $\tilde{M}(n) \to \infty$ and $\tilde{K}(n) \to \infty$, and

$$\delta_{1n} = \frac{\tilde{M}\tilde{K}}{\sqrt{n}} \left(\frac{1}{b_X} + \frac{1}{b_Y} + \frac{1}{h_X^2} + \frac{1}{h_Y^2} + \frac{1}{h_1 h_2} \right) \to 0.$$

(B2) Assume the remainder as $\tilde{M}(n) \to \infty$, $\tilde{K}(n) \to \infty$ as

$$\delta_{2n} = \sup_{(s,t)\in\mathcal{S}\times\mathcal{T}} \left| \sum_{m=\tilde{M}+1}^{\infty} \sum_{k=\tilde{K}+1}^{\infty} \frac{\tilde{\sigma}_{mk}}{\rho_m} \phi_m(s)\varphi_k(t) \right| \to 0.$$

(C1) The number of included eigenfunctions from local eigen-decompositions M_t or $M_r := M_{t_r}$ are integer valued sequences that depend on n with $\inf_{t\in[\Delta,T]M_t(n)} \to \infty$, and both $\inf_{t\in[\Delta,T]M_t(n)}$ and $\sup_{t\in[\Delta,T]M_t(n)}$ satisfy the rate conditions given in assumption (B5) of Yao et al. (2005b). Further, the linear operator regression coefficient $\varrho_\tau(s;t)$ in the model (S1.20) satisfies $\int_{t\in\mathcal{T}} \int_{s\in[0,\Delta]} \varrho_\tau^2(s;t)dsdt < \infty$.

(C2) Assume that the remainder as $M_t(n) \to \infty$ as

$$\delta_{3n} = \sup_{s\in[0,\Delta],t\in\mathcal{T}} \left| \sum_{m=M_t+1}^{\infty} \varrho_{\tau,m}^*(t)\phi_{tm}(s) \right| \to 0.$$

S5.1 For functional varying coefficient QR model

In the subsection, we present the proofs of Theorems 3.1-3.3.

Proof of Theorem 3.1 Recall that $\hat{\beta}_\tau^{(1)}(t) = \frac{\hat{r}_{X\tilde{Y}}(t,t)}{\hat{r}_{XX}(t,t)}$ and $\hat{\alpha}_\tau^{(1)}(t) = \hat{\mu}_{\tilde{Y}}(t) - \hat{\beta}_\tau^{(1)}(t)\hat{\mu}_X(t)$. By

Theorem 2.1, we have

$$
\begin{aligned}
|\hat{\beta}_{\tau}^{(1)}(t) - \beta_{\tau}^{*}(t)| &\leq \left|\hat{\beta}_{\tau}^{(1)}(t) - \beta_{\tau}^{(1)}(t)\right| + \left|\beta_{\tau}^{(1)}(t) - \beta_{\tau}^{*}(t)\right| \\
&= \left|\frac{\hat{r}_{X\widetilde{Y}}}{\hat{r}_{XX}} - \frac{r_{X\widetilde{Y}}}{r_{XX}}\right| + O\left(E\|X\|^3 \left|\beta_{\tau}^{(0)} - \beta_{\tau}^{*}\right|^2\right). \qquad \text{(S5.1)}
\end{aligned}
$$

Uniform consistency of $\hat{r}_{XX}$, μ_X and μ_Y follow from Theorem 1 of Yao et al. (2005a), that is,

$$
\begin{aligned}
\sup_{t\in\mathcal{T}} |\hat{\mu}_X(t) - \mu_X(t)| &= O_p\left(\frac{1}{\sqrt{n}b_X}\right), \\
\sup_{t\in\mathcal{T}} |\hat{\mu}_Y(t) - \mu_Y(t)| &= O_p\left(\frac{1}{\sqrt{n}b_Y}\right), \\
\sup_{s,t\in\mathcal{T}} |\hat{r}_{XX}(s,t) - r_{XX}(s,t)| &= O_p\left(\frac{1}{\sqrt{n}h_X^2}\right).
\end{aligned}
$$

We consider $\hat{r}_{X\widetilde{Y}}$. In the local linear estimator for covariance $r_{X\widetilde{Y}}$, we use the raw observations $R_{i,X\widetilde{Y}} = (U_{ij} - \hat{\mu}_X(T_{ij}))\left(\widetilde{V}_{ik} - \hat{\mu}_{\widetilde{Y}}(T_{ik})\right)$. Call that

$$
\widetilde{V}_{ij} = \hat{\alpha}_{\tau}^{(0)}(T_{ij}) + \hat{\beta}_{\tau}^{(0)}(T_{ij})U_{ij} - \left(\int_{\mathcal{T}} \hat{f}_t(0)dt\right)^{-1} \int_{\mathcal{T}} \frac{1}{n}\sum_{i=1}^{n} \hat{\omega}_i(t)dt,
$$

and $\hat{\omega}_i(t) = \mathbb{I}\left[\hat{V}_i^K(t) - \hat{\alpha}_{\tau}^{(0)}(t) - \hat{\beta}_{\tau}^{(0)}(t)\hat{U}_i^M(t) \leq 0\right] - \tau$. Under Condition (A6), by Theorems 2-3 of Yao et al. (2005a), we have $\lim_{M\to\infty}\lim_{n\to\infty} \hat{U}_i^M(t) = X_i(t) + O_p\left\{1/(\sqrt{n}b_X) + 1/(\sqrt{n}h_X^2)\right\}$ and

$$
\lim_{K\to\infty}\lim_{n\to\infty} \hat{V}_i^K(t) = Y_i(t) + O_p\left\{1/(\sqrt{n}b_Y) + 1/(\sqrt{n}h_Y^2)\right\}
$$

for all $t \in \mathcal{T}$. Thus,

$$
\lim_{M,K\to\infty}\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \hat{\omega}_i(t) = \omega(t) + O_p\left(\frac{1}{\sqrt{n}b_X} + \frac{1}{\sqrt{n}b_Y} + \frac{1}{\sqrt{n}h_X^2} + \frac{1}{\sqrt{n}h_Y^2}\right) \qquad \text{(S5.2)}
$$

for all $t \in \mathcal{T}$, where $\omega(t) = \mathbb{I}\left[Y(t) - \hat{\alpha}_\tau^{(0)}(t) - \hat{\beta}_\tau^{(0)}(t)X(t) \leq 0\right] - \tau$. Similar to the proof of Lemma 9 of Chen et al. (2020), we have

$$\sup_{t\in\mathcal{T}} |\hat{f}_t(0) - f_t(0)| = O_p\left(\sqrt{\frac{\log n}{nh}} + a_n + h\right)$$

for all $t \in \mathcal{T}$.

We know that

$$\begin{aligned} R_{i,X\widetilde{Y}} &= \{[U_{ij} - \mu_X(T_{ij})] + [\mu_X(T_{ij}) - \hat{\mu}_X(T_{ij})]\} \left(\widetilde{V}_{ik} - \hat{\mu}_{\widetilde{Y}}(T_{ik})\right) \\ &= \{[U_{ij} - \mu_X(T_{ij})] + [\mu_X(T_{ij}) - \hat{\mu}_X(T_{ij})]\} \\ &\quad \times \left([\widetilde{Y}(T_{ik}) - \mu_{\widetilde{Y}}(T_{ik})] + [\widetilde{V}_{ik} - \widetilde{Y}(T_{ik})] + [\mu_{\widetilde{Y}}(T_{ik}) - \hat{\mu}_{\widetilde{Y}}(T_{ik})]\right). \end{aligned}$$

Because $\max_{i,k}\left|\widetilde{V}_{ik} - \widetilde{Y}(T_{ik})\right| = O_p\left(\frac{1}{\sqrt{n}b_X} + \frac{1}{\sqrt{n}b_Y} + \frac{1}{\sqrt{n}h_X^2} + \frac{1}{\sqrt{n}h_Y^2}\right)$ by (S5.2) and

$$\sup_{t\in\mathcal{T}} \left|\mu_{\widetilde{Y}}(t) - \hat{\mu}_{\widetilde{Y}}(t)\right| = O_p\left(\frac{1}{\sqrt{n}b_Y}\right)$$

by Theorem 1 of Yao et al. (2005a), the local linear estimator, $\hat{r}_{X\widetilde{Y}}(s,t)$, of $r_{X\widetilde{Y}}(s,t)$ obtained from $R_{i,X\widetilde{Y}}(T_{ij}, T_{ik})$ is asymptotically equivalent to that obtained from $\tilde{R}_{i,X\widetilde{Y}}(T_{ij}, T_{ik}) = [U_{ij} - \mu_X(T_{ij})]\,[\widetilde{Y}(T_{ik}) - \mu_{\widetilde{Y}}(T_{ik})]$, denoted by $\tilde{r}_{X\widetilde{Y}}(s,t)$. So, by Lemma 2 and Theorem 1 of Yao et al. (2005a), we have

$$\sup_{s,t\mathcal{T}} \left|\hat{r}_{X\widetilde{Y}}(s,t) - r_{X\widetilde{Y}}(s,t)\right| = O_p\left(\frac{1}{\sqrt{n}b_X} + \frac{1}{\sqrt{n}b_Y} + \frac{1}{\sqrt{n}h_X^2} + \frac{1}{\sqrt{n}h_Y^2} + \frac{1}{\sqrt{n}h_1h_2}\right). \quad \text{(S5.3)}$$

