# Supplementary Materials for "Conditional analysis for mixed covariates, with application to feed intake of lactating sows"

S.Y. Park[1,*], C. Li[2,**], S.M. Mendoza Benavides[3], E. van Heugten[3] and A.M. Staicu[2]

Wednesday 29th May, 2019

The supplement contains two sections: Section S1 presents additional simulations results, Section S2 describes additional real data analysis for bike sharing data.

# S1 Additional simulation results

Section S1.1 provides simulation settings and the results for the case of having a single scalar covariate. Section S1.2 summarizes simulation results for the case of having a single functional covariate. Section S1.3 presents simulation results for the case of having a scalar covariate and a densely observed functional covariate.

## S1.1 Simulation study for scalar covariate only case

We compare the performance of the proposed method to the conventional quantile regression, and consider generating five models with different settings but all with scalar-only covariates as follows. The first three models follow Reich et al. (2010) and are all with linear quantiles,

$$\begin{array}{llcl} \text{Model 1} & Y_i & = & 1 + X_i + \epsilon_i, \\ \text{Model 2} & Y_i & = & 1 + X_i + \pi_i \epsilon_{1i} + (1-\pi_i)\epsilon_{2i}, \\ \text{Model 3} & Y_i & = & 1 + X_i + (1.1 - X_i)\epsilon_i, \end{array}$$

Model 1 is generated as above, where $X_i \overset{i.i.d}{\sim} N(0,1)$, $\epsilon_i \overset{i.i.d}{\sim} N(0,1)$. Since $Y_i \sim N(1+X_i, 1)$, the linear quantile can be analytically determined by $Q_\tau = 1 + X_i + \Phi^{-1}(\tau)$. Model 2 is similar to the first one, the setting here is a model with mixture errors, where $X_i \overset{i.i.d}{\sim} N(0,1)$, $\pi_i \overset{i.i.d}{\sim} Unif(0,1)$, $\epsilon_{1i} \overset{i.i.d}{\sim} N(0,1) \perp\!\!\!\perp \epsilon_{2i} \overset{i.i.d}{\sim} N(3,3)$. Since $Y_i \sim N(1+X_i, \pi_i^2 + 3(1-\pi_i)^2)$, $Q_\tau = 1 + X_i + \sqrt{\pi_i^2 + 3(1-\pi_i)^2}\Phi^{-1}(\tau)$. Model 3 is generated with heteroscedastic error, where $X_i \overset{i.i.d}{\sim} Unif(-1,1)$, $\epsilon_i \overset{i.i.d}{\sim} N(0,1)$. True linear quantile can be obtained according to $Y_i \sim N(1+X_i,(1.1-X_i)^2)$, $Q_\tau = 1+X_i+(1.1-X_i)\Phi^{-1}(\tau)$. Model 4 and model 5 are generated under the same frame of non-linear quantile considered in Bondell et al. (2010),

$$\begin{aligned}
& Y_i &=& \; f(X_i)+g(X_i)\epsilon_i, \\
\text{Model 4} \quad & f(X) &=& \; 0.5+2X+\sin(2\pi X-0.5) \qquad g(X)=1, \\
\text{Model 5} \quad & f(X) &=& \; 3X \qquad g(X)=0.5+2X+\sin(2\pi X-0.5),
\end{aligned}$$

where $X_i \overset{i.i.d}{\sim} Unif(-1,1)$, $\epsilon_i \overset{i.i.d}{\sim} N(0,1)$. For model 4, $Y_i \sim N(f(X_i),1)$, $Q_\tau = 0.5+2X_i+\sin(2\pi X_i-0.5)+\Phi^{-1}(\tau)$; for model 5 with heteroscedastic error, $Y_i \sim N(3X_i, g(X_i)^2)$, $Q_\tau = 3X_i+(0.5+2X_i+\sin(2\pi X_i-0.5))\Phi^{-1}(\tau)$. All covariates and error terms appearing in the models are mutually independent. The performance of each method is evaluated in terms of MAE again on seven quantile levels of interests ranging from 0.05 to 0.95.

We compare the proposed method with several alternatives, pointwise QR, LQR as introduced before, COBS (Constrained B-Spline Smoothing) implemented by `cobs` from `COBS` package (Ng and Maechler, 2007) in `R` and a variant of our approach by ignoring the binary-valued nature of the functional response, and thus using identity link function (Joint QR (G): `pffr`, Gaussian). Moreover we also consider the proposed methods (Pointwise QR and Joint QR) with non-linear modeling of the conditional distribution, i.e. $F_{Y|X}(y) = g^{-1}\{\beta_0(y)+h(X,y)\}$, where $\beta_0(\cdot)$ and $h(\cdot,\cdot)$ are unknown smooth functions; the methods are denoted by Pointwise QR (NL) and Joint QR (NL). Note that `pfr` cannot only take scalar covariates as input directly, thus we actually implement a generalized linear model.

As before, we set sample size $n = 1000$ as a training set, and use the additional 100 as a testing set. Totally 500 Monte Carlo samples are generated. Results of our numerical study are presented as in Table S1.

## S1.2 Simulation study for functional covariate only case

Consider when there is only a single sparsely observed functional covariate. The observed data for the $i$th subject is $\{Y_i,(W_{i1},t_{i1}),\cdots,(W_{im_i},t_{im_i})\}$. For brevity, we omit the same settings described previously.

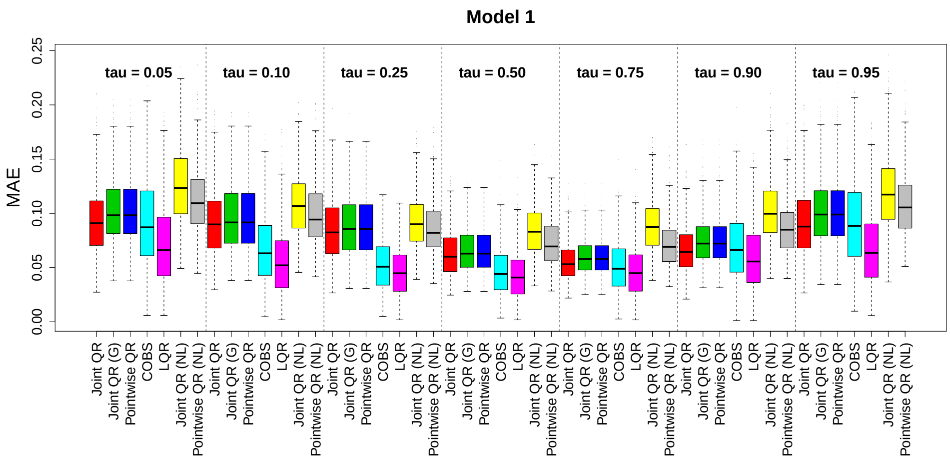
We consider two scenarios as before:

