Functional Censored Quantile Regression

Fei Jiang
Department of Statistics and Actuarial Science,
University of Hong Kong,
Qing Cheng
Center for Quantitative Medicine,
Duke-NUS Medical School
Guosheng Yin
Department of Statistics and Actuarial Science,
University of Hong Kong,
Haipeng Shen
Innovation and Information Management,
University of Hong Kong

Abstract

We propose a functional censored quantile regression model to describe the time-varying relationship between time-to-event outcomes and corresponding functional covariates. The time-varying effect is modeled as an unspecified function that is approximated via B-splines. A generalized approximate cross-validation method is developed to select the number of knots by minimizing the expected loss. We establish asymptotic properties of the method and the knot selection procedure. Furthermore, we conduct extensive simulation studies to evaluate the finite sample performance of our method. Finally, we analyze the functional relationship between ambulatory blood pressure trajectories and clinical outcome in stroke patients. The results reinforce the importance of the morning blood pressure surge phenomenon, whose effect has caught attention but remains controversial in the medical literature.

Keywords: B-spline; Censored quantile regression; Functional regression; Generalized approximate cross-validation; Time-varying effect.

1 Introduction

Censored quantile regression (CQR) has been studied to differentiate covariate effects at different quantiles of survival times (Ying et al. 1995, Lindgren 1997, Koenker & Geling 2001, Bang & Tsiatis 2002, Chernozhukov & Hong 2002, Portnoy 2003, Peng & Huang 2008, Wang & Wang 2009). For example, Portnoy (2003) developed a recursively reweighed estimation procedure based on the principle of self-consistency for the Kaplan–Meier estimator; Wang & Wang (2009) developed an estimation algorithm based on the redistribution-of-mass idea; Peng & Huang (2008) proposed a martingale-based estimation procedure, which automatically accommodates the monotone structure of the parameters of interest, and in turn naturally adapts to the quantile function estimation.

However, when the covariates are partially functional, and the time-varying effects are of interest, the existing methods are not readily applicable. This is because the unknown parameters are infinite dimensional, which violates the assumptions of the existing models. Our motivating application falls into this realm - where neurologists are interested in studying the functional relationship between ambulatory blood pressure trajectories (i.e. functions) and time to stroke recurrence for stroke patients. Stroke is one of the leading causes of death across the world, and hypertension is the primary risk factor for stroke. Hence, it is of great clinical importance to understand such functional relationship. This clinical question motivates us to propose a functional CQR model in this paper, which are useful for other similar survival analysis scenarios as well.

Mean-based regression methods have been extended to accommodate functional covariates. Qu et al. (2016) proposed a functional proportional hazards model, which however is insufficient to characterize the higher or lower quantiles of the survival times. Furthermore, the model relies upon the proportional hazards assumption, which may not be satisfied in practice. In particular, we applied the method in Qu et al. (2016) to analyze the motivating stroke blood pressure data. The estimated effect of the baseline blood pressure on the disease risk is -0.0106 with the 95% confidence interval $(-0.0168, -0.0045)$, implying that a higher blood pressure leads to a lower risk of stroke recurrence, which contradicts with physicians' understanding of the disease. The contradiction may arise due to the violation of the proportional hazards assumption. Based on physicians' classification of the blood

pressure patterns, we cluster the blood pressure curves into three groups and plot the logarithm of the cumulative hazard functions of the three groups in Figure 1. Clearly, the three cumulative hazard curves cross without clear separation, indicating that the proportional hazards model is not appropriate for the data. Other than modeling the hazard function, the semiparametric accelerated failure time (AFT) model can also be used to describe the mean structure of the survival times without imposing an error distribution (Prentice 1978, Buckley & James 1979, Louis 1981, Wei & Gail 1983, Wei et al. 1990). Besides the lack of ability to characterize the survival quantiles, the AFT model generally requires the errors to be identically distributed, and independent of covariates (Peng & Huang 2008).

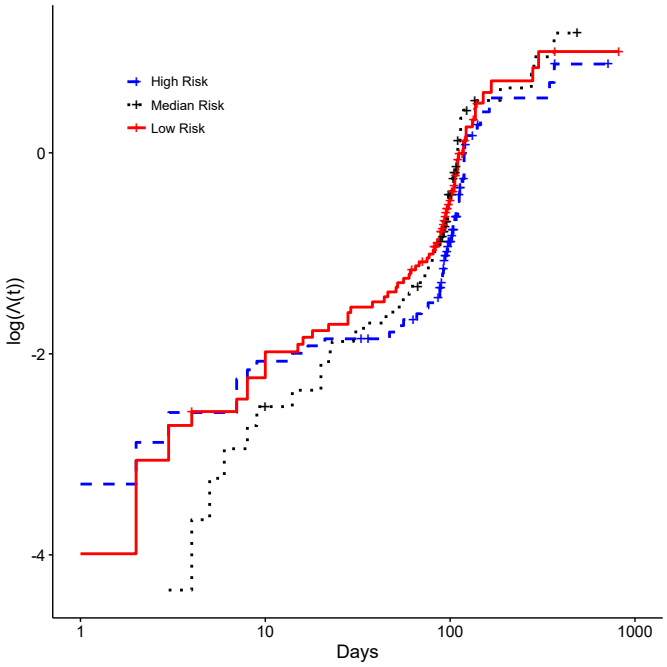

Figure 1: The logarithm of the three cumulative hazard functions calculated by the Kaplan–Meier method.

Functional quantile regression has been studied under the non-censoring framework (Cardot et al. 2005, Ferraty et al. 2005, Chen & Pouzo 2012, Kato 2012). For the censoring case, we propose a partially functional quantile regression model to handle a vector of covariates and a functional covariate jointly. Let T denote the survival time, $\mathbf{X}$ be the

vector of covariates, and $Z(s)$ be the functional covariate. In practice, $\mathbf{X}$ can be the baseline measurements, while $Z(s)$ can be the follow-up measurements, for example, the continuously monitored ambulatory blood pressure levels in our motivating application. We then assume that T, $\mathbf{X}$, and $Z(s)$ follow a linear location-scale design (Kato 2012). To fix ideas, when the survival time T is fully observed, the functional linear regression model can be written as

$$\log(T) = \mathbf{b}_1^{\mathrm{T}}\mathbf{X} + \int_0^1 \psi_1(s)Z(s)ds + \epsilon\left\{\mathbf{b}_2^{\mathrm{T}}\mathbf{X} + \int_0^1 \psi_2(s)Z(s)ds\right\},$$

where ϵ is a zero-mean random error independent of $\mathbf{X}$ and $Z(s)$.

The location-scale design not only defines the dependency of the quantile functions on the covariates, but also induces a quantile model with parameters varying with the quantiles. Furthermore, the linear structure is crucial for the computational feasibility in reality. Under this specification, the quantile function $Q_T(\tau|\mathbf{X}, Z; \boldsymbol{\beta}, \alpha) \equiv \inf_t\{\Pr(T \leq t|\mathbf{X}, Z) > \tau\}$ can be modelled as

$$Q_T(\tau|\mathbf{X}, Z; \boldsymbol{\beta}, \alpha) = \exp\left\{\boldsymbol{\beta}^{\mathrm{T}}(\tau)\mathbf{X} + \int_0^1 \alpha(s,\tau)Z(s)ds\right\}, \tag{1}$$

where $\boldsymbol{\beta}(\tau) = \mathbf{b}_1 + \mathbf{b}_2 Q_\epsilon(\tau)$ is the p-dimensional coefficient vector, and $\alpha(s,\tau) = \psi_1(s) + \psi_2(s)Q_\epsilon(\tau)$ is a function on $L_2[0,1]$, with $Q_\epsilon(\tau)$ being the quantile function for ϵ.

Simultaneous estimation of $\alpha(s,\tau)$ and $\boldsymbol{\beta}(\tau)$ is not an easy task, because $\alpha(s,\tau)$ is an infinite dimensional parameter. Simply treating $\alpha(s,\tau)$ as a multivariate covariate yields an extremely large parameter space; as such the parameters cannot be estimated theoretically with proper convergence rates and practically with reasonable computational power. To alleviate this curse of dimensionality, Kato (2012) adopted the functional principle component procedure of Yao et al. (2005). The functional expansion method is appealing, because through projection, the estimation reduces to a linear problem, which operationally falls in the convex optimization paradigm and can be implemented efficiently by existing simplex or interior point methods (Peng & Huang 2008).

However, the approach in Kato (2012) is not readily applicable to the censored quantile regression, especially when the censoring mechanism is random and the censoring times are conditionally independent of the survival times given covariates. To address these

issues, we develop a functional expansion approach for CQR, namely the functional censored quantile regression (FCQR), which adopts the martingale estimation procedure in Peng & Huang (2008). We approximate the unknown functionals using B-spline bases. By properly controlling the number of B-spline knots, we show that the B-spline approach yields the parametric and non-parametric convergence rates for the finite and functional estimates, respectively. Operationally, our estimation procedure is similar to that of Peng & Huang (2008); however, the development of the theoretical properties of our estimators is highly non-trivial because the number of unknown parameters increase with sample size. One critical issue of the B-spline functional approximation is to select the number of knots, especially with censored data. The usual cross-validation method based on martingale residuals may not work well due to changes in the censoring rates between training and validation datasets. We propose a generalized approximate cross-validation (GACV) method for selection of the B-spline knots. The GACV approach is computationally efficient and, more importantly, it avoids random partitioning of the samples, which may lead to invalid results due to inconsistency in the censoring rates and sample sizes between the subsamples and the original samples.

The rest of the article is organized as follows. In Section 2, we develop a class of semiparametric methods to estimate both $\boldsymbol{\beta}(\tau)$ and $\alpha(s,\tau)$. We derive the asymptotic properties of the estimators in Section 3, and introduce the generalized approximate cross-validation method in Section 4. We illustrate the finite sample performance of the proposed method through simulation studies in Section 5. The analysis of the blood pressure and stroke recurrence data is carried out in Section 6. We conclude the article in Section 7. The technical proofs are presented in the Appendix.

2 Model and Estimation

Let T_i be the survival time, C_i be the censoring time, $Y_i = T_i \wedge C_i$, $\Delta_i = I(T_i \leq C_i)$. Furthermore, we define $N_i(t) = I(T_i \leq t, \Delta_i = 1)$ to be the counting process for $i = 1, \ldots, n$, and $\Lambda(t)$ to be the cumulative hazard function. Moreover, let $Z_i(t)$ be a left-continuous process, and $\mathbf{X}_i$ be the baseline covariate vector. Without loss of generality, we assume

$t \in [0,1]$, a compact interval. For example, every subject in the stroke blood pressure data has continuous systolic blood pressure (SBP) measurements over a 24-hour period.

Consider Model (1), where $\boldsymbol{\beta}(\tau)$ and $\alpha(s,\tau)$ are unknown parameters, and define the following martingale

$$M_i(t) = N_i(t) - \Lambda(t \wedge Y_i), \tag{2}$$

which, according to Fleming & Harrington (2011), satisfies

$$E\left[M_i\left\{Q_T(\tau|\mathbf{X}_i, Z_i; \boldsymbol{\beta}_0, \alpha_0)\right\}\right] = 0,$$

where $\boldsymbol{\beta}_0(\tau)$ and $\alpha_0(\cdot,\tau)$ are the true parameters of interest. This mean zero property allows us to construct the estimating equation based on the sample version of the conditional mean.