From (S5.1) and (S5.3), we obtain

$$
\begin{aligned}
& |\hat{\beta}_\tau^{(1)}(t) - \beta_\tau^*(t)| \\
\leq & \left| \frac{\hat{r}_{X\widetilde{Y}}(t,t)}{\hat{r}_{XX}(t,t)} - \frac{r_{X\widetilde{Y}}(t,t)}{r_{XX}(t,t)} \right| + O\left(E\|X\|^3 \left|\beta_\tau^{(0)}(t) - \beta_\tau^*(t)\right|^2 \right) \\
= & \left| \frac{\left[\hat{r}_{X\widetilde{Y}}(t,t) - r_{X\widetilde{Y}}(t,t)\right] r_{XX}(t,t) + r_{X\widetilde{Y}}(t,t)\left[r_{XX}(t,t) - \hat{r}_{XX}(t,t)\right]}{r_{XX}(t,t)\hat{r}_{XX}(t,t)} \right| + O(a_n^2) \\
= & O_p\left(\frac{1}{\sqrt{n}b_X} + \frac{1}{\sqrt{n}b_Y} + \frac{1}{\sqrt{n}h_X^2} + \frac{1}{\sqrt{n}h_Y^2} + \frac{1}{\sqrt{n}h_1h_2} + a_n^2 \right).
\end{aligned}
$$

Further, we get $|\hat{\alpha}_\tau^{(1)}(t) - \alpha_\tau^*(t)|$ has the same order with $|\hat{\beta}_\tau^{(1)}(t) - \beta_\tau^*(t)|$. Thus, we complete the proof of Theorem 3.1.

Proof of Theorem 3.2 Let $b_n = O_p\left\{ \frac{1}{\sqrt{n}} \left[\frac{1}{b_X} + \frac{1}{b_Y} + \frac{1}{h_X^2} + \frac{1}{h_Y^2} + \frac{1}{h_1h_2} \right] \right\}$. From the proof of Theorem 3.1, we have

$$
|\hat{\beta}_\tau^{(1)}(t) - \beta_\tau^*(t)| = b_n + Ca_n|\hat{\beta}_\tau^{(0)}(t) - \beta_\tau^*(t)|.
$$

By the iteration algorithm, one gets

$$
\begin{aligned}
|\hat{\beta}_\tau^{(k)}(t) - \beta_\tau^*(t)| &= b_n[1 + Ca_n + \cdots + (Ca_n)^{k-1}] + (Ca_n)^k|\hat{\beta}_\tau^{(0)}(t) - \beta_\tau^*(t)| \\
&= \frac{b_n(1-(Ca_n)^k)}{1 - Ca_n} + C^k a_n^{k+1} \\
&= O_p\left(b_n + a_n^{k+1}\right).
\end{aligned}
$$

It completes the proof.

Proof of Theorem 3.3 For fixed M,

$$
\widetilde{Q}_{Y^*|X^*}^M(t;\tau) = \alpha_\tau(t) + \beta_\tau(t)\left(\mu_X(t) + \sum_{m=1}^M \tilde{\zeta}_m^* \phi_m(t) \right).
$$

Call that $\widetilde{Q}_{Y^*|X^*}(t;\tau) = \alpha_\tau(t) + \beta_\tau(t)\left(\mu_X(t) + \sum_{m=1}^{\infty} \tilde{\zeta}_m^* \phi_m(t)\right)$ and

$$\widehat{Q}^{\mathrm{K},M}_{Y^*|X^*}(t;\tau) = \hat{\alpha}^{\mathrm{K}}_\tau(t) + \hat{\beta}^{\mathrm{K}}_\tau(t)\left(\hat{\mu}_X(t) + \sum_{m=1}^{M} \hat{\zeta}_m^* \hat{\phi}_m(t)\right)$$

with $\tilde{\zeta}_m^*$ and $\hat{\zeta}_m^*$ defined in (3.11) and (3.12), respectively. Note that

$$\begin{aligned} \left|\widehat{Q}^{\mathrm{K},M}_{Y^*|X^*}(t;\tau) - \widetilde{Q}_{Y^*|X^*}(t;\tau)\right| &\leq \left|\widehat{Q}^{\mathrm{K},M}_{Y^*|X^*}(t;\tau) - \widetilde{Q}^{M}_{Y^*|X^*}(t;\tau)\right| \\ &+ \left|\widetilde{Q}^{M}_{Y^*|X^*}(t;\tau) - \widetilde{Q}_{Y^*|X^*}(t;\tau)\right|. \end{aligned} \tag{S5.4}$$

From Theorem 3.2, $\sup_{t\in\mathcal{T}}\left|\hat{\alpha}^{\mathrm{K}}_\tau(t) - \alpha_\tau(t)\right| = o_p(1)$ and $\sup_{t\in\mathcal{T}}\left|\hat{\beta}^{\mathrm{K}}_\tau(t) - \beta_\tau(t)\right| = o_p(1)$ for enough large K. By Theorem 1 of Yao et al. (2005a), Lemma A.1 and (B5), one gets $\sup_{t\in\mathcal{T}}|\hat{\mu}_X(t) - \mu_X(t)| = o_p(1)$ and $\left|\hat{\zeta}_m^* - \tilde{\zeta}_m^*\right| = o_p(1)$ as $n \to \infty$. Then by Slutsky's Theorem, we have

$$\left|\widehat{Q}^{\mathrm{K},M}_{Y^*|X^*}(t;\tau) - \widetilde{Q}^{M}_{Y^*|X^*}(t;\tau)\right| = o_p(1)$$

as $n \to \infty$ and sufficiently large K. On the other hand, it follows from Lemma 3 of Yao et al. (2005b) that $\widetilde{Q}^{M}_{Y^*|X^*}(t;\tau) \xrightarrow{p} \widetilde{Q}_{Y^*|X^*}(t;\tau)$. Therefore, Combining them with (S5.4), we complete the proof of Theorem 3.3.

Proof of Theorem 3.4 For a fixed $M \geq 1$, under the Gaussian assumption and conditional on N^* and $\boldsymbol{T}^*$, it is shown in Subsection 3.1.3 that $\tilde{\boldsymbol{\zeta}}_M^* - \boldsymbol{\zeta}_M^* \sim N(0, \boldsymbol{\Omega}_M)$. It then follows that

$$\widetilde{Q}^{M}_{Y^*|X^*}(t;\tau) - Q^{M}_{Y^*|X^*}(t;\tau) \xrightarrow{D} Z^M_\tau \sim \mathcal{N}(0, \omega^M_\tau(t)), \tag{S5.5}$$

where $Q^M_{Y^*|X^*}(t;\tau) = \alpha_\tau(t) + \beta_\tau(t)\left(\mu_X(t) + \sum_{m=1}^M \zeta^*_m \phi_m(t)\right)$. Note that

$$\begin{aligned}\widehat{Q}^{\mathrm{K},M}_{Y^*|X^*}(t;\tau) - Q_{Y^*|X^*}(t;\tau) &= \left(\widehat{Q}^{\mathrm{K},M}_{Y^*|X^*}(t;\tau) - \widetilde{Q}^{M}_{Y^*|X^*}(t;\tau)\right) \\ &+ \left(\widetilde{Q}^{M}_{Y^*|X^*}(t;\tau) - Q^{M}_{Y^*|X^*}(t;\tau)\right) \\ &+ \left(Q^{M}_{Y^*|X^*}(t;\tau) - Q_{Y^*|X^*}(t;\tau)\right). \qquad \text{(S5.6)}\end{aligned}$$

From the proof of the 1st term in (S5.4), we have the 1st term in (S5.6) for sufficiently large K and a fixed M,

$$\lim_{n\to\infty} \sup_{t\in\mathcal{T}} \left|\widehat{Q}^{\mathrm{K},M}_{Y^*|X^*}(t;\tau) - \widetilde{Q}^{M}_{Y^*|X^*}(t;\tau)\right| = o_p(1).$$

From Theorem 3.1 and Theorems 1 and 2 in Yao et al. (2005a), one gets $\hat{\omega}^{\mathrm{K},M}_\tau(t) \xrightarrow{p} \omega^M_\tau(t)$ as $n \to \infty$ and sufficient large K; and then by Assumption (A8)(i), we have $\lim_{M\to\infty} \lim_{n\to\infty} \hat{\omega}^{\mathrm{K},M}_\tau(t) = \omega_\tau(t)$ in probability for K enough large. Thus, letting $M \to \infty$ lead (S5.5), i.e. the 2nd term of (S5.6), to

$$\widetilde{Q}^{M}_{Y^*|X^*}(t;\tau) - Q^{M}_{Y^*|X^*}(t;\tau) \xrightarrow{D} Z^M_\tau \xrightarrow{D} Z_\tau \sim \mathcal{N}(0, \omega_\tau(t)).$$

For the 3rd term of (S5.6), by the Karhunen-Loéve theorem,

$$\widetilde{Q}^{M}_{Y^*|X^*}(t;\tau) - Q_{Y^*|X^*}(t;\tau) \xrightarrow{p} 0$$

as $M \to \infty$. Therefore, together with (S5.6), Theorem 3.4 follows by Slutsky's Theorem.

S5.2 For functional linear QR model

In the subsection, we give the proofs of Theorems 3.5-3.7.