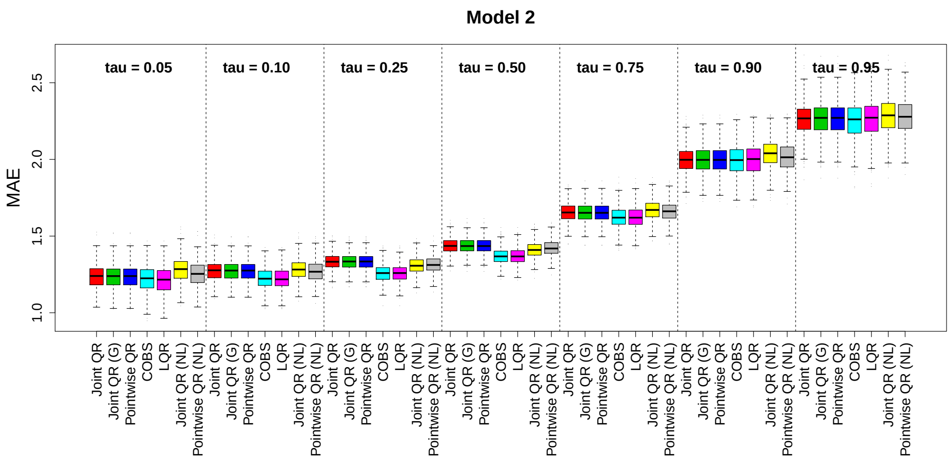
(i) normal distribution $Y_i|X_i(\cdot) \sim N(2\sum_{k=1}^4 \xi_{ik}, 5^2)$;

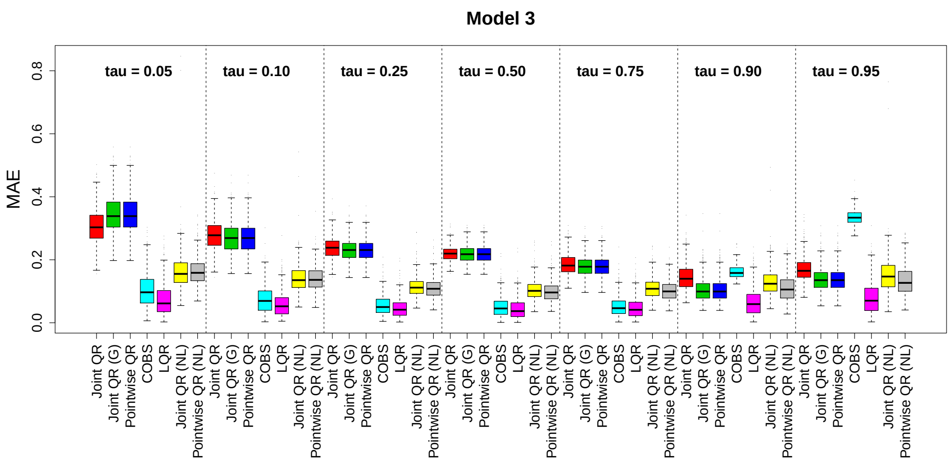
(ii) mixture distribution $Y_i|X_i(\cdot) \sim 0.5N(\sum_{k=1}^4 \xi_{ik}, 1^2)+0.5N(3\sum_{k=1}^4 \xi_{ik}, 4^2)$.

Table S1: Average MAE (standard error in parentheses) of the predicted $\tau$-level quantile for the case of having a scalar covariate. Sample size $n = 1000$.