In the martingale $M_i(t)$, although $N_i(t)$ is observed, the hazard $\Lambda(t)$ is not fully specified at every time point. However, the hazard function has the close form as

$$\Lambda\left\{Q_T(\tau|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha) \wedge Y_i\right\} \;=\; \int_0^{\tau} I\left[Y_i \geq Q_T(u|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha)\right] dH(u),$$

at the time of $Q_T(\tau|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha)$ with $H(u) = -\log(1-u)$ (Peng & Huang 2008). This quantity can be well approximated by

$$\widetilde{\Lambda}\left[Q_T(\tau_j|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha) \wedge Y_i\right] = \sum_{k=0}^{j-1} I\left[Y_i \geq Q_T(\tau_k|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha)\right]\left\{H(\tau_{k+1}) - H(\tau_k)\right\},$$

where $\tau_k \in \mathcal{T}_L$ with $\mathcal{T}_L = \{0 = \tau_0 < \tau_1 < \ldots < \tau_L < \tau_U\}$ for $\tau_U < 1$. This hazard approximation allows us to perform the estimation based on the mean zero property of the martingales at specified time points (Peng & Huang 2008).

Assume $\alpha(\cdot,u) \in C^q[0,1]$, a continuous differentiable function with order q. We approximate $\alpha(\cdot,u)$ by B-splines in the form of $\mathbf{B}_r(\cdot)^{\mathrm{T}}\boldsymbol{\gamma}(u)$ (De Boor 1978), where $\mathbf{B}_r(\cdot)$ is a known base function of time, and $\boldsymbol{\gamma}(u)$ is the unknown coefficient at quantile u. De Boor (1978) showed that there exists $\boldsymbol{\gamma}_0$ such that $\sup_{t\in[0,1]} |\mathbf{B}_r^{\mathrm{T}}(t)\boldsymbol{\gamma}_0(\tau) - \alpha_0(t,\tau)| = O_p(h_b^q)$ for $\alpha_0(t,\tau) \in C^q([0,1])$. Let $\mathbf{W}_i = \int_0^1 \mathbf{B}_r(s)Z_i(s)ds$ be the d-dimensional transformed covariate vector, $\widehat{\boldsymbol{\beta}}(\tau)$ and $\widehat{\boldsymbol{\gamma}}(\tau)$ be the consistent estimators for $\boldsymbol{\beta}(\tau)$ and $\boldsymbol{\gamma}(\tau)$, respectively. Since both $\mathbf{B}_r$ and Z_i are given, $\mathbf{W}_i$ is pre-defined and can be treated as observed.

Let

$$Q(\tau|\mathbf{X}_i, \mathbf{W}_i; \boldsymbol{\beta}, \boldsymbol{\gamma}) = \exp\{\boldsymbol{\beta}^{\mathrm{T}}(\tau)\mathbf{X}_i + \boldsymbol{\gamma}^{\mathrm{T}}(\tau)\mathbf{W}_i\}.$$

Combining the approximations for Λ and $\alpha(\cdot, u)$, we can estimate $\{\boldsymbol{\beta}^{\mathrm{T}}(\tau_j), \boldsymbol{\gamma}^{\mathrm{T}}(\tau_j)\}^{\mathrm{T}}$ sequentially, for $j = 1, \ldots, L$, via solving

$$\begin{aligned} \mathbf{0} \;=\; & \sum_{i=1}^{n} (\mathbf{X}_i^{\mathrm{T}}, \mathbf{W}_i^{\mathrm{T}}) \Bigg(N_i\{Q(\tau_j|\mathbf{X}_i, \mathbf{W}_i; \boldsymbol{\beta}, \boldsymbol{\gamma})\} \\ & - \sum_{k=0}^{j-1} I\{Y_i \geq Q(\tau_k|\mathbf{X}_i, \mathbf{W}_i; \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}})\}\{H(\tau_{k+1}) - H(\tau_k)\} \Bigg). \end{aligned} \tag{3}$$

In (3), because $Q_T(0|\mathbf{X}, Z; \boldsymbol{\beta}_0, \alpha_0) = 0$, we always set $Q(0|\mathbf{X}, \mathbf{W}; \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}}) = 0$. Operationally the procedure reduces to the censored linear quantile estimation described in Peng & Huang (2008). It is worth mentioning that solving (3) is equivalent to minimizing

$$\begin{aligned} & -\sum_{i=1}^{n} [\log(Y_i) - \log\{Q(\tau_j|\mathbf{X}_i, \mathbf{W}_i; \boldsymbol{\beta}, \boldsymbol{\gamma})\}] \Bigg(N_i\{Q(\tau_j|\mathbf{X}_i, \mathbf{W}_i; \boldsymbol{\beta}, \boldsymbol{\gamma})\} \\ & - \sum_{k=0}^{j-1} I[Y_i \geq Q(\tau_k|\mathbf{X}_i, \mathbf{W}_i; \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}})]\{H(\tau_{k+1}) - H(\tau_k)\} \Bigg), \end{aligned} \tag{4}$$

with respect to $\boldsymbol{\beta}(\tau_j)$ and $\boldsymbol{\gamma}(\tau_j)$ (Koenker 2008). The minimization algorithm has been implemented in the *R quantreg* package (Koenker 2012) and can be readily applied to estimate $\boldsymbol{\beta}(\tau)$ and $\boldsymbol{\gamma}(\tau)$ simultaneously.

The reasons for choosing the B-spline approximation are two-fold. First, under complete data settings, the B-spline based quantile function has the desirable finite-sample property that the quantile curves divide the sample by an approximate ratio of $\tau/(1-\tau)$ (He & Shi 1994). Second, we choose the B-spline bases instead of the eigen-bases used in Kato (2012) to incorporate heterogeneity in the functional covariates. In the stroke data, the blood pressure curves are not identically distributed across the patients; hence the covariance estimator is not valid for constructing unified eigen-bases across the samples.

3 Asymptotics

To derive asymptotic properties, we note the following technical challenge: as the dimensions for $\mathbf{W}_i$ and $\boldsymbol{\gamma}(\tau)$ grow with the sample size, the usual law of large numbers and central limit theorem are not applicable. To offer more insights, the score function on the right-hand side of (3) is not of order $O_p(1)$ because otherwise its L_2 norm is infinity as the dimension of $\mathbf{W}_i$ grows. Hence, properly controlling the number of knots is crucial.

For an order-r spline, let N be the number of interior knots, which yields the dimension for $\boldsymbol{\gamma}(\tau)$ to be $d = N + r$. Define $\|\cdot\|$ and $\|\cdot\|_\infty$ to be the L_2 and L_∞ vector norms, respectively. We select the knots uniformly in the interval of $[0, 1]$ with the distance $h_b = 1/N$. Furthermore, to control the order of convergence, we restrict the growing rate as $N = o(n)$. More detailed conditions are listed in Condition (A2) in Appendix A. We establish the estimation consistency in Theorem 1, the asymptotical normality in Theorem 2, and the convergence of the functional estimator in Corollary 1. The detailed derivations are all presented in Appendix A.

Theorem 1. *Assume that Conditions (A1)–(A6) hold, and $\widehat{\boldsymbol{\beta}}(\tau_j)$ and $\widehat{\boldsymbol{\gamma}}(\tau_j)$ satisfy*

$$\begin{aligned} 0 \;=\; & \sum_{i=1}^{n} (\mathbf{X}_i^{\mathrm{T}}, \mathbf{W}_i^{\mathrm{T}})^{\mathrm{T}} \Bigg(N_i[Q(\tau_j|\mathbf{X}_i, \mathbf{W}_i; \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}})] \\ & - \sum_{k=0}^{j-1} I[Y_i \geq Q(\tau_k|\mathbf{X}_i, \mathbf{W}_i; \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}})]\{H(\tau_{k+1}) - H(\tau_k)\} \Bigg). \end{aligned}$$

Then,

$$\|\{\widehat{\boldsymbol{\beta}}^{\mathrm{T}}(\tau_j), \widehat{\boldsymbol{\gamma}}^{\mathrm{T}}(\tau_j)\}^{\mathrm{T}} - \{\boldsymbol{\beta}_0^{\mathrm{T}}(\tau_j), \boldsymbol{\gamma}_0^{\mathrm{T}}(\tau_j)\}^{\mathrm{T}}\|_\infty = o_p(1),$$

and

$$\|\{\widehat{\boldsymbol{\beta}}^{\mathrm{T}}(\tau_j), \widehat{\boldsymbol{\gamma}}^{\mathrm{T}}(\tau_j)\}^{\mathrm{T}} - \{\boldsymbol{\beta}_0^{\mathrm{T}}(\tau_j), \boldsymbol{\gamma}_0^{\mathrm{T}}(\tau_j)\}^{\mathrm{T}}\| = o_p(1).$$

Furthermore, we define $F(t|Z) = \Pr(T \leq t|\mathbf{X}, Z)$, $\bar{F}(t|\mathbf{X}, Z) = 1 - F(t|\mathbf{X}, Z)$, $\widetilde{F}(t|\mathbf{X}, Z) = \Pr(Y \leq t, \Delta = 1|\mathbf{X}, Z)$, $f(t|\mathbf{X}, Z) = dF(t|\mathbf{X}, Z)/dt$, $\bar{f}(t|\mathbf{X}, Z) = d\bar{F}(t|\mathbf{X}, Z)/dt$, and $\widetilde{f}(t|\mathbf{X}, Z) = d\widetilde{F}(t|\mathbf{X}, Z)/dt$. In addition, we define the population score functions as

$$\mathbf{S}_{\boldsymbol{\beta}}(Y_i, \mathbf{X}_i, Z_i; \tau, \boldsymbol{\beta}, \alpha)$$

$$
\begin{aligned}
&= \mathbf{X}_i \left(N_i \left[Q_T(\tau|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha) \right] - \int_0^\tau I \left[Y_i \geq Q_T(u|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha) \right] dH(u) \right), \\
&\quad \mathbf{S}_{\boldsymbol{\gamma}}(Y_i, \mathbf{X}_i, Z_i; \tau, \boldsymbol{\beta}, \alpha) \\
&= \mathbf{W}_i \left(N_i \left\{ Q_T(\tau|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha) \right\} - \int_0^\tau I \left[Y_i \geq Q_T(u|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha) \right] dH(u) \right),
\end{aligned}
$$

and the corresponding Hessian matrix as

$$
\begin{aligned}
\mathbf{M}\{\boldsymbol{\beta}(\tau), \alpha(\cdot, \tau)\} &= E \Big[\left\{ (\mathbf{X}_i^{\mathrm{T}}, \mathbf{W}_i^{\mathrm{T}})^{\mathrm{T}} \right\}^{\otimes 2} \widetilde{f} \left\{ Q_T(\tau|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha) | \mathbf{X}_i, Z_i \right\} \\
&\quad \times Q_T(\tau|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha) \Big].
\end{aligned}
$$

Furthermore, we define the variance and covariance components as

$$
\begin{aligned}
\mathbf{J}\{\boldsymbol{\beta}(\tau), \alpha(\cdot, \tau)\} &= E \Big[\left\{ (\mathbf{X}_i^{\mathrm{T}}, \mathbf{W}_i^{\mathrm{T}})^{\mathrm{T}} \right\}^{\otimes 2} \bar{f} \left\{ Q_T(\tau|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha) | \mathbf{X}_i, Z_i \right\} \\
&\quad \times \exp \left\{ Q_T(\tau|\mathbf{X}_i, Z_i; \boldsymbol{\beta}, \alpha) \right\} \Big].
\end{aligned}
$$

Let $\prod$ be the product integral (Gill & Johansen 1990, Andersen et al. 2012) and define

$$
\boldsymbol{\phi} \left\{ \mathbf{G}(\tau), \mathbf{A}(\tau), \mathbf{B}(\tau) \right\} \equiv \int_0^\tau \prod\nolimits_{u \in (s,\tau]} [\mathbf{I}_q + \mathbf{A}(u)\mathbf{B}(u)^{-1} dH(u)] d\mathbf{G}(s),
$$

for arbitrary matrix valued functions $\mathbf{A}$ and $\mathbf{B}$.