Proof of Theorem 3.5 First, we consider the estimation of the 1st iterative algorithm. Call that $\hat{\beta}^{(1)}_\tau(s,t) = \sum_{m=1}^M \sum_{k=1}^K \frac{\hat{\tilde{\sigma}}_{mk}}{\hat{\rho}_m} \hat{\phi}_m(s)\hat{\varphi}_k(t)$, $\hat{\alpha}^{(1)}_\tau(t) = \hat{\mu}_{\widetilde{Y}}(t) - \int_{\mathcal{S}} \hat{\beta}^{(1)}_\tau(s,t)\hat{\mu}_X(s)ds$ and

$\beta_\tau^{(1)}(s,t) = \sum_{m=1}^\infty \sum_{k=1}^\infty \frac{\tilde\sigma_{mk}}{\rho_m}\phi_m(s)\varphi_k(t)$. By Theorem 2.1, we have

$$\begin{aligned}
& \left|\hat\beta_\tau^{(1)}(s,t) - \beta_\tau^*(s,t)\right| \\
\le\ & \left|\hat\beta_\tau^{(1)}(s,t) - \beta_\tau^{(1)}(s,t)\right| + \left|\beta_\tau^{(1)}(s,t) - \beta_\tau^*(s,t)\right| \\
\le\ & \left|\sum_{m=1}^{\tilde M}\sum_{k=1}^{\tilde K}\left[\frac{\hat{\tilde\sigma}_{mk}}{\hat\rho_m}\hat\phi_m(s)\hat\varphi_k(t) - \frac{\tilde\sigma_{mk}}{\rho_m}\phi_m(s)\varphi_k(t)\right]\right| + \left|\sum_{m=\tilde M+1}^{\infty}\sum_{k=\tilde K+1}^{\infty}\frac{\tilde\sigma_{mk}}{\rho_m}\phi_m(s)\varphi_k(t)\right| \\
& + O\left(E\|X\|^3\left\|\hat\beta_\tau^{(0)} - \beta_\tau^*\right\|^2\right) \\
=\ & Q_1(n) + Q_2(n) + O_p(a_n^2). \qquad \text{(S5.7)}
\end{aligned}$$

Since $\beta_\tau^{(1)}(s,t) = \arg\min_{\beta_\tau\in L_2(\mathcal{S}\times\mathcal{T})} E\|\widetilde{Y}^c - \mathcal{L}_{X^c}\beta_\tau\|^2$, we can write it as functional linear regression model

$$\widetilde{Y}^c(t) = \int_{\mathcal{S}} \beta_\tau(s,t)X^c(s)ds + e(t),$$

where e is mean zero random error, and independent of X. Under Condition 2, we have (S1.10), and $\sum_{m=1}^\infty \sum_{k=1}^\infty \frac{\tilde\sigma_{mk}}{\rho_m}\phi_m(s)\varphi_k(t)$ absolutely converges for all $s\in\mathcal{S}$ and $t\in\mathcal{T}$. Therefore,

$\sum_{m=1}^M \sum_{k=1}^K \frac{\tilde\sigma_{mk}}{\rho_m}\phi_m(s)\varphi_k(t)$ absolutely convergences to $\beta_\tau^{(1)}(s,t)$ for all $s\in\mathcal{S}$ and $t\in\mathcal{T}$ as $M,K\to\infty$. One has $Q_2(n)\to 0$ as $M,K\to\infty$. Next, we consider $Q_1(n)$.

Under Condition (A6), by Theorems 2-3 of Yao et al. (2005a), we have $\lim_{M\to\infty}\lim_{n\to\infty}\hat U_i^M(s) = X_i(s) + O_p\left\{1/(\sqrt{n}b_X) + 1/(\sqrt{n}h_X^2)\right\}$ and $\lim_{K\to\infty}\lim_{n\to\infty}\hat V_i^K(t) = Y_i(t) + O_p\left\{1/(\sqrt{n}b_Y) + 1/(\sqrt{n}h_Y^2)\right\}$ for all $s\in\mathcal{S}$ and $t\in\mathcal{T}$. Thus,

$$\lim_{M,K\to\infty}\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n \hat\omega_i(t) = \omega(t) + O_p\left(\frac{1}{\sqrt{n}b_X} + \frac{1}{\sqrt{n}b_Y} + \frac{1}{\sqrt{n}h_X^2} + \frac{1}{\sqrt{n}h_Y^2}\right)$$

for all $t \in \mathcal{T}$, where $\omega(t) = \mathbb{I}\left[Y(t) - \hat{\alpha}_\tau^{(0)}(t) - \int_\mathcal{S} \hat{\beta}_\tau^{(0)}(s,t)X(s)ds \le 0\right] - \tau$, and

$$\lim_{M\to\infty}\lim_{n\to\infty} \hat{Q}^{(0)}_{Y|X}(t;\tau) = Q^{(0)}_{Y|X}(t;\tau) + O_p\left(\frac{1}{\sqrt{n}b_X} + \frac{1}{\sqrt{n}h_X^2}\right),$$

where $Q^{(0)}_{Y|X}(t;\tau) = \hat{\alpha}_\tau^{(0)}(t) + \int_\mathcal{S} \hat{\beta}_\tau^{(0)}(s,t)X_i(s)ds$. Further,

$$\lim_{M\to\infty}\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \hat{U}_i^M(s) = X(s) + O_p\left\{1/(\sqrt{n}b_X) + 1/(\sqrt{n}h_X^2)\right\}.$$

Similar to the proofs of Theorem 3.1, $|\hat{f}_t(0) - f_t(0)| = O_p\left(\sqrt{\log n/(nh)} + a_n + h\right)$. So, for all $t \in \mathcal{T}$, we have

$$\lim_{M,K\to\infty}\lim_{n\to\infty} \hat{\mu}_{\widetilde{Y}}(t) = \mu_{\widetilde{Y}}(t) + O_p\left(\frac{1}{\sqrt{n}b_X} + \frac{1}{\sqrt{n}b_Y} + \frac{1}{\sqrt{n}h_X^2} + \frac{1}{\sqrt{n}h_Y^2}\right). \tag{S5.8}$$

In the local linear estimator for the covariance $r_{\tilde{Y}\tilde{Y}}$, we use the raw observations, $R_{i,\widetilde{Y}\widetilde{Y}}(T_{ij},T_{ik}) = \left(\widetilde{V}_{ij} - \hat{\mu}_{\widetilde{Y}}(T_{ij})\right)\left(\widetilde{V}_{ik} - \hat{\mu}_{\widetilde{Y}}(T_{ik})\right)$, instead of $\tilde{R}_{i,\widetilde{Y}\widetilde{Y}}(T_{ij},T_{ik}) = \left(\widetilde{Y}(T_{ij}) - \mu_{\widetilde{Y}}(T_{ij})\right)\left(\widetilde{Y}(T_{ik}) - \mu_{\widetilde{Y}}(T_{ik})\right)$. Note that

$$\begin{aligned}
& R_{i,\widetilde{Y}\widetilde{Y}}(T_{ij},T_{ik}) \\
= & \left\{\left(\widetilde{Y}(T_{ij}) - \mu_{\widetilde{Y}}(T_{ij})\right) + \left(\widetilde{V}_{ij} - \widetilde{Y}(T_{ij})\right) + \left(\hat{\mu}_{\widetilde{Y}}(T_{ij}) - \mu_{\widetilde{Y}}(T_{ij})\right)\right\} \\
& \times\left\{\left(\widetilde{Y}(T_{ik}) - \mu_{\widetilde{Y}}(T_{ik})\right) + \left(\widetilde{V}_{ik} - \widetilde{Y}(T_{ik})\right) + \left(\hat{\mu}_{\widetilde{Y}}(T_{ik}) - \mu_{\widetilde{Y}}(T_{ik})\right)\right\} \\
= & \left(\widetilde{Y}(T_{ij}) - \mu_{\widetilde{Y}}(T_{ij})\right)\left(\widetilde{Y}(T_{ik}) - \mu_{\widetilde{Y}}(T_{ik})\right) \\
& + \left(\widetilde{Y}(T_{ij}) - \mu_{\widetilde{Y}}(T_{ij})\right)\left[\left(\widetilde{V}_{ik} - \widetilde{Y}(T_{ik})\right) + \left(\hat{\mu}_{\widetilde{Y}}(T_{ik}) - \mu_{\widetilde{Y}}(T_{ik})\right)\right] \\
& + \left(\widetilde{Y}(T_{ik}) - \mu_{\widetilde{Y}}(T_{ik})\right)\left[\left(\widetilde{V}_{ij} - \widetilde{Y}(T_{ij})\right) + \left(\hat{\mu}_{\widetilde{Y}}(T_{ij}) - \mu_{\widetilde{Y}}(T_{ij})\right)\right] \\
& + \left[\left(\widetilde{V}_{ij} - \widetilde{Y}(T_{ij})\right) + \left(\hat{\mu}_{\widetilde{Y}}(T_{ij}) - \mu_{\widetilde{Y}}(T_{ij})\right)\right] \\
& \times\left[\left(\widetilde{V}_{ik} - \widetilde{Y}(T_{ik})\right) + \left(\hat{\mu}_{\widetilde{Y}}(T_{ik}) - \mu_{\widetilde{Y}}(T_{ik})\right)\right].
\end{aligned}$$