| Model | Method | $\tau = 0.05$ | $\tau = 0.1$ | $\tau = 0.25$ | $\tau = 0.5$ | $\tau = 0.75$ | $\tau = 0.9$ | $\tau = 0.95$ |
|---|---|---|---|---|---|---|---|---|
| Model 1 | Joint QR | 0.09 (0.14) | 0.09 (0.15) | 0.09 (0.13) | 0.06 (0.10) | 0.06 (0.08) | 0.07 (0.10) | 0.09 (0.15) |
| | Joint QR (G) | 0.38 (0.26) | 0.34 (0.19) | 0.23 (0.14) | 0.15 (0.13) | 0.29 (0.14) | 0.40 (0.19) | 0.44 (0.25) |
| | Pointwise QR | 0.10 (0.14) | 0.10 (0.14) | 0.09 (0.13) | 0.07 (0.10) | 0.06 (0.08) | 0.08 (0.10) | 0.10 (0.14) |
| | COBS | 0.09 (0.19) | 0.07 (0.15) | **0.05** (0.11) | 0.05 (0.11) | **0.05** (0.12) | 0.07 (0.15) | 0.09 (0.19) |
| | LQR | **0.07** (0.17) | **0.06** (0.14) | **0.05** (0.10) | **0.04** (0.10) | **0.05** (0.11) | **0.06** (0.14) | **0.07** (0.16) |
| | Joint QR (NL) | 0.13 (0.16) | 0.11 (0.13) | 0.09 (0.11) | 0.08 (0.10) | 0.09 (0.12) | 0.10 (0.13) | 0.12 (0.16) |
| | Pointwise QR (NL) | 0.11 (0.14) | 0.10 (0.13) | 0.09 (0.11) | 0.07 (0.10) | 0.07 (0.10) | 0.09 (0.11) | 0.11 (0.13) |
| Model 2 | Joint QR | 1.24 (0.35) | 1.27 (0.29) | 1.33 (0.23) | 1.44 (0.22) | 1.65 (0.28) | **2.00** (0.40) | 2.26 (0.51) |
| | Joint QR (G) | **0.94** (0.43) | **0.98** (0.34) | **1.17** (0.26) | 1.48 (0.23) | 1.83 (0.28) | 2.14 (0.40) | 2.43 (0.58) |
| | Pointwise QR | 1.24 (0.35) | 1.27 (0.29) | 1.33 (0.24) | 1.44 (0.23) | 1.65 (0.28) | **2.00** (0.40) | 2.27 (0.51) |
| | COBS | 1.22 (0.41) | 1.22 (0.33) | 1.26 (0.26) | **1.37** (0.25) | **1.62** (0.33) | **2.00** (0.48) | **2.25** (0.60) |
| | LQR | 1.21 (0.40) | 1.22 (0.33) | 1.26 (0.25) | **1.37** (0.24) | **1.62** (0.32) | **2.00** (0.46) | 2.26 (0.57) |
| | Joint QR (NL) | 1.28 (0.36) | 1.28 (0.29) | 1.31 (0.25) | 1.41 (0.24) | 1.67 (0.30) | 2.04 (0.42) | 2.29 (0.53) |
| | Pointwise QR (NL) | 1.26 (0.37) | 1.27 (0.30) | 1.31 (0.24) | 1.42 (0.23) | 1.66 (0.29) | 2.01 (0.42) | 2.27 (0.52) |
| Model 3 | Joint QR | 0.30 (0.25) | 0.28 (0.22) | 0.24 (0.15) | 0.22 (0.11) | 0.19 (0.15) | 0.15 (0.19) | 0.17 (0.17) |
| | Joint QR (G) | 0.68 (0.71) | 0.41 (0.25) | 0.15 (0.13) | 0.20 (0.14) | 0.16 (0.11) | 0.31 (0.23) | 0.49 (0.75) |
| | Pointwise QR | 0.34 (0.27) | 0.27 (0.22) | 0.23 (0.15) | 0.22 (0.12) | 0.18 (0.14) | 0.11 (0.17) | 0.14 (0.17) |
| | COBS | 0.10 (0.26) | 0.08 (0.21) | 0.06 (0.15) | 0.05 (0.15) | **0.05** (0.15) | 0.16 (0.10) | 0.34 (0.11) |
| | LQR | **0.07** (0.23) | **0.06** (0.18) | **0.05** (0.14) | **0.04** (0.14) | **0.05** (0.15) | **0.07** (0.19) | **0.08** (0.23) |
| | Joint QR (NL) | 0.17 (0.20) | 0.14 (0.18) | 0.11 (0.14) | 0.10 (0.14) | 0.10 (0.15) | 0.11 (0.19) | 0.13 (0.21) |
| | Pointwise QR (NL) | 0.16 (0.28) | 0.14 (0.20) | 0.11 (0.15) | 0.10 (0.14) | 0.11 (0.15) | 0.13 (0.20) | 0.15 (0.26) |
| Model 4 | Joint QR | 0.61 (0.22) | 0.60 (0.18) | 0.59 (0.15) | 0.58 (0.14) | 0.58 (0.15) | 0.60 (0.20) | 0.64 (0.26) |
| | Joint QR (G) | 0.73 (0.34) | 0.65 (0.22) | 0.58 (0.17) | 0.55 (0.14) | 0.59 (0.19) | 0.64 (0.25) | 0.69 (0.33) |
| | Pointwise QR | 0.62 (0.23) | 0.60 (0.19) | 0.58 (0.15) | 0.58 (0.14) | 0.58 (0.15) | 0.60 (0.20) | 0.65 (0.27) |
| | COBS | 0.35 (0.17) | 0.34 (0.15) | 0.32 (0.13) | 0.32 (0.12) | 0.32 (0.12) | 0.33 (0.15) | 0.34 (0.16) |
| | LQR | 0.62 (0.21) | 0.60 (0.19) | 0.58 (0.16) | 0.58 (0.14) | 0.60 (0.16) | 0.63 (0.19) | 0.65 (0.22) |
| | Joint QR (NL) | **0.14** (0.15) | **0.12** (0.13) | **0.11** (0.11) | **0.10** (0.10) | **0.10** (0.11) | **0.12** (0.13) | **0.14** (0.16) |
| | Pointwise QR (NL) | 0.18 (0.16) | 0.15 (0.13) | 0.12 (0.11) | 0.11 (0.10) | 0.11 (0.11) | 0.14 (0.13) | 0.17 (0.17) |
| Model 5 | Joint QR | 0.74 (0.34) | 0.57 (0.25) | 0.34 (0.19) | 0.22 (0.13) | 0.30 (0.13) | 0.55 (0.20) | 0.78 (0.29) |
| | Joint QR (G) | 1.11 (0.93) | 0.81 (0.33) | 0.57 (0.21) | 0.27 (0.14) | 0.60 (0.23) | 1.07 (0.42) | 1.52 (0.63) |
| | Pointwise QR | 0.76 (0.34) | 0.56 (0.24) | 0.33 (0.19) | 0.22 (0.14) | 0.30 (0.14) | 0.54 (0.21) | 0.79 (0.30) |
| | COBS | 0.38 (0.23) | 0.29 (0.19) | **0.16** (0.17) | **0.03** (0.11) | **0.16** (0.19) | 0.24 (0.22) | 0.33 (0.24) |
| | LQR | 0.84 (0.30) | 0.59 (0.21) | 0.30 (0.13) | **0.03** (0.09) | 0.30 (0.13) | 0.59 (0.20) | 0.84 (0.29) |
| | Joint QR (NL) | 0.34 (0.22) | 0.27 (0.18) | 0.17 (0.14) | 0.14 (0.15) | 0.17 (0.16) | 0.24 (0.21) | 0.30 (0.29) |
| | Pointwise QR (NL) | **0.32** (0.24) | **0.25** (0.21) | 0.18 (0.17) | 0.15 (0.14) | **0.16** (0.15) | **0.22** (0.18) | **0.28** (0.23) |


Model 1
MAE
tau = 0.05
tau = 0.10
tau = 0.25
tau = 0.50
tau = 0.75
tau = 0.90
tau = 0.95
Joint QR
Joint QR (G)
Pointwise QR
COBS
LQR
Joint QR (NL)
Pointwise QR (NL)


Model 2
MAE
tau = 0.05
tau = 0.10
tau = 0.25
tau = 0.50
tau = 0.75
tau = 0.90
tau = 0.95
Joint QR
Joint QR (G)
Pointwise QR
COBS
LQR
Joint QR (NL)
Pointwise QR (NL)


Model 3
MAE
tau = 0.05
tau = 0.10
tau = 0.25
tau = 0.50
tau = 0.75
tau = 0.90
tau = 0.95
Joint QR
Joint QR (G)
Pointwise QR
COBS
LQR
Joint QR (NL)
Pointwise QR (NL)