In the following theorem, we show that $\widehat{\boldsymbol{\beta}}(\tau)$ and $\widehat{\boldsymbol{\gamma}}(\tau)$ achieve parametric and non-parametric convergent rates, respectively.

Theorem 2. *Assume that Conditions (A1)–(A6) hold, and $a_n = o_p(n^{-1/2})$, then*

$$
\sup_{\tau \in [0, \tau_U]} \|\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\| = O_p(n^{-1/2}), \text{ and } \sup_{\tau \in [0, \tau_U]} \|\widehat{\boldsymbol{\gamma}}(\tau) - \boldsymbol{\gamma}_0(\tau)\| = O_p(n^{-1/2} h_b^{-1/2}).
$$

For each τ,

$$
\begin{aligned}
&n^{1/2} [\{\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\}^{\mathrm{T}}, \{\widehat{\boldsymbol{\gamma}}(\tau) - \boldsymbol{\gamma}_0(\tau)\}^{\mathrm{T}}]^{\mathrm{T}} \\
= \ &\mathbf{M}^{-1}\{\boldsymbol{\beta}_0(\tau), \alpha_0(\cdot, \tau)\} \cdot \boldsymbol{\phi}[\mathbf{G}(\tau), \mathbf{J}\{\boldsymbol{\beta}_0(\tau), \alpha_0(\cdot, \tau)\}, \mathbf{M}\{\boldsymbol{\beta}_0(\tau), \alpha_0(\cdot, \tau)\}] \cdot \{1 + o_p(1)\},
\end{aligned}
$$

where $\mathbf{G} = (\mathbf{G}_1, h_b^{1/2} \mathbf{G}_2)$, $\mathbf{G}_1(\cdot)$ is a vector of zero-mean Gaussian processes with covariance

$$
\boldsymbol{\Sigma}_1(s, t) = E\{\mathbf{S}_{\boldsymbol{\beta}}(Y_i, \mathbf{W}_i; s, \boldsymbol{\beta}_0, \alpha_0) \mathbf{S}_{\boldsymbol{\beta}}(Y_i, \mathbf{W}_i; t, \boldsymbol{\beta}_0, \alpha_0)^{\mathrm{T}}\},
$$

$\mathbf{G}_2(\cdot)$ *is a vector of zero-mean Gaussian processes with covariance*

$$\mathbf{\Sigma}_2(s,t) = h_b^{-1} E\{\mathbf{S}_{\boldsymbol{\gamma}}(Y_i, \mathbf{W}_i; s, \boldsymbol{\beta}_0, \alpha_0)\mathbf{S}_{\boldsymbol{\gamma}}^{\mathrm{T}}(Y_i, \mathbf{W}_i; t, \boldsymbol{\beta}_0, \alpha_0)\},$$

and furthermore,

$$\mathbf{\Sigma}_{12}(s,t) = h_b^{-1/2} E\{\mathbf{S}_{\boldsymbol{\beta}}(Y_i, \mathbf{W}_i; s, \boldsymbol{\beta}_0, \alpha_0)\mathbf{S}_{\boldsymbol{\gamma}}^{\mathrm{T}}(Y_i, \mathbf{W}_i; t, \boldsymbol{\beta}_0, \alpha_0)\}.$$

The variances of $\widehat{\boldsymbol{\beta}}(\tau)$ and $\widehat{\boldsymbol{\gamma}}(\tau)$ comprise of both integration and product integral. To alleviate computational instability, we use the bootstrap method for variance estimation.

Given the convergency of $\widehat{\boldsymbol{\gamma}}(\tau)$, the estimation error for $\alpha_0(s,\tau)$ is a combination of estimation bias and variance as shown in the following Corollary 1.

Corollary 1. *Assume that Conditions (A1)–(A6) hold, and* $a_n = o_p(n^{-1/2})$*, then*

$$\sup_{s\in[0,1]} |\mathbf{B}_r^{\mathrm{T}}(s)\widehat{\boldsymbol{\gamma}}(\tau) - \alpha_0(s,\tau)| = O_p(n^{-1/2}h_b^{-1/2} + h_b^q).$$

Furthermore, if $n^{1/2}N^{-1/2-q} \to 0$ *as* $n \to \infty$*, then for each* $s \in [0,1]$*,*

$$\begin{aligned} & n^{1/2}h_b^{1/2}\{\mathbf{B}_r^{\mathrm{T}}(s)\widehat{\boldsymbol{\gamma}}(\tau) - \alpha_0(s,\tau)\} \\ = \; & \mathbf{B}_r^{\mathrm{T}}(s)(\mathbf{0}_{d\times p}, h_b^{1/2}\mathbf{I}_{d\times d}) \cdot \mathbf{M}^{-1}\{\boldsymbol{\beta}_0(\tau), \alpha_0(\cdot,\tau)\} \cdot \boldsymbol{\phi}[\mathbf{G}(\tau), \mathbf{J}\{\boldsymbol{\beta}_0(\tau), \alpha_0(\cdot,\tau)\}, \mathbf{M}\{\boldsymbol{\beta}_0(\tau), \alpha_0(\cdot,\tau)\}] \\ & \times\{1 + o_p(1)\}. \end{aligned}$$

4 Generalized Approximate Cross-validation for Knot Selection

We consider the important problem of knot selection in a data-driven fashion, and study the selection procedure's theoretical property. For simplicity, the knots are uniformly selected on $[0,1]$ with the maximum number $s_n = O_p\{n^{1/(2q+1)}\}$. Then we extend the GACV procedure to select the number of knots, which is asymptotically equivalent to the leave-one-out cross-validation (Yuan 2006).

More specifically, let $\mathcal{S} = (j_1, \ldots, j_{N_S})$ be the indices of the interior knots with $d_S = N_S + r$ bases. Let $\mathcal{M}$ be the set containing all $\mathcal{S}$ such that $|\mathcal{S}| \leq s_n$, where $|\cdot|$ represent the cardinality of a set. We assume that N_S satisfies Condition (B1) in Appendix B for all $\mathcal{S} \in \mathcal{M}$. We modify the notation to reflect the dependence on the set $\mathcal{S}$. Furthermore, let $\boldsymbol{\gamma}_S$ be the corresponding B-spline coefficient when $\mathcal{S}$ is chosen and $\boldsymbol{\gamma}_{S0}(\tau)$ be the coefficient that satisfies

$$\sup_{t\in[0,1]} |\mathbf{B}_{rS}(t)^{\mathrm{T}}\boldsymbol{\gamma}_{S0}(\tau) - \alpha_0(t,\tau)| = O_p(N_S^{-q}),$$

and $\mathbf{W}_{Si} = \int_0^1 \mathbf{B}_{rS}(u)Z_i(u)du$.

Given that the estimation procedure is essentially to minimize (4), we define the objective function for GACV as

$$\begin{aligned}
&\mathrm{GACV}(\tau_j, d_S, \boldsymbol{\beta}, \boldsymbol{\gamma}_S) \\
= \; & -\sum_{i=1}^{n} \left[\log(Y_i) - \{\boldsymbol{\beta}(\tau_j)^{\mathrm{T}}\mathbf{X}_i + \boldsymbol{\gamma}_S(\tau_j)^{\mathrm{T}}\mathbf{W}_{Si}\}\right] \Bigg(N_i[\exp\{\boldsymbol{\beta}(\tau_j)^{\mathrm{T}}\mathbf{X}_i + \boldsymbol{\gamma}_S(\tau_j)^{\mathrm{T}}\mathbf{W}_{Si}\}] \\
& -\sum_{k=0}^{j-1} I[Y_i \geq \exp\{\widehat{\boldsymbol{\beta}}_S(\tau_k)^{\mathrm{T}}\mathbf{X}_i + \widehat{\boldsymbol{\gamma}}_S(\tau_k)^{\mathrm{T}}\mathbf{W}_{Si}\}]\{H(\tau_{k+1}) - H(\tau_k)\} \Bigg) / \{n - C_n(d_S + p)\}.
\end{aligned}$$

with $C_n = O_p[\{\log(n)\}^{\omega}]$, $0 \leq \omega \leq 1$, where $\omega = 1/2$ generally works well in our numerical studies. The GACV with $\omega = 0$ and $\omega = 1$ is asymptotically equivalent to the Akaike information criterion and Bayesian information criterion, respectively (Kato 2012).

We select d_S by minimizing the integrated GACV:

$$\mathrm{IGACV} = \sum_{j=1}^{L} \mathrm{GACV}(\tau_j, d_S, \widehat{\boldsymbol{\beta}}_S, \widehat{\boldsymbol{\gamma}}_S)/L,$$

where $\widehat{\boldsymbol{\beta}}_S$ and $\widehat{\boldsymbol{\gamma}}_S$ are the estimators solving (3) when $d = d_S$ and L is the number of evaluated quantiles.

Below we study the selection consistency of the IGACV criterion. Let $\mathcal{F}_0(\tau_j-)$ be the sigma-field generated by

$$\left[Y_i, \Delta_i, Q_T(u|\mathbf{X}_i, Z_i; \boldsymbol{\beta}_0, \alpha_0), \mathbf{W}_i, Z_i, 0 \leq u < \tau_j, i = 1, \ldots, n, \mathcal{S} \in \mathcal{M}\right].$$

Consider the problem of minimizing, over $\mathcal{S} \in \mathcal{M}$, the following criterion:

$$E\Bigg\{ -[\log(Y_i) - \log\{Q(\tau_j|\mathbf{X}_i, \mathbf{W}_{Si}; \boldsymbol{\beta}_S, \boldsymbol{\gamma}_S)\}] \\ \times \left(N_i\{Q(\tau_j|\mathbf{X}_i, \mathbf{W}_{Si}; \boldsymbol{\beta}_S, \boldsymbol{\gamma}_S)\} - \int_0^{\tau_j} I\left[Y_i \geq Q_T(u|\mathbf{X}_i, Z_i; \boldsymbol{\beta}_0, \alpha_0)\right] du\} \right) |\mathcal{F}_0(\tau_j-) \Bigg\}.$$

Finally, let $\mathcal{S}^*$ be the set of the interior knots that the corresponding $\boldsymbol{\gamma}_{S^*0}(\tau_j)$, together with $\boldsymbol{\beta}_0(\tau_j)$, is the unique minimizer of the above criterion.

Let $\mathcal{M}^O = \{\mathcal{S} \in \mathcal{M} : \mathcal{S} \supset \mathcal{S}^*, \mathcal{S} \neq \mathcal{S}^*\}$, and $\mathcal{M}^U = \{\mathcal{S} \in \mathcal{M} : \mathcal{S} \not\supset \mathcal{S}^*\}$.