Because $\max_{i,k}\left|\widetilde{V}_{ik}-\widetilde{Y}(T_{ik})\right| = O_p\left(\frac{1}{\sqrt{n}b_X}+\frac{1}{\sqrt{n}b_Y}+\frac{1}{\sqrt{n}h_X^2}+\frac{1}{\sqrt{n}h_Y^2}\right)$ and (S5.8), the local linear estimator, $\hat{r}_{\widetilde{Y}\widetilde{Y}}(s,t)$, of $r_{\widetilde{Y}\widetilde{Y}}(s,t)$ obtained from $R_{i,\widetilde{Y}\widetilde{Y}}(T_{ij},T_{ik})$ is asymptotically equivalent to that obtained from $\tilde{R}_{i,\widetilde{Y}\widetilde{Y}}(T_{ij},T_{ik})$, denoted by $\tilde{r}_{\widetilde{Y}\widetilde{Y}}(s,t)$. So, by Lemma 2 and Theorem 1 of Yao et al. (2005a), we have

$$\sup_{s,t\mathcal{T}}\left|\hat{r}_{\widetilde{Y}\widetilde{Y}}(s,t)-r_{\widetilde{Y}\widetilde{Y}}(s,t)\right| = O_p\left(\frac{1}{\sqrt{n}b_X}+\frac{1}{\sqrt{n}b_Y}+\frac{1}{\sqrt{n}h_X^2}+\frac{1}{\sqrt{n}h_Y^2}\right). \tag{S5.9}$$

By Theorem 2 of Yao et al. (2005a), one gets

$$\sup_{t\in\mathcal{T}}|\hat{\varphi}_k(t)-\varphi_k(t)| = O_p\left(\frac{1}{\sqrt{n}b_X}+\frac{1}{\sqrt{n}b_Y}+\frac{1}{\sqrt{n}h_X^2}+\frac{1}{\sqrt{n}h_Y^2}\right). \tag{S5.10}$$

Similar to the proof of (S5.9), we have

$$\sup_{s,t\mathcal{T}}\left|\hat{r}_{X\widetilde{Y}}(s,t)-r_{X\widetilde{Y}}(s,t)\right| = O_p\left(\frac{1}{\sqrt{n}b_X}+\frac{1}{\sqrt{n}b_Y}+\frac{1}{\sqrt{n}h_X^2}+\frac{1}{\sqrt{n}h_Y^2}+\frac{1}{\sqrt{n}h_1h_2}\right). \tag{S5.11}$$

In addition, by Theorem 2 of Yao et al. (2005a), we have

$$\sup_{s\in\mathcal{S}}\left|\hat{\phi}_m(t)-\phi_m(t)\right| = O_p\left(\frac{1}{\sqrt{n}h_X^2}\right). \tag{S5.12}$$

As a consequence of (S5.10), (S5.12) and (S5.11), one obtains

$$|\hat{\tilde{\sigma}}_{mk}-\tilde{\sigma}_{mk}| = O_p\left(\frac{1}{\sqrt{n}b_X}+\frac{1}{\sqrt{n}b_Y}+\frac{1}{\sqrt{n}h_X^2}+\frac{1}{\sqrt{n}h_Y^2}+\frac{1}{\sqrt{n}h_1h_2}\right). \tag{S5.13}$$

By Assumptions (B1)-(B2), results (S5.10), (S5.12) and (S5.11), and expression (S5.7), we have

$$\sup_{(s,t)\in\mathcal{S}\times\mathcal{T}}\left|\hat{\beta}_\tau^{(1)}(s,t)-\beta_\tau^*(s,t)\right| = O_p\left(\delta_{1n}+\delta_{2n}+a_n^2\right).$$

We can show $\sup_{(s,t)\in\mathcal{S}\times\mathcal{T}}\left|\hat{\alpha}_\tau^{(1)}(s,t)-\alpha_\tau^*(s,t)\right| = O_p\left(\delta_{1n}+\delta_{2n}+a_n^2\right)$ with similar proofs of $\hat{\beta}_\tau^{(1)}(s,t)$. By the k iterations, we obtain the results of theorem 3.5.

Proof of Theorem 3.6 Recall that

$$\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau) = \hat{\alpha}^{\mathrm{K}}_{\tau}(t) + \sum_{m=1}^{M}\sum_{k=1}^{K} \frac{\hat{\tilde{\sigma}}^{\mathrm{K}}_{mk}}{\hat{\rho}_k}\hat{\zeta}^*_m \hat{\varphi}^{\mathrm{K}}_k(t),$$

$$\widetilde{Q}_{Y^*|X^*}(t;\tau) = \alpha_{\tau}(t) + \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} \frac{\tilde{\sigma}_{mk}}{\rho_m}\tilde{\zeta}^*_m \varphi_k(t),$$

where $\tilde{\zeta}^*_m$ and $\hat{\zeta}^*_m$ are defined in (S1.15) and (S1.16), respectively. For given S^* and L^*, define

$$\widetilde{Q}^{M,K}_{Y^*|X^*}(t;\tau) = \alpha_{\tau}(t) + \sum_{m=1}^{M}\sum_{k=1}^{K} \frac{\tilde{\sigma}_{mk}}{\rho_m}\tilde{\zeta}^*_m \varphi_k(t).$$

Now, let's continue to prove after (S5.9)-(S5.13). First,

$$\begin{aligned}
& \widetilde{Y}(t) - \widetilde{Y}^*(t) \\
= & \left(Q^{(0)}_{Y|X}(t;\tau) - Q^*_{Y|X}(t;\tau)\right) \\
& - \left(\int_{\mathcal{T}} f_t(0)dt\right)^{-1} \int_{\mathcal{T}} \{\mathbb{I}[Y(t) - Q^{(0)}_{Y|X}(t;\tau) \le 0] - \mathbb{I}[Y(t) - Q^*_{Y|X}(t;\tau) \le 0]\}dt \\
= & O_p(a_n)
\end{aligned}$$

uniformly for $t \in \mathcal{T}$, since $\sup_{t\in\mathcal{T}}\left|Q^{(0)}_{Y|X}(t;\tau) - Q^*_{Y|X}(t;\tau)\right| = O_p(a_n)$ and $\mathbb{I}[Y(t) - Q^{(0)}_{Y|X}(t;\tau) \le 0] - \mathbb{I}[Y(t) - Q^*_{Y|X}(t;\tau) \le 0] = O_p(a_n)$ which is obtained by

$$E\left(\mathbb{I}[Y(t) - Q^{(0)}_{Y|X}(t;\tau) \le 0] - \mathbb{I}[Y(t) - Q^*_{Y|X}(t;\tau) \le 0]\right) = O(a_n),$$

$$E\left(\mathbb{I}[Y(t) - Q^{(0)}_{Y|X}(t;\tau) \le 0] - \mathbb{I}[Y(t) - Q^*_{Y|X}(t;\tau) \le 0]\right)^2 = O(a_n(1-a_n)) = O(a_n).$$

Thus, we also have $\sup_{t\in\mathcal{T}}|\mu_{\widetilde{Y}} - \mu_{\widetilde{Y}^*}| = O(a_n)$. Therefore, one gets

$$\begin{aligned}
& \sup_{s,t\mathcal{T}} \left| r_{\widetilde{Y}\widetilde{Y}}(s,t) - r_{\widetilde{Y}^*\widetilde{Y}^*}(s,t)\right| \\
= & \sup_{s,t\mathcal{T}} \left| E[\widetilde{Y}(s) - \mu_{\widetilde{Y}}(s)][\widetilde{Y}(t) - \mu_{\widetilde{Y}}(t)] - E[\widetilde{Y}^*(s) - \mu_{\widetilde{Y}^*}(s)][\widetilde{Y}^*(t) - \mu_{\widetilde{Y}^*}(t)]\right| \\
= & O(a_n^2). \qquad \text{(S5.14)}
\end{aligned}$$

By (S5.14) and (S5.9), we have

$$\sup_{s,t\mathcal{T}}\left|\hat{r}_{\widetilde{Y}\widetilde{Y}}(s,t)-r_{\widetilde{Y}^*\widetilde{Y}^*}(s,t)\right|=O_p\left(\frac{1}{\sqrt{n}b_X}+\frac{1}{\sqrt{n}b_Y}+\frac{1}{\sqrt{n}h_X^2}+\frac{1}{\sqrt{n}h_Y^2}+a_n^2\right). \tag{S5.15}$$

Similarly, by (S5.11), one gets

$$\begin{aligned}&\sup_{s,t\mathcal{T}}\left|\hat{r}_{X\widetilde{Y}}(s,t)-r_{X\widetilde{Y}^*}(s,t)\right|\\ =\;& O_p\left(\frac{1}{\sqrt{n}b_X}+\frac{1}{\sqrt{n}b_Y}+\frac{1}{\sqrt{n}h_X^2}+\frac{1}{\sqrt{n}h_Y^2}+\frac{1}{\sqrt{n}h_1h_2}+a_n^2\right).\end{aligned} \tag{S5.16}$$

From (S5.15), applying Theorem 2 of Yao et al. (2005a), we obtain

$$\sup_{t\in\mathcal{T}}|\hat{\varphi}_k(t)-\varphi_k(t)|=O_p\left(\frac{1}{\sqrt{n}b_X}+\frac{1}{\sqrt{n}b_Y}+\frac{1}{\sqrt{n}h_X^2}+\frac{1}{\sqrt{n}h_Y^2}+a_n^2\right).$$