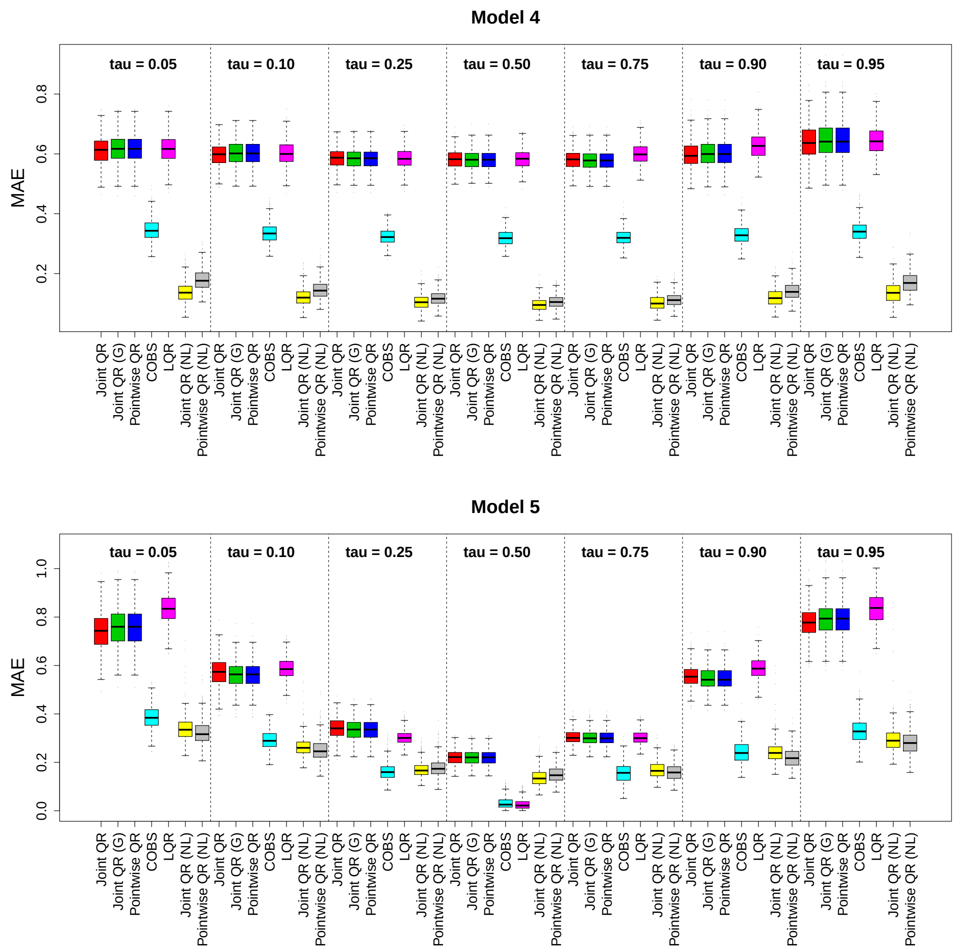



Figure S1: Boxplots of MAEs of the predicted $\tau$-level quantile for sample size $n = 1000$ for the case of having a scalar covariate only. Results are based on 500 replication.

Sample size and replication number are set as 1000 and 500, respectively. We compare the performance of our method with three alternatives: Pointwise QR, CM and Joint QR (G). The results are presented in Tables S2 and S3.

Table S2: Average MAE (standard error in parentheses) of the predicted $\tau$-level quantile for the case of having a sparsely observed functional covariate. Sample size $n = 1000$.

| Distribution | SNR | Method | $\tau = 0.05$ | $\tau = 0.1$ | $\tau = 0.25$ | $\tau = 0.5$ |
|---|---|---|---|---|---|---|
| Normal | 2 | Joint QR | **7.91** (0.03) | **7.04** (0.03) | **6.24** (0.02) | **6.05** (0.02) |
| | | Joint QR (G) | 10.95 (0.04) | 9.28 (0.03) | 7.18 (0.02) | 6.30 (0.02) |
| | | Pointwise QR | 8.06 (0.03) | 7.13 (0.03) | 6.28 (0.02) | 6.06 (0.02) |
| | | CM | 7.99 (0.03) | 7.14 (0.03) | 6.35 (0.02) | 6.15 (0.02) |
| Normal | 1 | Joint QR | **10.02** (0.04) | **8.83** (0.03) | **7.59** (0.03) | **7.23** (0.02) |
| | | Joint QR (G) | 11.80 (0.04) | 10.18 (0.04) | 8.17 (0.03) | 7.34 (0.03) |
| | | Pointwise QR | 10.17 (0.04) | 8.91 (0.03) | 7.62 (0.03) | 7.24 (0.03) |
| | | CM | 10.10 (0.04) | 8.95 (0.03) | 7.73 (0.03) | 7.35 (0.03) |
| Mixture | 2 | Joint QR | **9.61** (0.04) | **7.74** (0.03) | 6.32 (0.03) | 4.59 (0.02) |
| | | Joint QR (G) | 14.45 (0.07) | 10.28 (0.04) | 7.16 (0.02) | **4.29** (0.01) |
| | | Pointwise QR | 9.88 (0.04) | 7.77 (0.03) | **6.27** (0.03) | 4.53 (0.02) |
| | | CM | 9.81 (0.04) | 7.90 (0.03) | 6.43 (0.03) | 4.73 (0.02) |
| Mixture | 1 | Joint QR | **12.03** (0.04) | **9.60** (0.03) | 7.53 (0.03) | 5.27 (0.02) |
| | | Joint QR (G) | 15.15 (0.07) | 11.19 (0.04) | 8.12 (0.03) | **5.10** (0.02) |
| | | Pointwise QR | 12.39 (0.04) | 9.69 (0.03) | **7.51** (0.03) | 5.23 (0.02) |
| | | CM | 12.30 (0.04) | 9.85 (0.03) | 7.70 (0.03) | 5.43 (0.02) |

Table S3: Average computing times (in seconds).

| Distribution | Joint QR | Joint QR (G) | Pointwise QR | CM |
|---|---|---|---|---|
| Normal | 179 | 125 | 310 | 477 |
| Mixture | 206 | 134 | 313 | 314 |

## S1.3 Simulation results for the case of having a scalar covariate and a densely observed functional covariate

Tables S4 and S5 summarize the simulation results for the case of having a scalar covariate and a densely observed functional covariate. Simulation settings are described in the main manuscript.

Table S4: Average MAE (standard error in parentheses) of the predicted $\tau$-level quantile for the case of having a scalar covariate and a densely observed functional covariate. Sample size $n = 100$.

| Distribution | SNR | Method | $\tau = 0.05$ | $\tau = 0.1$ | $\tau = 0.25$ | $\tau = 0.5$ |
|---|---|---|---|---|---|---|
| Normal | 150 | Joint QR | 3.30(0.03) | 3.15(0.03) | 2.90(0.02) | 2.71(0.02) |
| | | Pointwise QR | 4.41 (0.03) | 4.06 (0.03) | 3.67 (0.02) | 3.59 (0.02) |
| | | Mod CM | 4.32 (0.03) | 4.04 (0.03) | 3.67 (0.02) | 3.53 (0.02) |
| | | PQR | **2.97** (0.04) | **2.72** (0.04) | **2.50** (0.04) | **2.45** (0.04) |
| Normal | 2 | Joint QR | **8.13** (0.04) | **7.57** (0.03) | **7.13** (0.03) | 7.02 (0.03) |
| | | Pointwise QR | 9.37 (0.05) | 8.46 (0.04) | 7.63 (0.03) | 7.36 (0.03) |
| | | Mod CM | 8.88 (0.04) | 8.69 (0.04) | 8.52 (0.04) | 8.60 (0.04) |
| | | PQR | 8.76 (0.06) | 8.04 (0.05) | 7.13 (0.04) | **6.76** (0.03) |
| Normal | 1 | Joint QR | **9.90** (0.05) | **9.05** (0.04) | **8.32** (0.03) | 8.13 (0.03) |
| | | Pointwise QR | 11.07 (0.06) | 9.90 (0.05) | 8.78 (0.03) | 8.41 (0.03) |
| | | Mod CM | 10.27 (0.04) | 10.00 (0.04) | 9.79 (0.04) | 9.84 (0.04) |
| | | PQR | 10.67 (0.07) | 9.71 (0.06) | 8.40 (0.04) | **7.86** (0.03) |
| Mixture | 150 | Joint QR | **6.59** (0.06) | **6.27** (0.06) | **6.07** (0.06) | **4.45** (0.05) |
| | | Pointwise QR | 7.53 (0.07) | 6.13 (0.06) | 6.07 (0.06) | 4.64 (0.06) |
| | | Mod CM | 6.66 (0.06) | 6.44 (0.06) | 6.37 (0.06) | 4.95 (0.06) |
| | | PQR | 8.08 (0.07) | 7.06 (0.05) | 6.32 (0.04) | 5.91 (0.14) |
| Mixture | 2 | Joint QR | **10.40** (0.06) | **9.06** (0.05) | **8.60** (0.05) | 6.89 (0.05) |
| | | Pointwise QR | 11.70 (0.08) | 9.65 (0.06) | 8.61 (0.05) | 6.88 (0.05) |
| | | Mod CM | 11.40 (0.06) | 10.89 (0.06) | 10.68 (0.06) | 9.34 (0.07) |
| | | PQR | 12.00 (0.08) | 10.28 (0.06) | 8.68 (0.04) | **6.09** (0.06) |
| Mixture | 1 | Joint QR | **11.79** (0.07) | **10.15** (0.05) | 9.38 (0.05) | 7.53 (0.05) |
| | | Pointwise QR | 12.94 (0.09) | 10.68 (0.06) | **9.37** (0.05) | 7.48 (0.05) |
| | | Mod CM | 12.61 (0.07) | 11.95 (0.06) | 11.60 (0.06) | 10.23 (0.07) |
| | | PQR | 13.40 (0.10) | 11.47 (0.07) | 9.48 (0.04) | **6.59** (0.05) |