Theorem 3. *Assuming that Conditions (B1)–(B3) hold, then*

$$\lim_{n\to\infty} \Pr\left\{ \min_{\mathcal{S}\in\mathcal{M}^O,\, S\in\mathcal{S}} \mathrm{GACV}(\tau_j, d_S, \widehat{\boldsymbol{\beta}}_S, \widehat{\boldsymbol{\gamma}}_S) > \mathrm{GACV}(\tau_j, d_{S^*}, \boldsymbol{\beta}_0, \boldsymbol{\gamma}_{S^*0}) \right\} = 1,$$

and

$$\lim_{n\to\infty} \Pr\left\{ \min_{\mathcal{S}\in\mathcal{M}^U,\, S\in\mathcal{S}} \mathrm{GACV}(\tau_j, d_S, \widehat{\boldsymbol{\beta}}_S, \widehat{\boldsymbol{\gamma}}_S) > \mathrm{GACV}(\tau_j, d_{S^*}, \boldsymbol{\beta}_0, \boldsymbol{\gamma}_{S^*0}) \right\} = 1.$$

The above Theorem 3 suggests the selection consistency of the GACV procedure: both over-selection and under-selection yield a smaller GACV than the optimal one with probability tending to one. Furthermore, because the theorem holds for every τ_k, as long as $\boldsymbol{\beta}_0(\tau_k)$ and $\alpha_0(\cdot, \tau_k)$, $k < j$, are consistently estimated; and the grid on τ is sufficiently dense (see the proof of Theorem 3 in Appendix B), the IGACV procedure also guarantees the selection consistency asymptotically.

5 Simulation

5.1 Finite sample performance without IGACV procedure

We conduct extensive simulation studies to assess the finite-sample performance of the FCQR procedure. We generate the event time from the model,

$$\log(T_i) = b_1 X_{1i} + \int_0^1 Z_i(s)\psi(s)ds + \left\{ b_2 X_{2i} + \int_0^1 Z_i(s)ds \right\} \epsilon_i,$$

where X_{1i} is a standard normal variate, X_{2i} is a variate uniformly distributed on $[0, 1]$, and ϵ_i is a normal random error with mean zero and variance σ_ϵ^2. The functional covariate $Z_i(s)$ and $\psi(s)$ are given by

$$Z_i(s) = \left| \sum_{k=1}^{K} \zeta_k U_{ik} \phi_k(s) \right|, \quad \psi(s) = -4 + \sum_{k=2}^{K} 4(-1)^k k^{-2} [\sqrt{2} \cos\{(k-1)\pi s\}],$$

with $\zeta_k = (-1)^{k+1} k^{-v/2}$, $K = 50$, $v = 1, 1.5, 2, 2.5$, $\phi_1 = 1$ and $\phi_k = \sqrt{2}\cos\{(k-1)\pi s\}$ for $k > 1$. Furthermore, U_{ki} is uniformly distributed on $[-\sqrt{3}, \sqrt{3}]$, and the coefficients are set to be $b_1 = 1$ and $b_2 = 2$.

Let $Q_\epsilon(\tau)$ be the τth quantile of ϵ_i. The quantities in (1) can be written as $\mathbf{X}_i = (X_{1i}, X_{2i})^{\mathrm{T}}$, $\alpha(s, \tau) = \psi(s) + Q_\epsilon(\tau)$, and $\boldsymbol{\beta}(\tau) = \{b_1, b_2 Q_\epsilon(\tau)\}^{\mathrm{T}}$. We generate 1000 simulated datasets of sample size $n = 200$ and 500 for each value of $v = 1, 1.5, 2, 2.5$. We use cubic splines to approximate $\alpha(s, \tau)$, while setting the number of knots d as the smallest integer greater than $1.5n^{1/5}$. The censoring times are generated independently to yield desired censoring rates.

Tables 1–3 show the estimation results for censoring rates of 10%, 20% and 40%, respectively. The empirical confidence interval is defined as the 95% confidence intervals based on the quantiles of the bootstrap estimate of the sampling distribution (Jiang & Haneuse 2017). For the bootstrap confidence intervals, we take 200 bootstrap samples to obtain the bootstrap standard error and then use normal approximation. Both the biases and standard errors of our estimators are small and the 95% coverage probabilities are close to the nominal level. The estimation errors decrease as the sample size increases. The coverage probabilities are closer to the normal level for a lower censoring rate (10%) than those for a higher censoring rate (40%).

We also evaluate the performance of estimating $\alpha(s, \tau)$. To save space, we only show the results for estimating $\alpha(s, 0.5)$ with $v = 2$. (The results under other settings are similar.) Figure 2 displays the time-varying estimators, the empirical and bootstrap 95% confidence intervals for $\alpha(s, 0.5)$. The estimated curves are close to the true ones; the bootstrap and empirical confidence intervals nearly overlap across all the scenarios.

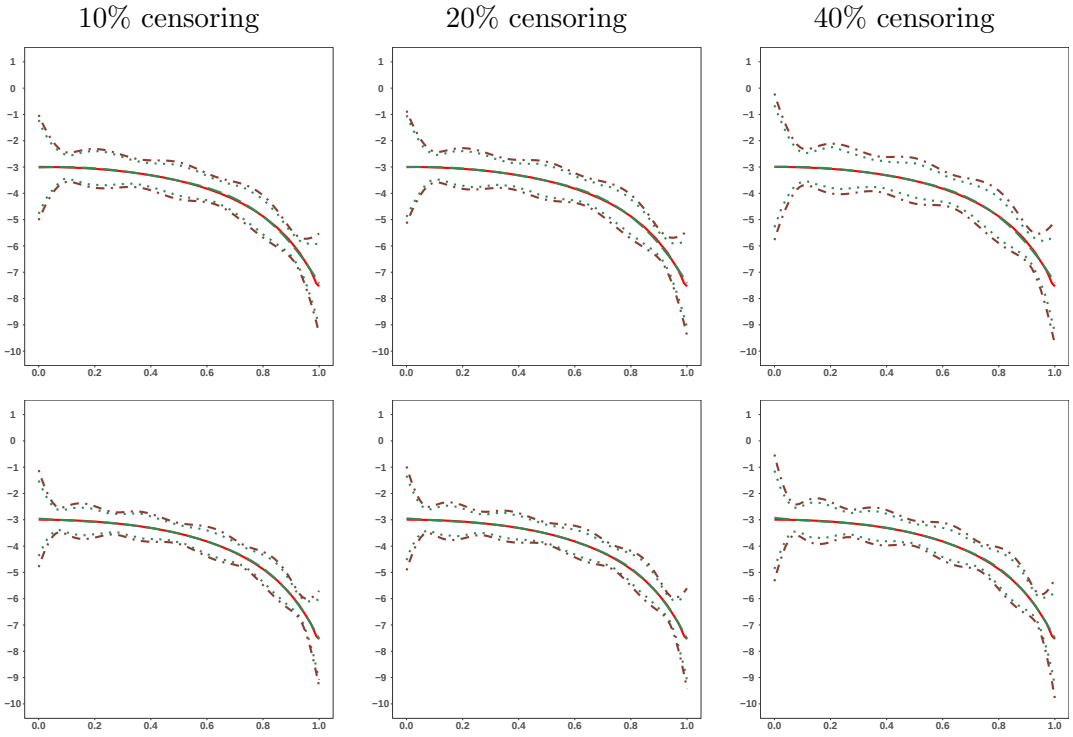

Figure 2: Estimated coefficient functions $\mathbf{B}_r^{\mathrm{T}}(s)\widehat{\boldsymbol{\gamma}}(0.5)$ with $\sigma_\epsilon = 0.2$, $v = 2$, and sample sizes 200 (top panels) and 500 (bottom panels), respectively. The red solid lines are the true functions, the green long-dashed lines represent the estimated mean functions, the green dotted are the empirical confidence intervals and the brown dot-dashed lines are the bootstrap confidence intervals.

5.2 Performance of IGACV for knot selection

To illustrate the performance of the IGACV procedure, we mimic the real data in the second set of simulations. In particular, we generate $n = 297$ survival times from the model

$$\log(T_i) = \int_0^1 Z_i(s)\psi(s)ds + \left\{ b_1 Z_{1i} + \int_0^1 Z_i(s)ds \right\} \epsilon_i,$$

where $Z_i(s) = \mathrm{SBP}_i(s) + U_i(s)$ and $Z_{1i} = \min(\mathrm{SBP}_i) + U_{1i}$ with $\mathrm{SBP}_i(s)$ corresponding to the systolic blood pressure (SBP) trajectory for the ith individual, and $U_i(s)$ and U_{1i} are

uniform random variables on $[-10, 10]$. We set $\psi(s) = 5\{\sin(\pi/2 + 13\pi s/4)\}$, $b_1 = 0.1$, and generate ϵ_i from a normal distribution with mean 0 and variance 0.04. Under this setting, we have $\boldsymbol{\beta}(\tau) = b_1 Q_\epsilon(\tau)$ and $\alpha(s, \tau) = \psi(s) + Q_\epsilon(\tau)$. The censoring times are generated to achieve censoring rates of 10%, 20% and 40%, respectively.

For the selection of the number of B-spline knots, we also consider minimizing the integrated Bayesian information criterion (IBIC) (Kato 2012, Lee et al. 2014). Table 4 presents the number of times a specific d_S is chosen when each criterion reaches the minimum value in 500 replications. We evaluate IGACV and IBIC over the choices of $d_S = 4, 5, 6, 7, 8$. In all the settings, IGACV chooses $d_S = 6$ as the optimal number of basis functions, while IBIC tends to select the boundary value $d_S = 4$, the minimal number of knots for the cubic spline. This comparison result suggests that IBIC is conservative in the knots selection.

We present the estimation results for $d_S = 6$ in Table 5, Figures 3 and 4, which show that both $\widehat{\boldsymbol{\beta}}_S(\tau)$ and $\mathbf{B}_{rS}^{\mathrm{T}}(s)\widehat{\boldsymbol{\gamma}}_S(\tau)$ are close to the true values, the coverage probabilities reach the nominal level, and the estimated confidence intervals approach the empirical counterparts under various censoring rates for τ between 0.3 and 0.7. Due to sparse event information during the initial follow-up time and insufficient samples at the end of the study, the variance estimation and coverage probabilities at $\tau = 0.2$ and 0.8 are less satisfactory. Hence, quantile estimation tends to be more reliable in the middle range of quantiles.