As a consequence of (S5.17), (S5.12) and (S5.16),

$$\max_{mk}|\hat{\tilde{\sigma}}_{mk}-\tilde{\sigma}_{mk}|=O_p\left(\frac{1}{\sqrt{n}b_X}+\frac{1}{\sqrt{n}b_Y}+\frac{1}{\sqrt{n}h_X^2}+\frac{1}{\sqrt{n}h_Y^2}+\frac{1}{\sqrt{n}h_1h_2}+a_n^2\right).$$

By the K iterations, we have

$$\sup_{t\in\mathcal{T}}\left|\hat{\varphi}_k^{\mathrm{K}}(t)-\varphi_k(t)\right|=O_p\left(\delta_{1n}+\delta_{2n}+a_n^{\mathrm{K}+1}\right), \tag{S5.17}$$

$$\max_{mk}|\hat{\tilde{\sigma}}_{mk}^{\mathrm{K}}-\tilde{\sigma}_{mk}|=O_p\left(\delta_{1n}+\delta_{2n}+a_n^{\mathrm{K}+1}\right). \tag{S5.18}$$

Note that

$$\begin{aligned}\left|\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau)-\widetilde{Q}_{Y^*|X^*}(t;\tau)\right| \;\le\;& \left|\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau)-\widetilde{Q}^{M,K}_{Y^*|X^*}(t;\tau)\right|\\ &+\left|\widetilde{Q}^{M,K}_{Y^*|X^*}(t;\tau)-\widetilde{Q}_{Y^*|X^*}(t;\tau)\right|.\end{aligned} \tag{S5.19}$$

From Theorem 3.5,

$$\sup_{t\in\mathcal{T}} \left|\hat{\alpha}_\tau^{\mathrm{K}}(t) - \alpha_\tau(t)\right| = o_p(1)$$

and $\sup_{(s,t)\in\mathcal{S}\times\mathcal{T}} \left|\hat{\beta}_\tau^{\mathrm{K}}(s,t) - \beta_\tau(s,t)\right| = o_p(1)$ for enough large K. By Theorem 1 of Yao et al. (2005a), Lemma A.1 and (B5), one gets

$\sup_{t\in\mathcal{T}} |\hat{\mu}_X(t) - \mu_X(t)| = o_p(1)$ and $\left|\hat{\zeta}_m^* - \tilde{\zeta}_m^*\right| = o_p(1)$ as $n \to \infty$. Then under Assumptions (B1)-(B2), by (S5.18)-(S5.19) and Slutsky's Theorem, we have $\left|\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau) - \widetilde{Q}^{M,K}_{Y^*|X^*}(t;\tau)\right| = o_p(1)$ as $n \to \infty$ and sufficiently large K. On the other hand, it follows from Lemma 3 of Yao et al. (2005b) that $\widetilde{Q}^{M,K}_{Y^*|X^*}(t;\tau) \xrightarrow{p} \widetilde{Q}_{Y^*|X^*}(t;\tau)$. Therefore, Combining them with (S5.19), we complete the proof of Theorem 3.6.

Proof of Theorem 3.7 For a fixed $M, K \geq 1$, under the Gaussian assumption and conditional on N^* and $\boldsymbol{T}^*$, it is shown in Subsection 3.2.3 that $\tilde{\boldsymbol{\zeta}}_M^* - \boldsymbol{\zeta}_M^* \sim N(0, \boldsymbol{\Omega}_M)$. It then follows that

$$\widetilde{Q}^{M,K}_{Y^*|X^*}(t;\tau) - Q^{M,K}_{Y^*|X^*}(t;\tau) \xrightarrow{D} Z_\tau^{M,K} \sim \mathcal{N}(0, \omega_\tau^{M,K}(t)), \tag{S5.20}$$

where

$$Q^{M,K}_{Y^*|X^*}(t;\tau) = \alpha_\tau(t) + \sum_{m=1}^{M}\sum_{k=1}^{K} \frac{\tilde{\sigma}_{mk}}{\rho_m}\zeta_m^*\varphi_k(t),$$

. Note that

$$\begin{aligned}
\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau) - Q_{Y^*|X^*}(t;\tau) \;\; = \;\; & \left(\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau) - \widetilde{Q}^{M,K}_{Y^*|X^*}(t;\tau)\right) \\
& + \left(\widetilde{Q}^{M,K}_{Y^*|X^*}(t;\tau) - Q^{M,K}_{Y^*|X^*}(t;\tau)\right) \\
& + \left(Q^{M,K}_{Y^*|X^*}(t;\tau) - Q_{Y^*|X^*}(t;\tau)\right).
\end{aligned} \tag{S5.21}$$

From the proof of the 1st term in (S5.19), we have the 1st term in (S5.21) for sufficiently large K and a fixed M, K,

$$\lim_{n\to\infty}\sup_{t\in\mathcal{T}}\left|\widehat{Q}^{\mathrm{K},M,K}_{Y^*|X^*}(t;\tau)-\widetilde{Q}^{M,K}_{Y^*|X^*}(t;\tau)\right|=o_p(1).$$

From Theorem 3.1, (S5.17)-(S5.18), and Theorems 1 in Yao et al. (2005a), one gets $\hat{\omega}^{\mathrm{K},M,K}_{\tau}(t)\xrightarrow{p}\omega^{M,K}_{\tau}(t)$ as $n\to\infty$ and sufficient large K; and then by Assumption (A8)(ii), we have $\lim_{M,K\to\infty}\lim_{n\to\infty}\hat{\omega}^{\mathrm{K},M,K}_{\tau}(t)=\omega_\tau(t)$ in probability for K enough large. Thus, letting $M, K\to\infty$ lead (S5.20), i.e. the 2nd term of (S5.21), to

$$\widetilde{Q}^{M,K}_{Y^*|X^*}(t;\tau)-Q^{M,K}_{Y^*|X^*}(t;\tau)\xrightarrow{D}Z^{M,K}_{\tau}\xrightarrow{D}Z_\tau\sim\mathcal{N}(0,\omega_\tau(t)).$$

For the 3rd term of (S5.21), by the Karhunen-Loéve theorem,

$$\widetilde{Q}^{M,K}_{Y^*|X^*}(t;\tau)-Q_{Y^*|X^*}(t;\tau)\xrightarrow{p}0$$

as $M, K\to\infty$. Therefore, together with (S5.21), Theorem 3.7 follows by Slutsky's Theorem.

S5.3 For functional varying coefficient QR model with history index

In the subsection, we give the proofs of Theorems

Proof of Theorem S1.3.1 Let $b_n=\dfrac{1}{\sqrt{n}b_X}+\dfrac{1}{\sqrt{n}b_Y}+\dfrac{1}{\sqrt{n}h_X^2}+\dfrac{1}{\sqrt{n}h_Y^2}+\dfrac{1}{\sqrt{n}h_1h_2}$. To prove uniform consistency of $\hat{\gamma}_\tau^{(1)}(s)$, we first prove one of $\hat{\varrho}_\tau^{(1)}(s;t)$. Note that by Theorem 2.1,

$$\begin{aligned}
& \sup_{s\in[0,\Delta]} |\hat{\varrho}_\tau^{(1)}(s;t) - \varrho_\tau^*(s;t)| \\
\leq\ & \sup_{s\in[0,\Delta]} |\hat{\varrho}_\tau^{(1)}(s;t) - \varrho_\tau^{(1)}(s;t)| + \sup_{s\in[0,\Delta]} |\varrho_\tau^{(1)}(s;t) - \varrho_\tau^*(s;t)| \\
\leq\ & \sup_{s\in[0,\Delta]} \left| \sum_{m=1}^{M_t} \hat{\varrho}_{\tau,m}^{(1)}(t)\hat{\phi}_{tm}(s) - \sum_{m=1}^{M_t} \varrho_{\tau,m}^{(1)}(t)\phi_{tm}(s) \right| + O_p(a_n^2) \\
& + \sup_{s\in[0,\Delta]} \left| \sum_{m=1}^{M_t} \varrho_{\tau,m}^{(1)}(t)\phi_{tm}(s) - \sum_{m=1}^{\infty} \varrho_{\tau,m}^{(1)}(t)\phi_{tm}(s) \right| \\
=:\ & Q_1(n) + Q_2(n) + O_p(a_n^2)
\end{aligned}$$

for $t \in \mathcal{T}$. Similar to the proof of (S5.11) , we have

$$\sup_{s\in[0,\Delta],t\in\mathcal{T}} \left| \hat{r}_{X\widetilde{Y}}(t-s,t) - r_{X\widetilde{Y}}(t-s,t) \right| = O_p\left(b_n\right). \tag{S5.22}$$

By Theorem 2 of Yao et al. (2005a), one gets

$$\begin{aligned}
\sup_{s\in[0,\Delta],t\in\mathcal{T}} \left| \hat{\phi}_{tm}(s) - \phi_{tm}(s) \right| &= O_p\left(\frac{1}{\sqrt{n}h_X^2}\right), \\
\sup_{t\in\mathcal{T}} |\hat{\rho}_{\tau,m}(t) - \rho_{\tau,m}(t)| &= O_p\left(\frac{1}{\sqrt{n}h_X^2}\right).
\end{aligned} \tag{S5.23}$$

From (S5.22)-(S5.23), we have

$$\sup_{t\in\mathcal{T}} \left| \hat{\varrho}_{\tau,m}^{(1)}(t) - \varrho_{\tau,m}^{(1)}(t) \right| = O_p\left(b_n\right). \tag{S5.24}$$

Together with (S5.24), we have $Q_1(n) = O_p\left(b_n\right)$ uniformly.