Table S5: Average MAE (standard error in parentheses) of the predicted $\tau$-level quantile for the case of having a scalar covariate and a densely observed functional covariate. Sample size $n = 1000$.

| Distribution | SNR | Method | $\tau = 0.05$ | $\tau = 0.1$ | $\tau = 0.25$ | $\tau = 0.5$ |
|---|---|---|---|---|---|---|
| Normal | 150 | Joint QR | **1.27** (0.01) | **1.34** (0.01) | **1.33** (0.01) | **1.22** (0.01) |
| | | Pointwise QR | 1.61 (0.01) | 1.59 (0.01) | 1.54 (0.01) | 1.43 (0.01) |
| | | Mod CM | 1.43 (0.01) | 1.40 (0.01) | 1.39 (0.01) | 1.36 (0.01) |
| | | PQR | 1.74 (0.02) | 1.71 (0.02) | 1.67 (0.02) | 1.67 (0.02) |
| Normal | 2 | Joint QR | **7.79** (0.03) | **6.91** (0.02) | **6.22** (0.02) | **6.11** (0.02) |
| | | Pointwise QR | 8.03 (0.03) | 7.06 (0.03) | 6.31 (0.02) | 6.19 (0.02) |
| | | Mod CM | 7.94 (0.03) | 7.11 (0.03) | 6.43 (0.02) | 6.32 (0.02) |
| | | PQR | 8.36 (0.04) | 7.55 (0.04) | 6.60 (0.03) | 6.22 (0.02) |
| Normal | 1 | Joint QR | **9.84** (0.04) | **8.59** (0.03) | **7.52** (0.03) | **7.30** (0.03) |
| | | Pointwise QR | 10.09 (0.04) | 8.76 (0.03) | 7.61 (0.03) | 7.35 (0.03) |
| | | Mod CM | 9.95 (0.04) | 8.79 (0.03) | 7.77 (0.03) | 7.55 (0.03) |
| | | PQR | 10.39 (0.05) | 9.26 (0.04) | 7.89 (0.03) | 7.33 (0.02) |
| Mixture | 150 | Joint QR | 4.33 (0.03) | 3.89 (0.02) | 3.66 (0.03) | 3.54 (0.03) |
| | | Pointwise QR | **4.04** (0.03) | **3.82** (0.02) | **3.65** (0.03) | **3.45** (0.03) |
| | | Mod CM | 4.12 (0.03) | 3.88 (0.02) | 3.68 (0.03) | 3.65 (0.03) |
| | | PQR | 7.74 (0.04) | 6.40 (0.03) | 5.31 (0.02) | 3.75 (0.10) |
| Mixture | 2 | Joint QR | **9.62** (0.04) | 7.46 (0.03) | 7.10 (0.03) | 5.74 (0.03) |
| | | Pointwise QR | 9.64 (0.04) | **7.37** (0.03) | **6.86** (0.03) | 5.67 (0.03) |
| | | Mod CM | 9.64 (0.04) | 7.60 (0.03) | 7.18 (0.03) | 6.12 (0.03) |
| | | PQR | 11.74 (0.05) | 9.65 (0.04) | 7.82 (0.03) | **4.64** (0.02) |
| Mixture | 1 | Joint QR | **11.47** (0.04) | 8.94 (0.03) | 8.00 (0.03) | 6.33 (0.03) |
| | | Pointwise QR | 11.55 (0.05) | **8.91** (0.03) | **7.82** (0.03) | 6.22 (0.03) |
| | | Mod CM | 11.48 (0.04) | 9.17 (0.04) | 8.22 (0.04) | 6.81 (0.03) |
| | | PQR | 13.33 (0.06) | 10.87 (0.04) | 8.65 (0.03) | **5.37** (0.02) |

# S2 Bike sharing data application

In this section we illustrate the proposed method using the bike sharing data (Fanaee-T and Gama, 2014) available at the University of California, Irvine (UCI) Machine Learning Repository (Lichman, 2013). The data set consists of hourly counts of bike rentals in Washington, D.C. recorded by the Capital Bike Sharing (CBS) system from January 1, 2011 to December 31, 2012, with additional information on corresponding users, such as membership (casual vs. registered), and on corresponding days, such as season, temperature, and etc. More detailed descriptions on the data set are provided in Fanaee-T and Gama (2014).

The CBS system is an automated bike sharing system, where users can rent bikes from (and return them to) any bike docks located across Washington, D.C. on an hourly basis. One of main challenges associated with the system is to understand and forecast a daily demand of bike rentals in a supply chain. In most cases, such demand analysis involves studying a quantile of the response instead of its average; for example, a supplier would probably be more interested in covering 90% of a demand than only an average demand. Here we use the proposed method to study the total number of bikes rented by casual users on Saturday ($Y$), using average of hourly 'feeling' temperatures on Saturday ($X_1$) and the hourly counts of bike rentals by casual users on the previous day ($X_2(t)$ for hour of a day $t = 1, 2, \dots, 24$). In this analysis short term weather-related forecasts are assumed to be accurate.

There are total of $n = 104$ Saturdays during the period that the data were collected, with about four to five weekly measurements per each month. Raw responses of total counts of Saturday bike rentals, $\{Y_i : i = 1, 2, \dots, 104\}$, suggest strong seasonal and year effects; see Figure S4. To remove these effects prior to the analysis, we fit a linear model to the log-transformed response, $\log(Y+1)$, using three dummy variables for season and one dummy variable for year and obtain residuals shown in Figure S5. The resulting residuals are then used as our response variable in the subsequent analysis procedure; henceforth, with abuse of notation, denote the resulting residuals by $Y$. Alternatively, we also consider fitting a generalized linear model with a Poisson distribution to the original responses. However, as two approaches give a similar conclusion in our analysis, we exclude the results obtained using the second approach.