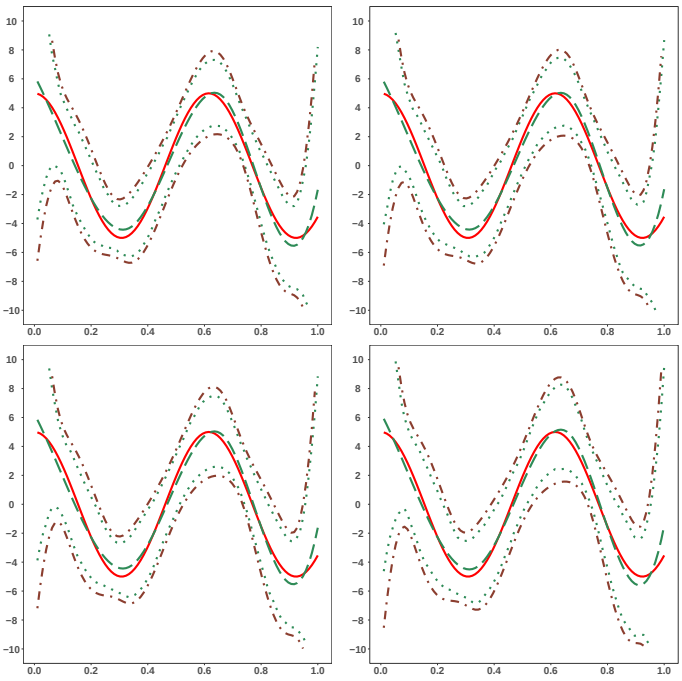

Figure 3: Estimated coefficient functions $\mathbf{B}_{rS}^{\mathrm{T}}(s)\widehat{\gamma}_S(0.5)$ with $d_S = 6$ under censoring rates 0% (top left), 10% (top right), 20% (bottom left), and 40% (bottom right), respectively. The red solid lines represent the true functions, the green long-dashed lines are the estimated mean functions, the green dotted lines are the empirical confidence intervals, and the brown dot-dashed lines are the bootstrap confidence intervals.

Furthermore, we compare our FCQR method with that of Kato (2012) based on the same simulated dataset under the non-censoring scenario. For the method of Kato (2012), we select the five functional bases that give the smallest integrated absolute error for the functional estimator. Figure 5 exhibits the functional estimators at $\tau = 0.3, 0.5, 0.7$ using the method of Kato (2012), which clearly shows under-smoothness. To obtain more reasonable estimators, we utilize the local regression method (Chapter 8 in Chambers et al. (1992)) with bandwidth 0.4 to obtain the smoothed fitted curves and the corresponding confidence intervals as shown in Figure 6. The bandwidth 0.4 is chosen so that the resulted curves are sufficiently smooth with the minimal loss of the integrated absolute error compared to the original fitted curves in Figure 5. Table 6 shows that the baseline effect

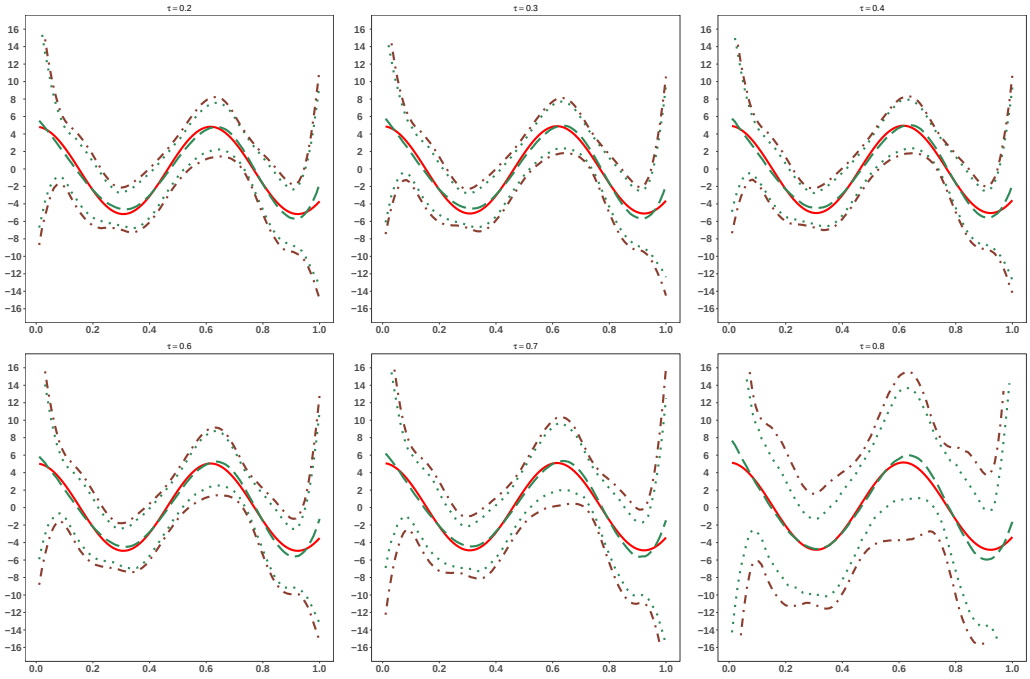

Figure 4: Estimated coefficient function $\mathbf{B}_{rS}^{\mathrm{T}}(s)\widehat{\gamma}_S(\tau)$ with $d_S = 6$ under a censoring rate of 40%. The red solid lines represent the true function, the green long-dashed lines are the estimated mean function, the green dotted lines are the empirical confidence intervals, and the brown dot-dashed lines are the bootstrap confidence intervals.

estimation of Kato (2012) is slightly better than FCQR under non-censoring scenarios in terms of bias. We also compute the integrated absolute errors of the functional estimators over 1000 simulations, which yield 1.38 and 0.41 at $\tau = 0.5$ for Kato's method and FCQR, respectively, demonstrating the superiority of our proposed method.

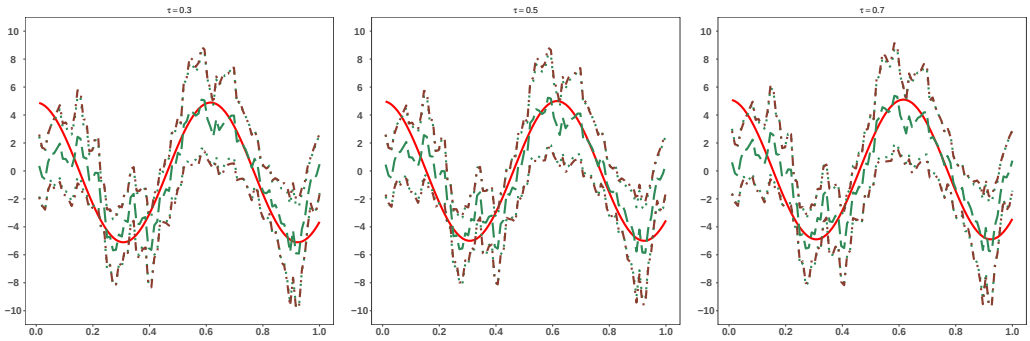

Figure 5: The functional coefficient estimators of Kato (2012) at $\tau = 0.3, 0.5, 0.7$ with five functional bases under no censoring. The red solid lines represent the true function, the green long-dashed lines are the estimated mean function, the green dotted lines are the empirical confidence intervals, and the brown dot-dashed lines are the bootstrap confidence intervals.

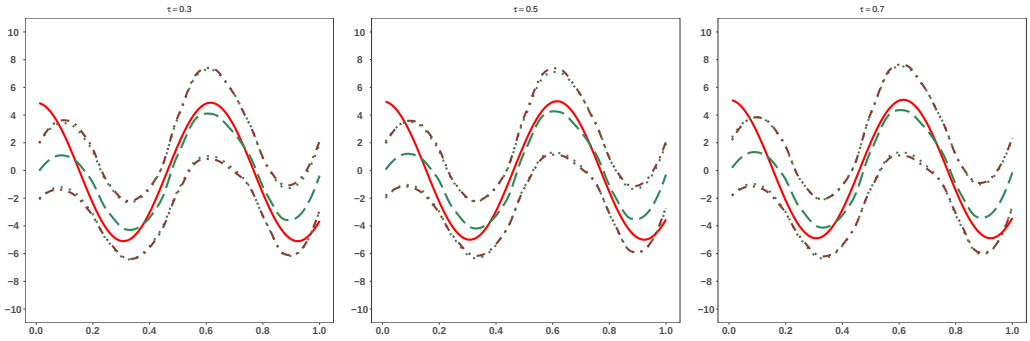

Figure 6: The functional coefficient estimators of Kato (2012) at $\tau = 0.3, 0.5, 0.7$ with five functional bases under no censoring after the local regression smoothing with bandwidth 0.4. The red solid lines represent the true function, the green long-dashed lines are the estimated mean function, the green dotted lines are the empirical confidence intervals, and the brown dot-dashed lines are the bootstrap confidence intervals.

6 Stroke Application: Time to Stroke Recurrence and Blood Pressure Trajectory

Our motivating application is from a nation wide hospital-based, prospective cohort study about ischemic stroke and transient ischemic attack patients, aiming at identifying func-

tional relationship between ambulatory blood pressure trajectories and clinical outcomes in stroke patients. Stroke is the top killer in China and the fifth leading killer in North America. Yearly, 15 million people suffer from stroke; 5 million of these people die, another 5 million people suffer from long-term disability. The data contain 24-hour ambulatory blood pressure measurements and interesting clinical outcomes about 297 stroke patients. The primary endpoint is the time to the composite stroke recurrent event, including death, disability, or vascular events. The censoring rate is 40%. Each patient's systolic blood pressure (SBP) is measured every 15 minutes starting from 19:00 for 24 hours.

For the ith patient, we consider his/her minimal SBP as the baseline covariate denoted as X_{1i}, and the SBP trajectory minus X_{1i} as the functional covariate, denoted as $Z_i(s)$. We standardize the logarithm of the survival time to stabilize the computation.

As shown in Table 7, we apply the IGACV and IBIC criteria to select d_S, respectively. The IGACV reaches the minimum value at $d_S = 7$, while IBIC does at $d_S = 4$. Based on $d_S = 7$, Table 8 shows that there are significant negative effects of Z_{1i} at all considered quantiles, which suggests that hypertension is an important risk factor for stroke.

Figure 7 displays the estimated $\alpha(s, \tau)$ and the corresponding 95% pointwise bootstrap confidence intervals at different values of τ. For the quantiles τ between 0.3 and 0.6, the magnitude of the effect starts to decrease around 4:00, becomes negatively significant around 8:00, and reaches the minimum at 10:30 on average. This trend coincides with the morning surge phenomenon that often occurs when patients wake up in the morning: the blood pressure increases quickly between 6:00 and 10:00. The results suggest that the early morning surge of the blood pressure level has a significant effect on the disease risk. Results for $\tau = 0.2,\ 0.7,\ 0.8$ are presented in Appendix C.

The estimates in the panels of Figure 7 have an obvious bimodal feature. We also comment on the first valley between 20:00 and 24:00 . During this time period, high blood pressure increases the risk for the medical outcome, and the effect is the strongest at the peak. This is consistent with the medical knowledge that high blood pressure at night increases the stroke risk.