Further, similar to the proof of (S5.14), we have

$$\sup_{s\in[0,\Delta],t\in\mathcal{T}} \left| r_{X\widetilde{Y}}(t-s,t) - r_{X\widetilde{Y}^*}(t-s,t) \right| = O_p\left(a_n^2\right),$$

thus, by Theorem 2 of Yao et al. (2005a), $\sup_{t\in\mathcal{T}}\left|\varrho^{(1)}_{\tau,m}(t)-\varrho^{*}_{\tau,m}(t)\right| = O_p\left(a_n^2\right)$. By Assumption (C1), one gets $Q_2(n)=\delta_{3n}+O_p\left(a_n^2\right)$. Thus, we obtain

$$\sup_{s\in[0,\Delta],t\in\mathcal{T}}|\hat{\varrho}^{(1)}_{\tau}(s;t)-\varrho^{*}_{\tau}(s;t)| = O_p(\delta_{3n}+b_n+a_n^2).$$

So, the rate of uniform consistency of $\hat{\varrho}^{(1)}_{\tau}(s;t)$ leads to the one of $\hat{\gamma}^{(1)}_{\tau}(s)$. The uniform consistency of $\hat{\alpha}^{(1)}_{\tau}(t)$ and $\hat{\beta}^{(1)}_{\tau}(t)$ follows analogously. By the k iteration, we complete the proof of Theorem S1.3.1.

Proof of Theorem S1.3.2 For a fixed $\mathcal{M}$, M_t, let

$$\begin{aligned}\widehat{Q}^{(\mathcal{M},M_t)}_{Y^*|X^*}(t;\tau) &= \hat{\alpha}_{\tau}(t)+\int_0^{\Delta}\hat{\varrho}(s;\tau)\hat{\mu}_X(t-s)ds \\ &\quad+\frac{\hat{\beta}_{\tau}(t)\sum_{m=1}^{M_t}\hat{\zeta}^{*}_{tm}\sum_{r=1}^{R}\sum_{m'=1}^{M_r}\hat{\varrho}_{\tau,m'}(t)\int_0^{\Delta}\hat{\phi}_{tm}(s)\hat{\phi}_{tm'}(s)ds}{\left[\int_0^{\Delta}\left(\sum_{r=1}^{R}\hat{\varrho}_{\tau}(s;t_r)\right)^2ds\right]^{1/2}}.\end{aligned}$$

Similar to the proof of Theorem 3.6, note that

$$\begin{aligned}\left|\widehat{Q}^{(\mathrm{K},\mathcal{M},M_t)}_{Y^*|X^*}(t;\tau)-\widetilde{Q}_{Y^*|X^*}(t;\tau)\right| &\le \left|\widehat{Q}^{(\mathrm{K},\mathcal{M},M_t)}_{Y^*|X^*}(t;\tau)-\widetilde{Q}^{(\mathcal{M},M_t)}_{Y^*|X^*}(t;\tau)\right| \\ &\quad+\left|\widetilde{Q}^{(\mathcal{M},M_t)}_{Y^*|X^*}(t;\tau)-\widetilde{Q}_{Y^*|X^*}(t;\tau)\right|. \qquad \text{(S5.25)}\end{aligned}$$

By similar to arguments as in the proof of Lemma 3 in Yao et al. (2005b), $\widetilde{Q}^{(\mathcal{M},M_t)}_{Y^*|X^*}(t;\tau)\xrightarrow{p}\widetilde{Q}_{Y^*|X^*}(t;\tau)$ as M_t, $M_1,\cdots,M_R\to\infty$ and $n\to\infty$. On the other hand, from Theorem S1.3.1, $\sup_{s\in[0,\Delta]}|\hat{\gamma}^{(k)}_{\tau}(s)-\gamma^{*}_{\tau}(s)|\to 0$, $\sup_{t\in\mathcal{T}}|\hat{\alpha}^{(k)}_{\tau}(t)-\alpha^{*}_{\tau}(t)|\to 0$ and $\sup_{t\in\mathcal{T}}|\hat{\beta}^{(0)}_{\tau}(t)-\beta^{*}_{\tau}(t)|\to 0$ in probability for enough large K. By Theorem 1 of Yao et al. (2005a), Lemma A.1 and (B5), one gets $\sup_{t\in\mathcal{T}}|\hat{\mu}_X(t)-\mu_X(t)| = o_p(1)$ and $\left|\hat{\zeta}^{*}_{m}-\tilde{\zeta}^{*}_{m}\right| = o_p(1)$ as $n\to\infty$. Then by Slutsky's Theorem, we have $\sup_{t\in\mathcal{T}}\left|\widehat{Q}^{(\mathrm{K},\mathcal{M},M_t)}_{Y^*|X^*}(t;\tau)-\widetilde{Q}^{(\mathcal{M},M_t)}_{Y^*|X^*}(t;\tau)\right|\to 0$ in probability for sufficiently large K and $n\to\infty$. Thus, it follows Theorem S1.3.2.

Proof of Theorem S1.3.3 The proof is similar to the proofs of Theorems 3.4 and 3.7. Note that the decomposition

$$\begin{aligned}\widehat{Q}^{\mathrm{K},\mathcal{M},M_t}_{Y^*|X^*}(t;\tau) - Q_{Y^*|X^*}(t;\tau) &= \left(\widehat{Q}^{\mathrm{K},\mathcal{M},M_t}_{Y^*|X^*}(t;\tau) - \widetilde{Q}^{\mathcal{M},M_t}_{Y^*|X^*}(t;\tau)\right) \\ &+ \left(\widetilde{Q}^{\mathcal{M},M_t}_{Y^*|X^*}(t;\tau) - Q^{\mathcal{M},M_t}_{Y^*|X^*}(t;\tau)\right) \\ &+ \left(Q^{\mathcal{M},M_t}_{Y^*|X^*}(t;\tau) - Q^{M_t}_{Y^*|X^*}(t;\tau)\right) \\ &+ \left(Q^{M_t}_{Y^*|X^*}(t;\tau) - Q_{Y^*|X^*}(t;\tau)\right) \end{aligned} \tag{S5.26}$$

where

$$\widetilde{Q}^{\mathcal{M},M_t}_{Y^*|X^*}(t;\tau) = \alpha_\tau(t) + \int_0^\Delta \varrho(s;\tau)\mu_X(t-s)ds + \left(\tilde{\boldsymbol{\zeta}}^{M_t}_{*,t}\right)^T \boldsymbol{\phi}_{t,\mathcal{M}},$$

$$Q^{\mathcal{M},M_t}_{Y^*|X^*}(t;\tau) = \alpha_\tau(t) + \int_0^\Delta \varrho(s;\tau)\mu_X(t-s)ds + \left(\boldsymbol{\zeta}^{M_t}_{*,t}\right)^T \boldsymbol{\phi}_{t,\mathcal{M}}.$$

For a fixed $(\mathcal{M}, M_t)$, under the Gaussian assumption and conditional on N^* and $\boldsymbol{T}^*$, it is shown in Subsection S1.3 that $\tilde{\boldsymbol{\zeta}}^{M_t}_{*,t} - \boldsymbol{\zeta}^{M_t}_{*,t} \sim N(0, \boldsymbol{\Omega}_{tM_t})$. It then follows that

$$\widetilde{Q}^{\mathcal{M},M_t}_{Y^*|X^*}(t;\tau) - Q^{\mathcal{M},M_t}_{Y^*|X^*}(t;\tau) \xrightarrow{D} Z^{\mathcal{M},M_t}_\tau \sim \mathcal{N}(0, \omega^{\mathcal{M},M_t}_\tau(t)). \tag{S5.27}$$

From the proof of the 1st term in (S5.25), we have the 1st term in (S5.26) for sufficiently large K and a fixed $(\mathcal{M}, M_t)$,

$$\lim_{n\to\infty}\sup_{t\in\mathcal{T}} \left|\widehat{Q}^{\mathrm{K},\mathcal{M},M}_{Y^*|X^*}(t;\tau) - \widetilde{Q}^{\mathcal{M},M}_{Y^*|X^*}(t;\tau)\right| = o_p(1).$$

When $M_1(n), \cdots, M_R(n) \to \infty$ as $n \to \infty$, one gets

$$\lim_{n\to\infty}\sup_{t\in\mathcal{T}} \left|Q^{\mathcal{M},M_t}_{Y^*|X^*}(t;\tau) - Q^{M_t}_{Y^*|X^*}(t;\tau)\right| = o_p(1).$$

From Theorem S1.3.1 and Theorems 1 and 2 in Yao et al. (2005a), one gets $\hat{\omega}_{\tau}^{\mathrm{K},\mathcal{M},M_t}(t) \xrightarrow{p} \omega_{\tau}^{\mathcal{M},M_t}(t)$ as $n \to \infty$ and sufficient large K; and then by Assumption (A8)(iii), we have $\lim_{M_t,M_1,\cdots,M_R\to\infty} \lim_{n\to\infty} \hat{\omega}_{\tau}^{\mathrm{K},\mathcal{M},M_t}(t) = \omega_{\tau}(t)$ in probability for K enough large. Thus, letting $M \to \infty$ lead (S5.27), i.e. the 2nd term of (S5.26), to

$$\widetilde{Q}_{Y^*|X^*}^{\mathcal{M},M_t}(t;\tau) - Q_{Y^*|X^*}^{\mathcal{M},M_t}(t;\tau) \xrightarrow{D} Z_{\tau}^{\mathcal{M},M_t} \xrightarrow{D} Z_{\tau} \sim \mathcal{N}(0, \omega_{\tau}(t)).$$

For the 4th term of (S5.26), by the Karhunen-Loéve theorem,

$$Q_{Y^*|X^*}^{M_t}(t;\tau) - Q_{Y^*|X^*}(t;\tau) \to 0$$

as $M_t \to \infty$. Therefore, together with (S5.26), Theorem S1.3.3 follows by Slutsky's Theorem.