While the data set includes hourly 'feeling' temperatures on Saturday, we use their average as one of the predictors because a forecast for the average daily temperature is expected to be less variable than that for the hourly temperatures. Specifically we consider the *centered* average daily 'feeling' temperature of Saturday as our scalar covariate in the analysis. Whereas hourly counts of bike rentals by casual users on Fridays are considered in the analysis as noisy functional observations with mild missingness (about 0.35% of the data is missing). Prior to the analysis we pre-smooth the observed functional predictor and center it; by following an approach taken by Goldsmith et al. (2011) and Ivanescu et al. (2015), pre-smoothing is done using the principal analysis by conditional expectation (PACE) method (Yao et al., 2005) with percentage of variance explained (PVE) equal to 0.99. With abuse of notation, we denote the resulting, pre-processed scalar and functional covariates for the $i$th Saturday with $X_{1,i}$ and $X_{2,i}(t)$, respectively; see Figures S6 - S8.

In the following we first consider three different quantile models:

(M1) a model with a scalar covariate $X_1$ only

(M2) a model with a functional covariate $X_2(t)$ only

(M3) a model with both covariates $X_1$ and $X_2(t)$

When applicable, we also compare the prediction accuracy of our proposed method with other available methods that are closely related to ours, namely LQR and CM. We quantify prediction accuracy using leave-one-out cross validation (LOO-CV) with the loss function typically used in a quantile regression: $\rho_\tau(q) = (1-\tau)\sum_{y_i<q}(y_i-q)+\tau\sum_{y_i\geq q}(y_i-q)$. Specifically we define out-of-sample prediction error for given $\tau$ as $\text{OUT-PE}_\tau = n^{-1}\sum_{i=1}^{n}\rho_\tau\left(\widehat{Q}^{[-i]}_{Y|X}(\tau)\right)$, where $\widehat{Q}^{[-i]}_{Y|X}(\tau)$ is the predicted $\tau$th quantile obtained using the $i$th training set; we prepare the $i$th training set by leaving the $i$th observations out and then pre-processing the rest as described earlier in this section. Covariates in the testing sets are also appropriately pre-processed. For example, consider the $i$th testing set; the scalar covariate, $X_{1,i}$, is centered by subtracting the average of $X_{1,i}$'s in the $i$th training set, whereas the noisy functional covariate, $X_{2,i}(t)$, is smoothed and is centered using the mean and principal component (PC) estimates obtained using $X_{2,i}(t)$'s in the $i$th training set.

In terms of prediction accuracy we compare the proposed method (Joint QR) with Pointwise QR for all of the three models (M1) - (M3) considered. Additionally LQR and CM are considered for the models (M1) and (M2), respectively. $\text{OUT-PE}_\tau$ is obtained for $\tau = 0.1, 0.2, \ldots, 0.9$ and the results are summarized in Figure S2. The results are consistent with the simulation results given in Sections 4 and S1. As shown in Figure S2, Pointwise QR and Joint QR have very similar prediction accuracy for all three models (M1)-(M3). When considering only a functional covariate, prediction accuracy of the CM method also lies in the same range as Pointwise QR and

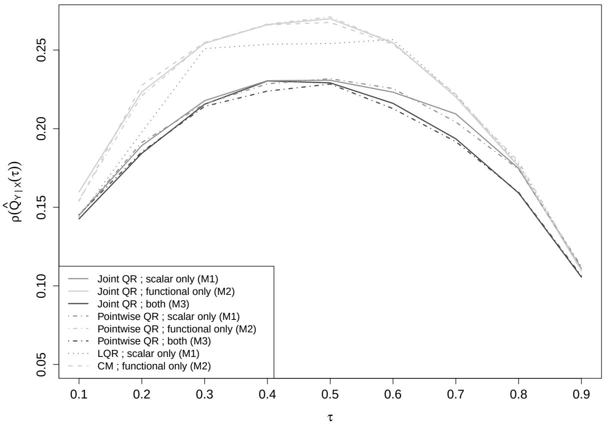
Joint QR; this is consistent with the simulation results given in Section S1.2. When considering a scalar covariate only, the LQR estimation method has substantially higher OUT-PE$_\tau$ than Pointwise QR and Joint QR for all $\tau$'s. It implies that the relationship between quantiles of the total number of bike rentals of Saturday and average daily 'feeling' temperature on Saturday may not be linear. In summary for models (M1) and (M2), the propose method provides as reliable prediction as other existing methods and there is no obvious estimation method that outperforms the others.

However the existing methods are not applicable for model (M3), when there is both scalar and functional covariates $X_1$ and $X_2(t)$. One of advantages of the proposed methods (Joint and Pointwise QR) is that now we can compare among models (M1) - (M3) and investigate which covariates have more effects on the conditional distribution of the response, and consequently on the quantiles. For example, based on OUT-PE$_\tau$'s for Joint QR for (M1)-(M3) shown in Figure S9 we can see that average daily 'feeling' temperature on Saturday ($X_1$) provides more information on the total number of bike rental demands on Saturday ($Y$) than hourly counts of bike rentals on Friday ($X_2(t)$) does for all quantile levels considered. It seems that incorporating a functional covariate $X_2(t)$ in addition to a scalar covariate, $X_1$, does not have any effect on prediction accuracy for low quantile levels ($\tau < 0.5$). However, a model that incorporates both covariates, $X_1$ and $X_2(t)$, has slightly better prediction accuracy than a model with only scalar covariate, $X_1$, for higher quantile levels ($\tau > 0.5$).

As we are interested in studying demand of bike rentals, predicting quantiles at *high* quantile levels are particularly of interest in this analysis. Thus based on prediction accuracy of different models and estimation methods for $\tau > 0.5$, we choose a model that incorporates covariates of both types ($X_1$ and $X_2(t)$): $\mathrm{E}\left[I(Y < y)|X_1, X_2(t)\right] = \beta_0 + \beta_{X_1}(y)X_1 + \int X_2(t)\beta_{X_2}(t, y)dt$ and use the proposed method for estimation. The estimated coefficients, $\widehat{\beta}_{X_1}(y)$ and $\widehat{\beta}_{X_2}(t, y)$, are given in Figure S10. As in the analysis of the sow data in Section 5, we investigate the relationship between covariates and quantiles of the response (i.e. bike rental demands) for different quantile levels by studying predicted quantiles. While fixing $X_2(t)$ equal to the pointwise average of hourly Friday bike rentals, we predict quantiles of total demands of Saturday's bike rentals for fine grids of average feeling temperatures ($X_1$); again $\tau = 0.1, 0.2, \ldots, 0.9$ are considered. The predicted quantiles, $\widehat{Q}_{Y|X}(\tau)$, are given in Figure S3.