For comparison, we also implement the method of Kato (2012) on the 174 uncensored observations. The IGACV and IBIC in Kato (2012) suggest using four functional bases to approximate the functional coefficients. Table 8 shows the estimated baseline coefficients,

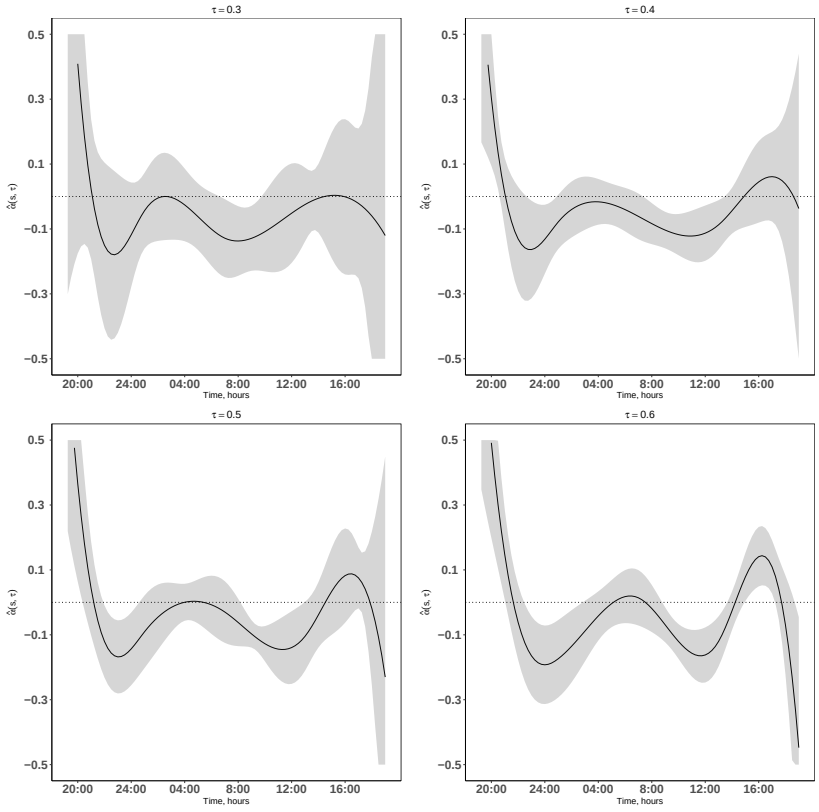

Figure 7: Estimated coefficient functions $\mathbf{B}_{rS}^{\mathrm{T}}(s)\widehat{\boldsymbol{\gamma}}_S(\tau)$ using 297 observations with $d_S = 7$. The solid lines represent $\mathbf{B}_{rS}^{\mathrm{T}}(s)\widehat{\boldsymbol{\gamma}}_S(\tau)$ and shaded areas are the corresponding 95% pointwise bootstrap confidence intervals for $\tau = 0.3, 0.4, 0.5$, and 0.6, respectively. The peak times are 8:00, 10:45, 11:15, 11:45 in the morning for $\tau = 0.3, 0.4, 0.5, 0.6$, respectively.

which are close to the estimates from FCQR. Moreover, the functional estimates of Kato (2012) in Figure 8 have similar patterns to those in Figure 7, but less smooth. Because the sample size is smaller, the estimators of Kato (2012) have wider confidence intervals, and in turn do not reach the significance level.

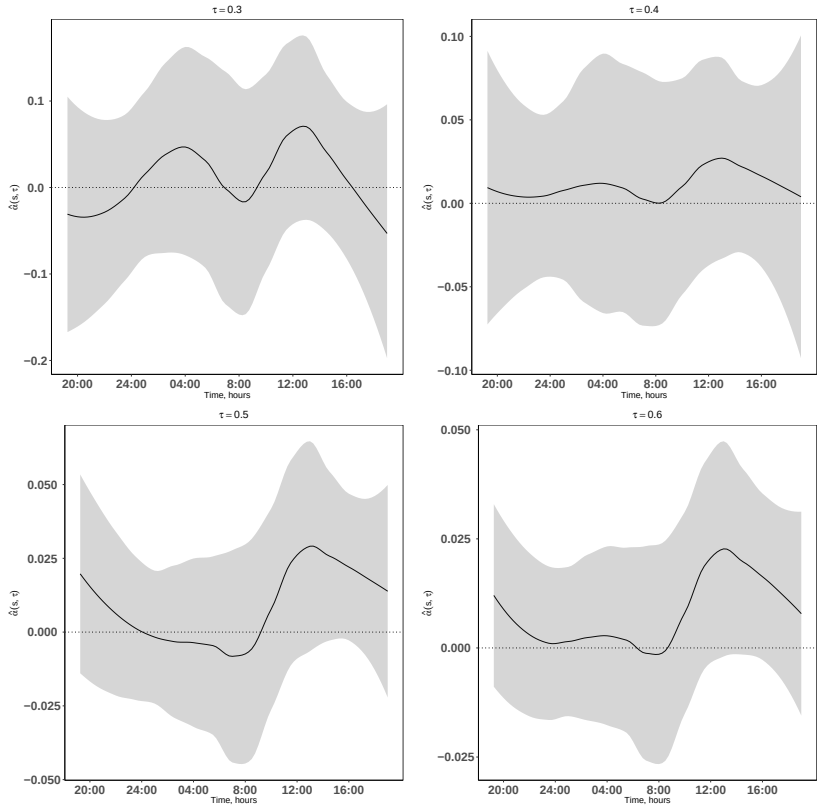

Figure 8: Functional coefficient estimates from the Kato (2012) method with four functional bases given 174 complete observations. The estimators are smoothed using the local regression method (Chambers et al. 1992) with bandwidth 0.4. The solid lines represent the estimated functional curves and shaded areas are the corresponding 95% pointwise bootstrap confidence intervals for $\tau = 0.3, 0.4, 0.5$, and 0.6, respectively.

To evaluate the adequacy of the FCQR, we follow Peng & Huang (2008) to define

$$K_n(\tau) = n^{-1/2} \sum_{i=1}^{n} q(X_{1i}) M_i \left\{ Q_T(\tau | \mathbf{X}_i, Z_i; \widehat{\boldsymbol{\beta}}, \widehat{\alpha}) \right\}$$

where M_i is defined in (2), We choose $q()$ as a quadratic function of baseline covariate, that is $q(X_{1i}) = \{(X_{1i} - 113.5)/16.8\}^2$, where 113.5 and 16.8 are the mean and standard

deviation of X_{1i}. When Model (1) is correct, $K_n(\tau)$ converges to a zero-mean Gaussian process. We utilize the resampling approach in Peng & Huang (2008) to approximate the null distribution of $K_n(\tau)$ and obtain the p-value of the supremum-based lack-of-fit test to be 0.815. It suggests that Model (1) provides a reasonably good fit for the blood pressure data.

7 Conclusion

We consider a survival analysis setting with a functional covariate, and develop a functional censored quantile regression approach to study the time-varying effect of the functional covariate on the surivival times. Our motivating stroke application is about identifying potential time-varying effect of ambulatory blood pressure on time to stroke recurrence. We use B-spline bases to approximate the coefficient function of the functional covariate, and establish consistency and asymptotic normality of the method. We develop a flexible IGACV method to select the number of knots in the B-spline approximation. We conduct extensive simulations to study the finite sample properties of the estimators and evaluate the performance of the IGACV knot selection procedure. The application to the stroke study reveals a couple of interesting findings that are of clinical interests. First, the overall magnitude of the blood pressure trajectory has a significant negative effect on stroke recurrence. Furthermore, the estimated time-varying effect suggests that the morning blood pressure surge (from 6:00am to 10:00am) indeed significantly increases the risk of stroke recurrence, consistently across multiple quantiles of the time-to-recurrence distribution. This phenomenon has been noticed in the medical literature, but hasn't been validated statistically before.

Although the quantile function $Q_T(\tau|\mathbf{X}, Z, \boldsymbol{\beta}, \alpha)$ is monotone, its estimator may not be. Should this happen, we could adopt Peng & Huang (2008)'s martingale estimation procedure to ensure the monotonicity of the estimator. In special cases, when the covariates are nonnegative and one can use non-negative B-spline bases, a simpler adjustment can be done. To be more specific, we can adjust the estimators so that $\widehat{\beta}_l(\tau_j) \geq \widehat{\beta}_l(\tau_k)$ and $\widehat{\gamma}_l(\tau_j) \geq \widehat{\gamma}_l(\tau_k)$ for $\tau_j \geq \tau_k$ to achieve the monotonicity of the quantile function, where $\widehat{\beta}_l$ and

$\widehat{\gamma}_l$ are the lth element of $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\gamma}}$, respectively. Note that the aforementioned adjustments are not necessary in our numerical studies as the default estimates obtained are already monotonic.

In practice, the discretization mechanism of $Z(\cdot)$ affects the estimation results. To see the effect, in the real data, instead of the default 15-minute sampling interval, we sample the blood pressure $Z(\cdot)$ every 30 minutes (resulting in 48 time points), and every 45 minutes (32 time points). We then implement our method on the newly constructed covariates. Figure S.2 in the supplementary material shows that when $Z(\cdot)$ contains 48 time points, the functional estimates are similar to those in Figure 7 with wider confidence intervals. On the other hand, when $Z(\cdot)$ contains only 32 time points (Figure S.3 in the supplementary material), the functional estimates fail to capture the time varying pattern of the covariate effects. This phenomenon implies that one needs a minimal number of time points to achieve estimation consistency. To derive such a lower bound requires in-depth analyses on the structure of the functional parameters, which is out of the scope of the current research. Further research along this line is needed.

References

Andersen, P. K., Borgan, O., Gill, R. D. & Keiding, N. (2012), *Statistical Models Based on Counting Processes*, Springer Science & Business Media.

Bang, H. & Tsiatis, A. A. (2002), 'Median regression with censored cost data', *Biometrics* **58**(3), 643–649.

Buckley, J. & James, I. (1979), 'Linear regression with censored data', *Biometrika* **66**(3), 429–436.

Cardot, H., Crambes, C. & Sarda, P. (2005), 'Quantile regression when the covariates are functions', *Nonparametric Statistics* **17**(7), 841–856.

Chambers, J. M., Hastie, T. J. et al. (1992), *Statistical Models in S*, Vol. 251, Wadsworth & Brooks/Cole Advanced Books & Software Pacific Grove, CA.

Chen, X. & Pouzo, D. (2012), 'Estimation of nonparametric conditional moment models with possibly nonsmooth generalized residuals', *Econometrica* **80**(1), 277–321.

Chernozhukov, V. & Hong, H. (2002), 'Three-step censored quantile regression and extramarital affairs', *Journal of the American Statistical Association* **97**(459), 872–882.

De Boor, C. (1978), *A Practical Guide to Splines*, Vol. 27, Springer-Verlag, New York.

Ferraty, F., Rabhi, A. & Vieu, P. (2005), 'Conditional quantiles for dependent functional data with application to the climatic El niño phenomenon', *Sankhyā: The Indian Journal of Statistics* **67**(2), 378–398.

Fleming, T. R. & Harrington, D. P. (2011), *Counting Processes and Survival Analysis*, Vol. 169, John Wiley & Sons.

Gill, R. D. & Johansen, S. (1990), 'A survey of product-integration with a view toward application in survival analysis', *Annals of Statistics* **18**(4), 1501–1555.

He, X. & Shi, P. (1994), 'Convergence rate of b-spline estimators of nonparametric conditional quantile functions', *Journal of Nonparametric Statistics* **3**(3-4), 299–308.

Jiang, F. & Haneuse, S. (2017), 'A semi-parametric transformation frailty model for semi-competing risks survival data', *Scandinavian Journal of Statistics* **44**(1), 112–129.

Kato, K. (2012), 'Estimation in functional linear quantile regression', *Annals of Statistics* **40**(6), 3108–3136.

Koenker, R. (2008), 'Censored quantile regression redux', *Journal of Statistical Software* **27**(6), 1–25.

Koenker, R. (2012), 'quantreg: Quantile regression. R package version 4.79'.

Koenker, R. & Geling, O. (2001), 'Reappraising medfly longevity: a quantile regression survival analysis', *Journal of the American Statistical Association* **96**(454), 458–468.