S6 Boxplots of MSEs

Boxplots of MSEs for example 2 are presented in Figures S1 and S2.

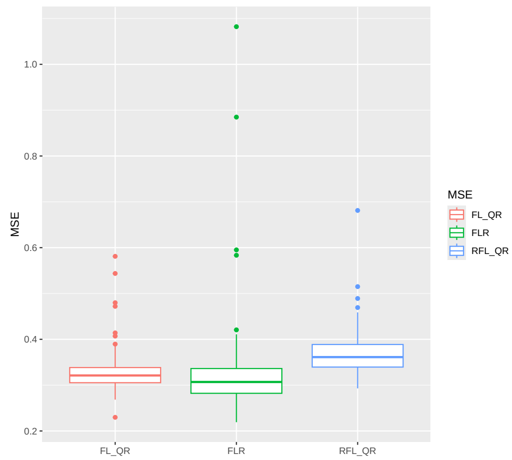
S7 Estimated pointwise coefficient of determination

For evaluating performance of our FL-QR and FLR, we also give the curve of estimated pointwise functional coefficients of determination $R_Q^2(t)$ based on FL-QR with the definition

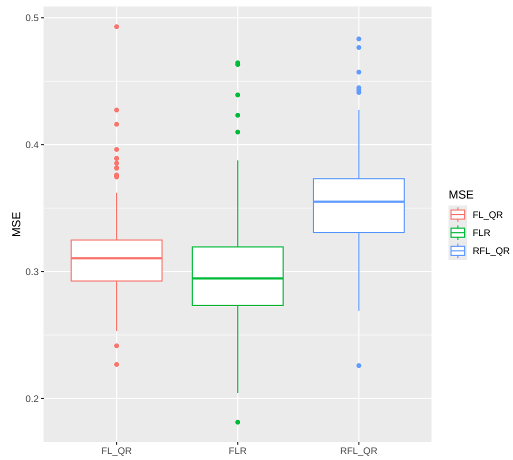
$$R_Q^2(t) = \frac{\mathrm{Var}[Q_{Y|X}(Y(t;\tau) \mid X)]}{\mathrm{Var}[Y(t)]},$$

and compare with that of determination $R_M^2(t)$ based on FLR with the similar definition

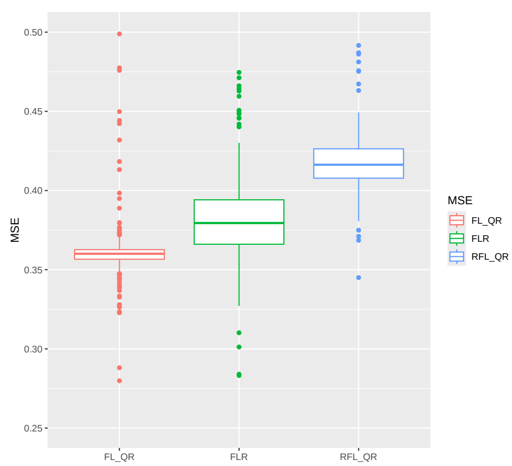
$$R_M^2(t) = \frac{\mathrm{Var}(E[Y(t) \mid X])}{\mathrm{Var}(Y(t))}.$$



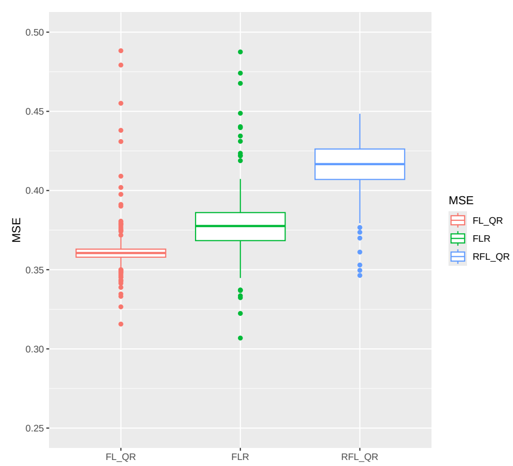
(a) $n = 50$



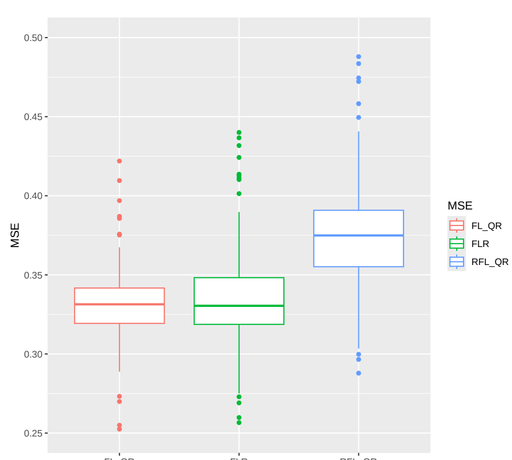
(b) $n = 100$



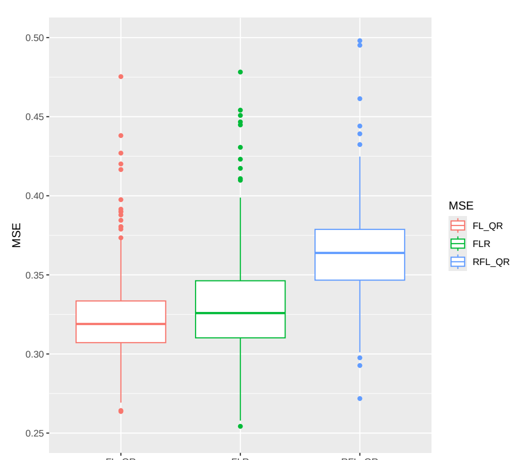
(c) $n = 50$



(d) $n = 100$

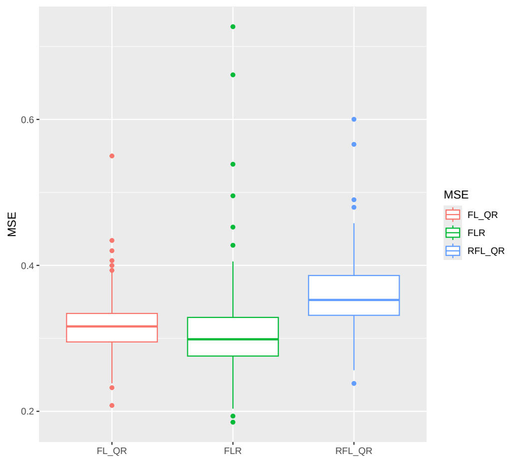


(e) $n = 50$

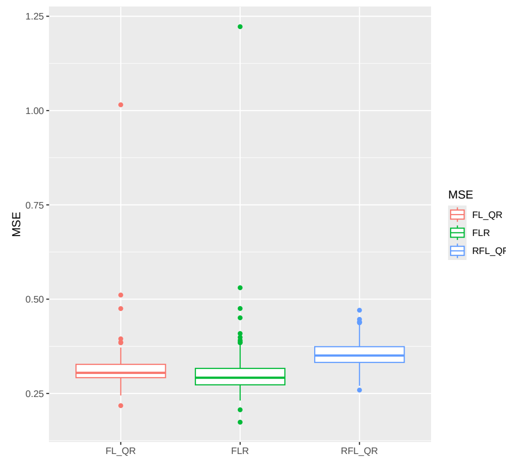


(f) $n = 100$

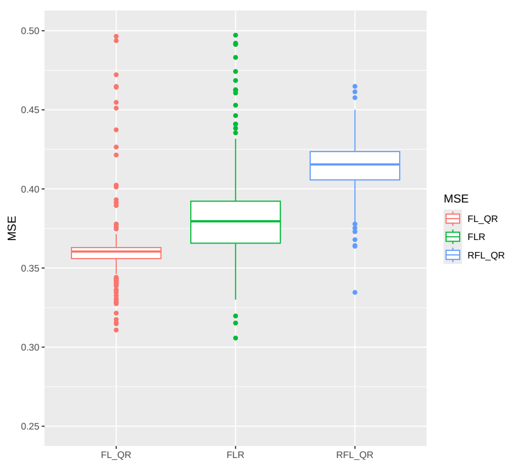
Figure S1: Boxplots of MSE when $\tau = 0.25$ for Example 2. blue: FLR method, : FL-QR method, green: RFL-QR method. The first row corresponds to the normal errors, the second row to the Cauchy errors, and the third row to the Chi-square errors.