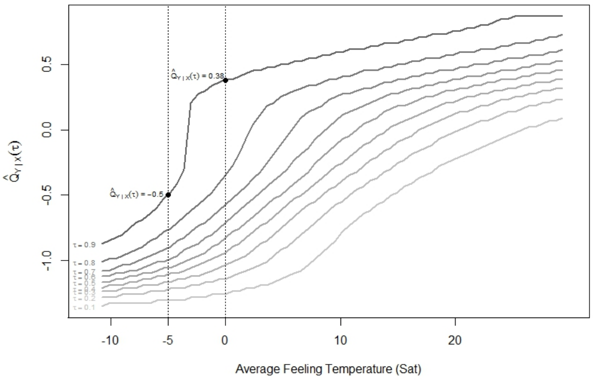
First, it is clear from the predicted quantiles curves that the relationship between average feeling temperature and quantile of $Y$ is positive and nonlinear for all $\tau$'s we consider. This is probably why prediction accuracy of

Figure S2: Out-of-sample prediction error, $\rho_\tau\left(\widehat{Q}_{Y|X}(\tau)\right)$ for three models (M1) - (M3) and four different methods (Joint QR, Pointwise QR, CM, LQR)




LQR is noticeably inferior to the proposed approach when only scalar covariate ($X_1$) in earlier comparison. Secondly, gaps between curves are wider for higher $\tau$'s imply that the density function of total bike rental demands on Saturday is right skewed; while there is implication of right-skewness across all values of average feeling temperature, we notice particularly large right-skewness around $-3°$C average feeling temperature and a steep increase in 90% quantiles between $-5°$C to 0°C. Specifically, the predicted 90% quantiles, $\widehat{Q}_{Y|X}(0.9)$, for $-5°$C and 0°C average Saturday feeling temperatures are about $-0.5$ and 0.38 respectively; the result suggest that, with increase in average Saturday feeling temperature from $-5°$C to 0°C, we expect 90% quantile of Saturday bike rentals demands to increase by more than twice (from 168 to 407) for winter of 2012.

Figure S3: Predicted quantiles against average Saturday feeling temperatures for $X_2(t)$ equal to pointwise average of hourly Friday bike rentals




## S2.1 Additional Figures for the Bike Sharing Data Analysis

Figure S4: Total counts of bike rentals on Saturday, $Y_i$

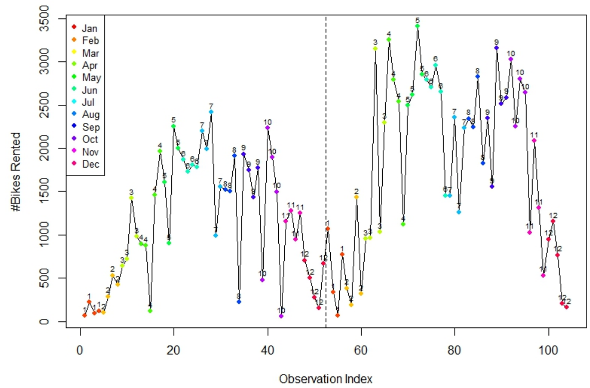


Figure S5: Pre-processed response; transformed responses, $\log(1 + Y)$ (top) and transformed responses without season and year effects (bottom)

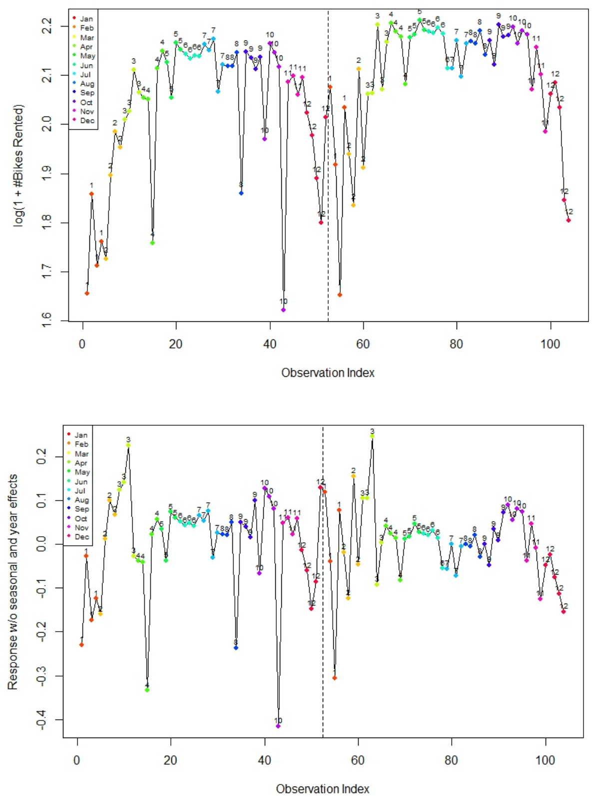



Figure S6: Average feeling temperatures on Saturday

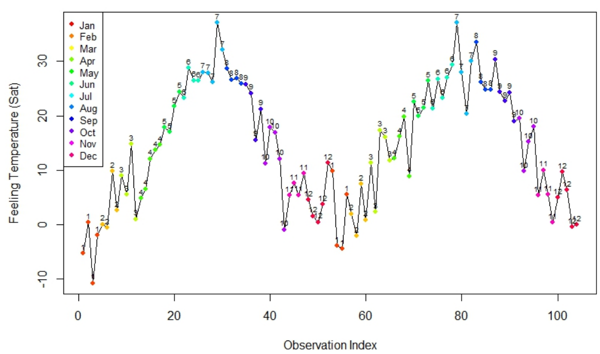


Figure S7: Spaghetti plots of hourly number of bike rentals on Friday; observed (top) and smoothed (bottomed)