Lee, E. R., Noh, H. & Park, B. U. (2014), 'Model selection via Bayesian information criterion for quantile regression models', *Journal of the American Statistical Association* **109**(505), 216–229.

Lindgren, A. (1997), 'Quantile regression with censored data using generalized L1 minimization', *Computational Statistics & Data Analysis* **23**(4), 509–524.

Louis, T. A. (1981), 'Nonparametric analysis of an accelerated failure time model', *Biometrika* **68**(2), 381–390.

Peng, L. & Huang, Y. (2008), 'Survival analysis with quantile regression models', *Journal of the American Statistical Association* **103**(482), 637–649.

Portnoy, S. (2003), 'Censored regression quantiles', *Journal of the American Statistical Association* **98**(464), 1001–1012.

Prentice, R. L. (1978), 'Linear rank tests with right censored data', *Biometrika* **65**(1), 167–179.

Qu, S., Wang, J. L. & Wang, X. (2016), 'Optimal estimation for the functional Cox model', *Annals of Statistics* **44**(4), 1708–1738.

Wang, H. J. & Wang, L. (2009), 'Locally weighted censored quantile regression', *Journal of the American Statistical Association* **104**(487), 1117–1128.

Wei, L. J. & Gail, M. H. (1983), 'Nonparametric estimation for a scale-change with censored observations', *Journal of the American Statistical Association* **78**(382), 382–388.

Wei, L. J., Ying, Z. & Lin, D. (1990), 'Linear regression analysis of censored survival data based on rank tests', *Biometrika* **77**(4), 845–851.

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

Ying, Z., Jung, S. & Wei, L. J. (1995), 'Survival analysis with median regression models', *Journal of the American Statistical Association* **90**(429), 178–184.

Yuan, M. (2006), 'GACV for quantile smoothing splines', *Computational Statistics & Data Analysis* **50**(3), 813–829.

Table 1: Estimation results from 1000 simulations for $\widehat{\boldsymbol{\beta}}(\tau) = \{\widehat{\beta}_1(\tau), \widehat{\beta}_2(\tau)\}^{\mathrm{T}}$ at different quantiles with $\sigma_\epsilon = 0.2$ and a censoring rate of 10%. SD represents the empirical standard deviation, SE is the standard error, and CP stands for the 95% coverage probability.

		$n = 200$				$n = 500$			
v	$\widehat{\boldsymbol{\beta}}(\tau)$	BIAS	SD	SE	CP	BIAS	SD	SE	CP
1	$\widehat{\beta}_1(0.3)$	0.001	0.054	0.060	0.939	0.000	0.033	0.039	0.960
	$\widehat{\beta}_2(0.3)$	0.017	0.179	0.205	0.964	0.008	0.112	0.132	0.960
	$\widehat{\beta}_1(0.4)$	0.001	0.050	0.058	0.960	0.000	0.033	0.037	0.952
	$\widehat{\beta}_2(0.4)$	0.015	0.174	0.197	0.956	0.006	0.109	0.128	0.958
	$\widehat{\beta}_1(0.5)$	0.000	0.051	0.058	0.949	0.001	0.034	0.038	0.956
	$\widehat{\beta}_2(0.5)$	0.011	0.178	0.195	0.946	0.007	0.109	0.127	0.968
	$\widehat{\beta}_1(0.6)$	0.002	0.052	0.061	0.947	0.000	0.034	0.039	0.956
	$\widehat{\beta}_2(0.6)$	0.006	0.182	0.203	0.954	0.007	0.112	0.128	0.957
	$\widehat{\beta}_1(0.7)$	0.002	0.054	0.063	0.952	0.000	0.035	0.041	0.955
	$\widehat{\beta}_2(0.7)$	0.007	0.187	0.214	0.953	0.005	0.117	0.139	0.954
1.5	$\widehat{\beta}_1(0.3)$	0.001	0.042	0.048	0.952	0.000	0.026	0.031	0.961
	$\widehat{\beta}_2(0.3)$	0.016	0.142	0.163	0.954	0.009	0.089	0.104	0.961
	$\widehat{\beta}_1(0.4)$	0.001	0.041	0.046	0.953	0.001	0.026	0.029	0.955
	$\widehat{\beta}_2(0.4)$	0.013	0.139	0.157	0.961	0.007	0.087	0.100	0.959
	$\widehat{\beta}_1(0.5)$	0.001	0.041	0.046	0.948	0.000	0.027	0.030	0.957
	$\widehat{\beta}_2(0.5)$	0.010	0.143	0.153	0.937	0.009	0.088	0.100	0.961
	$\widehat{\beta}_1(0.6)$	0.001	0.041	0.048	0.954	0.000	0.026	0.030	0.950
	$\widehat{\beta}_2(0.6)$	0.006	0.144	0.160	0.953	0.007	0.089	0.102	0.955
	$\widehat{\beta}_1(0.7)$	0.001	0.044	0.050	0.953	0.000	0.027	0.032	0.960
	$\widehat{\beta}_2(0.7)$	0.008	0.149	0.168	0.945	0.008	0.094	0.109	0.959
2	$\widehat{\beta}_1(0.3)$	0.000	0.038	0.044	0.958	0.001	0.024	0.029	0.967
	$\widehat{\beta}_2(0.3)$	0.016	0.127	0.144	0.943	0.009	0.081	0.095	0.960
	$\widehat{\beta}_1(0.4)$	0.000	0.037	0.042	0.950	0.000	0.024	0.027	0.952
	$\widehat{\beta}_2(0.4)$	0.015	0.123	0.139	0.953	0.007	0.079	0.091	0.955
	$\widehat{\beta}_1(0.5)$	0.001	0.037	0.042	0.954	0.001	0.024	0.028	0.956
	$\widehat{\beta}_2(0.5)$	0.009	0.127	0.136	0.938	0.008	0.080	0.091	0.962
	$\widehat{\beta}_1(0.6)$	0.001	0.037	0.043	0.957	0.001	0.025	0.028	0.952
	$\widehat{\beta}_2(0.6)$	0.005	0.127	0.141	0.947	0.009	0.082	0.093	0.950
	$\widehat{\beta}_1(0.7)$	0.001	0.040	0.045	0.954	0.001	0.025	0.030	0.966
	$\widehat{\beta}_2(0.7)$	0.008	0.133	0.149	0.946	0.008	0.086	0.100	0.957
2.5	$\hat{\beta}^{(1)}$	0.000	0.036	0.042	0.967	0.000	0.022	0.027	0.967
	$\hat{\beta}^{(2)}$	0.014	0.121	0.134	0.951	0.009	0.074	0.087	0.958
	$\hat{\beta}^{(1)}$	0.000	0.035	0.040	0.955	0.000	0.022	0.025	0.962
	$\hat{\beta}^{(2)}$	0.014	0.115	0.130	0.954	0.008	0.073	0.083	0.954
	$\hat{\beta}^{(1)}$	0.001	0.036	0.040	0.951	0.000	0.023	0.026	0.954
	$\hat{\beta}^{(2)}$	0.008	0.118	0.127	0.942	0.008	0.074	0.083	0.953
	$\hat{\beta}^{(1)}$	0.001	0.035	0.041	0.961	0.000	0.023	0.026	0.953
	$\hat{\beta}^{(2)}$	0.006	0.118	0.132	0.942	0.009	0.076	0.084	0.941
	$\hat{\beta}^{(1)}$	0.001	0.037	0.043	0.953	0.000	0.023	0.028	0.964
	$\hat{\beta}^{(2)}$	0.007	0.126	0.140	0.951	0.009	0.080	0.091	0.950

Table 2: Estimation results from 1000 simulations for $\widehat{\boldsymbol{\beta}}(\tau) = \{\widehat{\beta}_1(\tau), \widehat{\beta}_2(\tau)\}^{\mathrm{T}}$ at different quantiles with $\sigma_\epsilon = 0.2$ and a censoring rate of 20%. SD represents the empirical standard deviation, SE is the standard error, and CP stands for the 95% coverage probability.

		$n = 200$				$n = 500$			
v	$\widehat{\boldsymbol{\beta}}(\tau)$	BIAS	SD	SE	CP	BIAS	SD	SE	CP
1	$\widehat{\beta}_1(0.3)$	0.001	0.051	0.057	0.956	0.001	0.031	0.035	0.953
	$\widehat{\beta}_2(0.3)$	0.014	0.172	0.197	0.960	0.009	0.107	0.124	0.957
	$\widehat{\beta}_1(0.4)$	0.000	0.048	0.054	0.961	0.001	0.030	0.034	0.949
	$\widehat{\beta}_2(0.4)$	0.010	0.171	0.189	0.950	0.006	0.105	0.119	0.963
	$\widehat{\beta}_1(0.5)$	0.000	0.048	0.054	0.956	0.000	0.030	0.034	0.946
	$\widehat{\beta}_2(0.5)$	0.010	0.170	0.186	0.948	0.009	0.102	0.118	0.960
	$\widehat{\beta}_1(0.6)$	0.001	0.049	0.056	0.940	0.001	0.030	0.034	0.951
	$\widehat{\beta}_2(0.6)$	0.007	0.172	0.192	0.948	0.008	0.104	0.118	0.964
	$\widehat{\beta}_1(0.7)$	0.001	0.051	0.057	0.933	0.000	0.032	0.036	0.949
	$\widehat{\beta}_2(0.7)$	0.007	0.178	0.200	0.947	0.006	0.110	0.126	0.954
1.5	$\widehat{\beta}_1(0.3)$	0.001	0.041	0.045	0.955	0.001	0.024	0.028	0.951
	$\widehat{\beta}_2(0.3)$	0.015	0.137	0.156	0.952	0.009	0.085	0.099	0.957
	$\widehat{\beta}_1(0.4)$	0.000	0.038	0.043	0.941	0.001	0.024	0.027	0.959
	$\widehat{\beta}_2(0.4)$	0.009	0.136	0.149	0.952	0.006	0.082	0.094	0.951
	$\widehat{\beta}_1(0.5)$	0.000	0.038	0.043	0.955	0.000	0.024	0.027	0.954
	$\widehat{\beta}_2(0.5)$	0.008	0.137	0.147	0.948	0.008	0.083	0.094	0.962
	$\widehat{\beta}_1(0.6)$	0.001	0.039	0.044	0.941	0.001	0.024	0.027	0.950
	$\widehat{\beta}_2(0.6)$	0.007	0.139	0.152	0.949	0.009	0.084	0.094	0.966
	$\widehat{\beta}_1(0.7)$	0.001	0.041	0.046	0.943	0.000	0.025	0.029	0.959
	$\widehat{\beta}_2(0.7)$	0.008	0.142	0.159	0.940	0.008	0.088	0.100	0.948
2	$\widehat{\beta}_1(0.3)$	0.001	0.036	0.040	0.954	0.000	0.022	0.025	0.946
	$\widehat{\beta}_2(0.3)$	0.014	0.122	0.136	0.938	0.008	0.074	0.086	0.954
	$\widehat{\beta}_1(0.4)$	0.000	0.034	0.039	0.952	0.001	0.021	0.023	0.953
	$\widehat{\beta}_2(0.4)$	0.011	0.118	0.130	0.943	0.006	0.072	0.081	0.955
	$\widehat{\beta}_1(0.5)$	0.001	0.033	0.038	0.960	0.000	0.021	0.024	0.949
	$\widehat{\beta}_2(0.5)$	0.007	0.120	0.129	0.942	0.008	0.073	0.081	0.963
	$\widehat{\beta}_1(0.6)$	0.001	0.034	0.039	0.949	0.000	0.021	0.024	0.956
	$\widehat{\beta}_2(0.6)$	0.006	0.120	0.132	0.945	0.009	0.073	0.082	0.955
	$\widehat{\beta}_1(0.7)$	0.001	0.037	0.041	0.941	0.000	0.022	0.025	0.958
	$\widehat{\beta}_2(0.7)$	0.007	0.125	0.138	0.945	0.008	0.077	0.087	0.948
2.5	$\widehat{\beta}_1(0.3)$	0.001	0.033	0.038	0.959	0.000	0.020	0.023	0.956
	$\widehat{\beta}_2(0.3)$	0.013	0.113	0.125	0.939	0.008	0.068	0.078	0.959
	$\widehat{\beta}_1(0.4)$	0.000	0.032	0.036	0.955	0.001	0.020	0.022	0.951
	$\widehat{\beta}_2(0.4)$	0.011	0.108	0.120	0.953	0.006	0.066	0.074	0.958
	$\widehat{\beta}_1(0.5)$	0.000	0.032	0.036	0.958	0.000	0.020	0.022	0.948
	$\widehat{\beta}_2(0.5)$	0.006	0.109	0.119	0.948	0.008	0.067	0.075	0.960
	$\widehat{\beta}_1(0.6)$	0.001	0.032	0.036	0.953	0.000	0.020	0.022	0.954
	$\widehat{\beta}_2(0.6)$	0.005	0.110	0.120	0.950	0.008	0.068	0.075	0.948
	$\widehat{\beta}_1(0.7)$	0.001	0.034	0.038	0.940	0.000	0.020	0.024	0.958
	$\widehat{\beta}_2(0.7)$	0.007	0.115	0.127	0.953	0.009	0.071	0.080	0.952