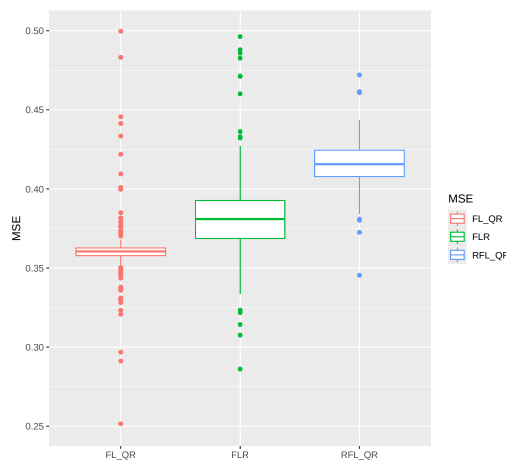
(a) $n = 50$



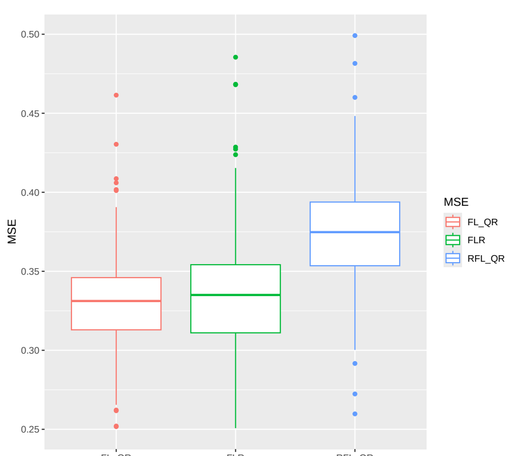
(b) $n = 100$



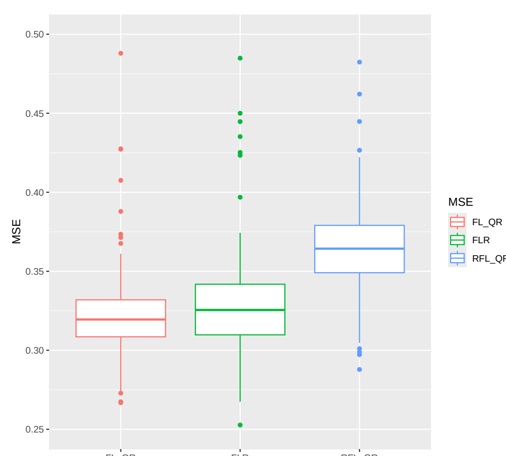
(c) $n = 50$



(d) $n = 100$



(e) $n = 50$



(f) $n = 100$

Figure S2: Boxplots of MSE when $\tau = 0.90$ for Example 2. blue: FLR method, : FL-QR method, green: RFL-QR method. The first row corresponds to the normal errors, the second row to the Cauchy errors, and the third row to the Chi-square errors.

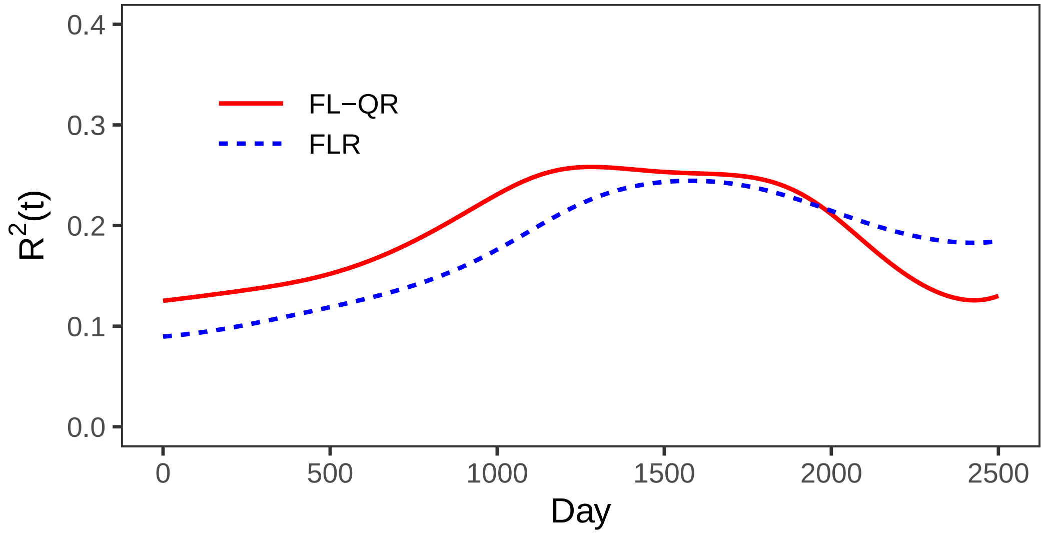


Figure S3: Estimated pointwise coefficient of determination $\hat{R}^2(t)$ for FLR and FL-QR.

They are displayed in Figure S3, indicating that the dynamics of albumin in FL-QR are more capable of explaining the total variation of prothrombin time trajectories over a more time range (from 0 to 1975 days), than the one in FLQ. In addition, it indicates generally stronger linear association at intermediate days (1000 to 2000 days) compare to the earlier days (0 to 500 days) and later days (2250 to 2500 days).

Last, we reconstruct mean trajectories of prothrombin times by using FLR and quantile trajectories of prothrombin times by applying FL-QR with the levels of quantile $\tau = 0.1, 0.5$ and 0.9, which is presented in Figure S4. We see that these trajectories have the same growth mode; mean and quantile with $\tau = 0.5$ trajectories of prothrombin times are almost identical, which implies that the conditional distributions of prothrombin time given albumin at each day don't skew; our FL-QR can capture lower (e.g. $\tau = 0.1$) and upper (e.g. $\tau = 0.9$) conditional quantiles of the trajectories of prothrombin time, which cannot be characterized

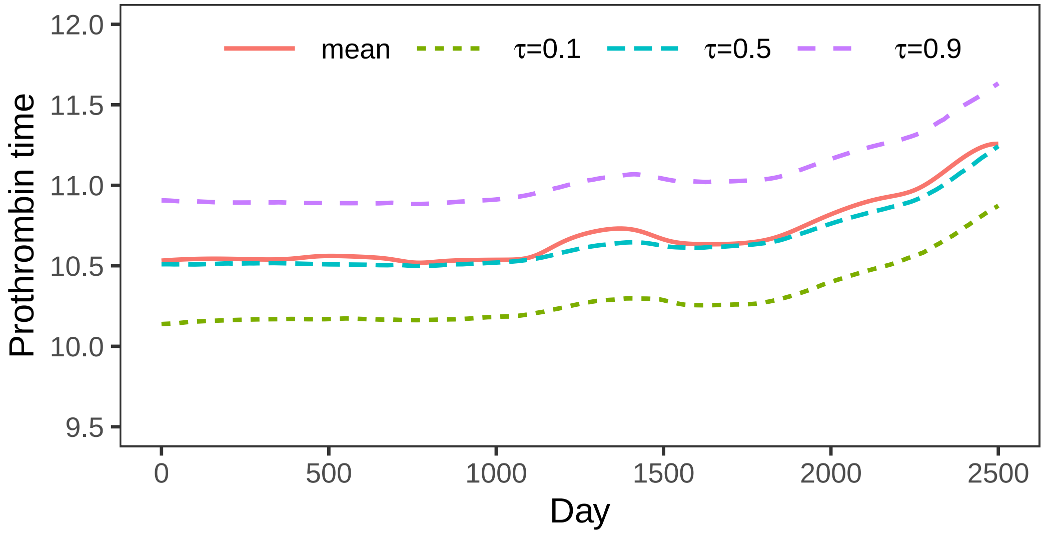


Figure S4: Mean regression function of Y in FLR and quantile regression functions of Y in FL-QR with $\tau = 0.1, 0.5$ and 0.9.

by analyzing the conditional mean of FLR model alone.

References

Chen, X., W. D. Liu, X. J. Mao, and Z. Y. Yang (2020). Distributed high-dimensional regression under a quantile loss function. *Journal of Machine learning Research 21*, 1–43.

Conway, J. B. (1985). *A Course in Functional Analysis*. New York: Springer-Verlag.

Şentürk, D. and H.-G. Müller (2010). Functional varying coefficient models for longitudinal data. *Journal of the American Statistical Association 105*, 1256–1264.

He, G., H. G. Müller, and L. Wang (2000). Extending correlation and regression from multivariate to functional data. In M. L. Puri (Ed.), *Asymptotics in Statistics and Probability*, pp. pp. 301–315. VSP International Science Publishers.

Li, Y. and T. Hsing (2010). Uniform convergence rates for nonparametric regression and principal component analysis in function/longitudinal data. *Ann. Statist. 38*, 3321–3351.

Yao, F. (2007). Asymptotic distributions of nonparametric regression estimators for longitudinal or functional data. *J. Mult. Anal. 98*, 40–56.

Yao, F., H. G. Müller, and J. L. Wang (2005a). Functional data analysis for sparse longitudinal data. *J. Am. Statist. Assoc. 100*, 577–590.

Yao, F., H. G. Müller, and J. L. Wang (2005b). Functional linear regression analysis for longitudinal data. *Ann. Statist. 33*, 2873–2903.