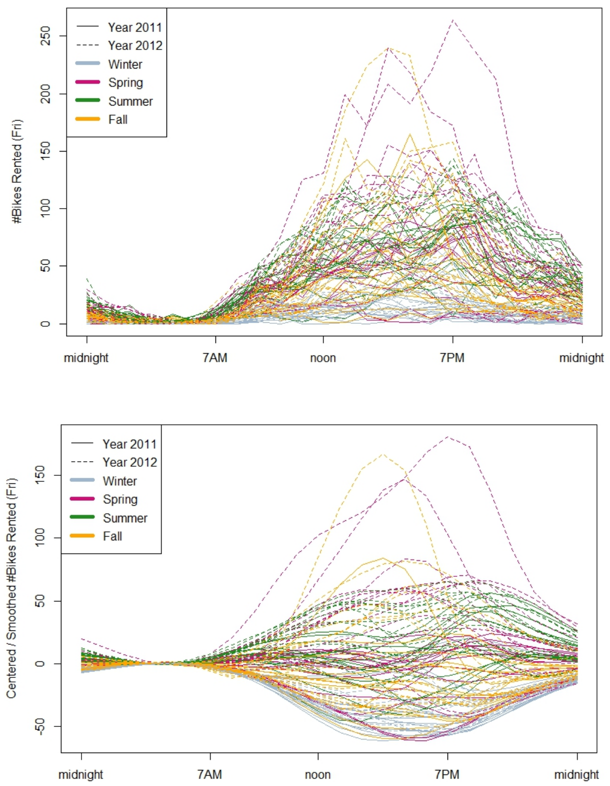
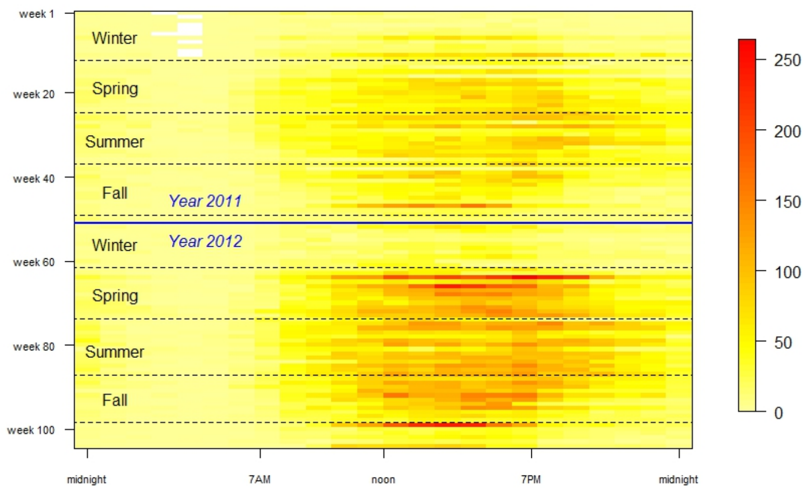


Figure S8: Lasagna plots of hourly number of bike rentals on Friday; observed (top) and smoothed (bottomed)



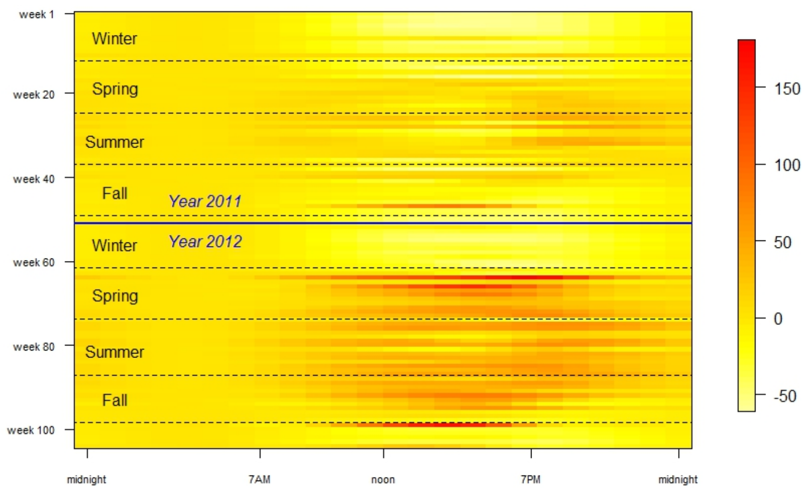



Figure S9: Out-of-sample prediction error, $\rho_\tau\left(\widehat{Q}_{Y|X}(\tau)\right)$, when the proposed estimation method (Joint QR) is used

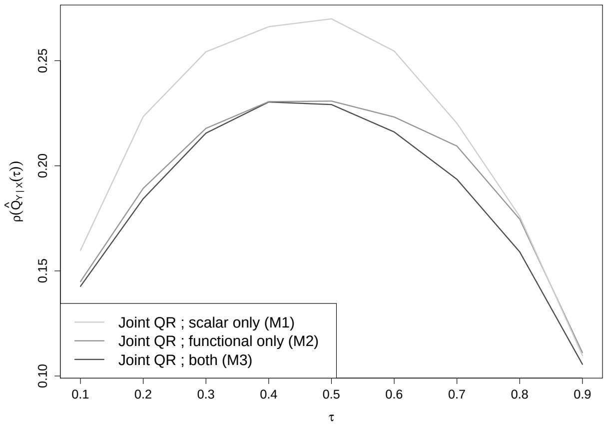
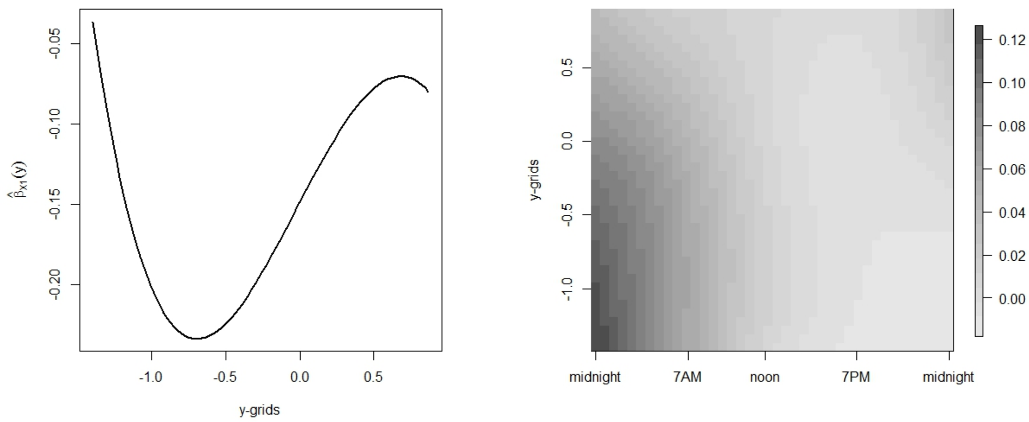



Figure S10: Estimated coefficients, $\widehat{\beta}_{X_1}(y)$ and $\widehat{\beta}_{X_2}(y,t)$



# References


Bondell, H., Reich, B., and Wang, H. (2010). Noncrossing quantile regression curve estimation. *Biometrika*, 97(4):825–838.

Fanaee-T, H. and Gama, J. (2014). Event labeling combining ensemble detectors and background knowledge. *Progress in Artificial Intelligence*, 2(2-3):113–127.

Goldsmith, J., Bobb, J., Crainiceanu, C. M., Caffo, B., and Reich, D. (2011). Penalized functional regression. *Journal of Computational and Graphical Statistics*, 20(4):830–851.

Ivanescu, A. E., Staicu, A.-M., Scheipl, F., and Greven, S. (2015). Penalized function-on-function regression. *Computational Statistics*, 30(2):539–568.

Lichman, M. (2013). UCI machine learning repository.

Ng, P. and Maechler, M. (2007). A fast and efficient implementation of qualitatively constrained quantile smoothing splines. *Statistical Modelling*, 7(4):315–328.

Reich, B., Bondell, H., and Wang, H. (2010). Flexible bayesian quantile regression for independent and clustered data. *Biostatistics*, 11(2):337–352.

Yao, F., Müller, H.-G., and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association*, 100(470):577–590.