Table 3: Estimation results from 1000 simulations for $\widehat{\boldsymbol{\beta}}(\tau) = \{\widehat{\beta}_1(\tau), \widehat{\beta}_2(\tau)\}^{\mathrm{T}}$ at different quantiles with $\sigma_\epsilon = 0.2$ and a censoring rate of 40%. SD represents the empirical standard deviation, SE is the standard error, and CP stands for the 95% coverage probability.

		$n = 200$				$n = 500$			
v	$\widehat{\boldsymbol{\beta}}(\tau)$	BIAS	SD	SE	CP	BIAS	SD	SE	CP
1	$\widehat{\beta}_1(0.3)$	0.003	0.065	0.075	0.961	0.003	0.040	0.050	0.970
	$\widehat{\beta}_2(0.3)$	0.021	0.203	0.245	0.968	0.009	0.130	0.157	0.959
	$\widehat{\beta}_1(0.4)$	0.003	0.063	0.074	0.960	0.003	0.040	0.048	0.964
	$\widehat{\beta}_2(0.4)$	0.014	0.202	0.238	0.959	0.004	0.126	0.152	0.964
	$\widehat{\beta}_1(0.5)$	0.003	0.064	0.076	0.957	0.003	0.042	0.049	0.955
	$\widehat{\beta}_2(0.5)$	0.009	0.210	0.239	0.951	0.005	0.127	0.157	0.977
	$\widehat{\beta}_1(0.6)$	0.003	0.068	0.080	0.961	0.002	0.043	0.051	0.963
	$\widehat{\beta}_2(0.6)$	0.004	0.216	0.255	0.954	0.004	0.133	0.162	0.960
	$\widehat{\beta}_1(0.7)$	0.006	0.071	0.087	0.953	0.002	0.045	0.056	0.961
	$\widehat{\beta}_2(0.7)$	0.002	0.232	0.280	0.959	0.002	0.144	0.179	0.958
1.5	$\widehat{\beta}_1(0.3)$	0.004	0.054	0.066	0.972	0.002	0.033	0.041	0.967
	$\widehat{\beta}_2(0.3)$	0.024	0.172	0.213	0.970	0.009	0.112	0.134	0.952
	$\widehat{\beta}_1(0.4)$	0.003	0.053	0.064	0.967	0.001	0.033	0.040	0.947
	$\widehat{\beta}_2(0.4)$	0.019	0.177	0.206	0.965	0.007	0.110	0.129	0.964
	$\widehat{\beta}_1(0.5)$	0.002	0.055	0.065	0.961	0.002	0.035	0.041	0.954
	$\widehat{\beta}_2(0.5)$	0.012	0.180	0.207	0.955	0.008	0.111	0.132	0.968
	$\widehat{\beta}_1(0.6)$	0.002	0.057	0.068	0.958	0.002	0.036	0.043	0.958
	$\widehat{\beta}_2(0.6)$	0.006	0.186	0.218	0.962	0.009	0.114	0.137	0.963
	$\widehat{\beta}_1(0.7)$	0.004	0.061	0.073	0.958	0.001	0.037	0.046	0.954
	$\widehat{\beta}_2(0.7)$	0.003	0.200	0.236	0.956	0.006	0.123	0.151	0.959
2	$\widehat{\beta}_1(0.3)$	0.003	0.048	0.059	0.967	0.002	0.030	0.038	0.969
	$\widehat{\beta}_2(0.3)$	0.023	0.154	0.188	0.956	0.014	0.100	0.121	0.964
	$\widehat{\beta}_1(0.4)$	0.003	0.048	0.057	0.965	0.001	0.030	0.035	0.963
	$\widehat{\beta}_2(0.4)$	0.023	0.157	0.181	0.954	0.010	0.098	0.116	0.960
	$\widehat{\beta}_1(0.5)$	0.002	0.048	0.057	0.955	0.002	0.031	0.037	0.963
	$\widehat{\beta}_2(0.5)$	0.016	0.160	0.184	0.947	0.010	0.099	0.120	0.968
	$\widehat{\beta}_1(0.6)$	0.002	0.049	0.060	0.958	0.002	0.032	0.038	0.948
	$\widehat{\beta}_2(0.6)$	0.012	0.168	0.191	0.959	0.011	0.104	0.123	0.956
	$\widehat{\beta}_1(0.7)$	0.002	0.052	0.064	0.960	0.001	0.033	0.040	0.965
	$\widehat{\beta}_2(0.7)$	0.006	0.178	0.206	0.952	0.008	0.111	0.134	0.955
2.5	$\widehat{\beta}_1(0.3)$	0.003	0.047	0.058	0.968	0.002	0.028	0.036	0.967
	$\widehat{\beta}_2(0.3)$	0.024	0.151	0.179	0.951	0.015	0.096	0.117	0.972
	$\widehat{\beta}_1(0.4)$	0.002	0.046	0.055	0.964	0.001	0.029	0.034	0.955
	$\widehat{\beta}_2(0.4)$	0.022	0.150	0.176	0.957	0.012	0.096	0.111	0.957
	$\widehat{\beta}_1(0.5)$	0.002	0.047	0.055	0.958	0.002	0.029	0.035	0.962
	$\widehat{\beta}_2(0.5)$	0.016	0.152	0.178	0.956	0.012	0.097	0.115	0.961
	$\widehat{\beta}_1(0.6)$	0.002	0.048	0.057	0.957	0.001	0.030	0.036	0.946
	$\widehat{\beta}_2(0.6)$	0.013	0.158	0.186	0.953	0.011	0.099	0.118	0.962
	$\widehat{\beta}_1(0.7)$	0.000	0.049	0.061	0.956	0.001	0.031	0.039	0.971
	$\widehat{\beta}_2(0.7)$	0.006	0.173	0.198	0.953	0.009	0.107	0.127	0.957

Table 4: Optimal d_S based on the IGACV and IBIC criteria under different censoring rates over 500 simulations.

d_S	4	5	6	7	8	4	5	6	7	8
	0% censoring					10% censoring				
IGACV	88	122	**250**	36	4	71	121	**256**	40	12
IBIC	330	102	67	1	0	316	112	72	0	0
	20% censoring					40% censoring				
IGACV	67	121	**250**	42	20	89	104	**169**	74	64
IBIC	288	124	85	3	0	218	135	127	18	2

Table 5: Estimation results for $\boldsymbol{\beta}(\tau)$ based on 1000 simulations with $n = 297$ and $\sigma_\epsilon = 0.2$ under different censoring rates.

τ	BIAS	SD	SE	CP	BIAS	SD	SE	CP
	0% censoring				10% censoring			
0.2	0.018	0.179	0.202	0.972	0.019	0.180	0.203	0.966
0.3	0.012	0.167	0.187	0.961	0.011	0.168	0.189	0.959
0.4	0.012	0.157	0.177	0.962	0.013	0.160	0.182	0.957
0.5	0.011	0.158	0.180	0.953	0.009	0.162	0.182	0.955
0.6	0.007	0.164	0.181	0.941	0.005	0.168	0.188	0.945
0.7	0.009	0.173	0.193	0.951	0.007	0.182	0.202	0.950
0.8	0.016	0.188	0.209	0.964	0.012	0.204	0.231	0.962
	20% censoring				40% censoring			
0.2	0.017	0.182	0.205	0.966	0.014	0.186	0.217	0.965
0.3	0.012	0.170	0.192	0.960	0.008	0.178	0.205	0.959
0.4	0.011	0.162	0.183	0.955	0.006	0.173	0.197	0.959
0.5	0.012	0.165	0.188	0.959	0.002	0.177	0.207	0.955
0.6	0.007	0.175	0.197	0.951	0.009	0.195	0.231	0.959
0.7	0.001	0.192	0.215	0.951	0.018	0.231	0.285	0.967
0.8	0.003	0.222	0.261	0.974	0.046	0.366	0.489	0.989

Table 6: Estimation results for $\boldsymbol{\beta}(\tau)$ based on 1000 simulations with $n = 297$ and $\sigma_\epsilon = 0.2$ with complete data by the method in Kato (2012).

τ	BIAS	SSE	SEE	CP
0.2	0.004	0.177	0.197	0.970
0.3	0.002	0.169	0.180	0.959
0.4	0.006	0.159	0.173	0.960
0.5	0.005	0.157	0.172	0.955
0.6	0.008	0.166	0.175	0.959
0.7	0.006	0.177	0.183	0.944
0.8	0.004	0.189	0.196	0.956

Table 7: Selection of the optimal number of basis functions using the IGACV and IBIC criteria respectively for the stroke data.

	d_S				
Criterion	4	5	6	7	8
IGACV	0.35431	0.35557	0.35365	**0.35329**	0.35535
IBIC	-0.61645	-0.60511	-0.60075	-0.59379	-0.58103

Table 8: Estimates $\hat{\boldsymbol{\beta}}(\tau)$ and the 95% bootstrap confidence intervals (CIs) at different values of τ for the stroke data.

Method	Quantile τ	0.3	0.4	0.5	0.6
FCQR		-0.0086	-0.0074	-0.0073	-0.0068
	95% CI	(-0.0120, -0.0052)	(-0.0101, -0.0046)	(-0.0097, -0.0048)	(-0.0081, -0.0055)
Kato (2012)		-0.0142	-0.0110	-0.0095	-0.0088
	95% CI	(-0.0180, -0.0105)	(-0.0127, -0.0092)	(-0.0104, -0.0085)	(-0.0093, -0.0083)