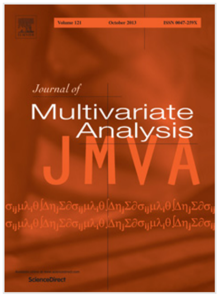
Accepted Manuscript



Quantile regression of longitudinal data with informative observation times

Xuerong Chen, Niansheng Tang, Yong Zhou

PII: S0047-259X(15)00275-4
DOI: http://dx.doi.org/10.1016/j.jmva.2015.11.007
Reference: YJMVA 4047

To appear in: *Journal of Multivariate Analysis*

Received date: 15 April 2013

Please cite this article as: X. Chen, N. Tang, Y. Zhou, Quantile regression of longitudinal data with informative observation times, *Journal of Multivariate Analysis* (2015), http://dx.doi.org/10.1016/j.jmva.2015.11.007

Quantile regression of longitudinal data with informative observation times

Xuerong Chen[1], Niansheng Tang[4] and Yong Zhou[2,3]

[1]School of Statistics, Southwestern University of Finance and Economics, Chengdu, Sichuan, China

[2]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

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

[4]Department of Statistics, Yunnan University, Kunming, Yunnan, China

Abstract

Longitudinal data are frequently encountered in medical follow-up studies and economic research. Conditional mean regression and conditional quantile regression are often used to fit longitudinal data. Many methods focused on the cases where the observation times are independent of the response variables or conditionally independent of them given the covariates. Few papers have considered the case where the the response variables depend on the observation times or observation times are random variables associated with a counting process. In this paper, we propose a marginally conditional quantile regression approach for modeling longitudinal data with random observing times and informative observation times. Estimators of the parameters in the proposed conditional quantile regression are derived by constructing non-smooth estimating equations when the observation times follow a counting process. Consistency and asymptotic normality for these estimators are established. Asymptotic variance is estimated based on a resampling method. A simulation study is conducted and suggests that the finite sample performance of the proposed approach is very good, and an illustrative approach is provided.

Keywords: Estimating equation; Informative observation times; Longitudinal data; Quantile regression; Resampling method.

1 Introduction

Longitudinal data arise frequently in many types of studies, for example, medical follow-up studies and observational investigations. In these longitudinal studies, observations from an individual are collected repeatedly over time. Various methods including generalized estimating equation and random effects model have been developed to analyze longitudinal data, see, Diggle et al. (1994) and Laird & Ware (1982). Recently, non-parametric and semiparametric models for longitudinal data have attracted much attention. For nonparametric methods see Hoover et al. (1998), Wu et al. (1998), Scheike & Zhang (1998), Wu & Zhang (2002b), Wu & Liang (2004), and Sun & Wu(2003); for semiparametric approaches see Moyeed & Diggle (1994), Wu & Zhang (2000a), Liang et al. (2003). One major difficulty in analyzing longitudinal data is that the observation times are often different across subjects. In Martinussen & Scheike (1999, 2000, 2001) and Lin & Ying (2001), these authors considered time-varying coefficient regression models for longitudinal data, based on modeling the observation times by counting process,. Under this framework, the observation times were allowed to have arbitrary pattern and to depend on covariates. These seminal work provided ways for modeling the time-dependent observations, survival data and recurrent event data in a unified framework.

Most existing methods assumed that the response variable is independent of observation times completely or conditionally independent given the covariates. This assumption may be unrealistic in applications. Informative observation times often occur when they are subject- or response variable- dependent. For example, consider the bladder cancer study conducted by Veterans Administration Cooperative Urological Research Group (See Sun & Wei (2000)). In the beginning of this study, all subjects who participated in the study had superficial bladder tumors and these tumors were removed. During the study, many patients suffered from multiple recurrences of tumor, and the recurrent tumors were removed during clinical visits. The clinical visit times and the number of tumors occurred between clinical visits were collected. One aim of this study was to compare the recurrence rates of the tumors of patients in different treatment groups. It is worth noting that, some subjects had significantly more clinical visits than others, which suggests that the number of clinical visits may contain some information about the tumor recurrence rates and the clinic visit times may depend on subject or covariate. It is important to make use of these information for inference on the recurrence rate of tumor. This motivated several authors to consider to incorporating the informative observation times in longitudinal data analysis. For example, Sun et al. (2005) considered a semiparametric regression approach by using the estimating equation approach when the response variable depends on the observation times. They proposed a marginal model for the response variable process conditional on the covariates and the observation times, and the observation times was assumed to follow a counting process. Their model is a generalization of the marginal model proposed by Lin & Ying (2001). Almost all papers on longitudinal data whose observation times follow a counting process were conducted by conditional mean regression method. Besides the traditional conditional mean regression method, conditional quantile regression is another important approach used in longitudinal data analysis. When data contain some outliers or the error distribution is skewed or has heavy tails, the latter method is more robust and efficient than the former one.

Quantile regression method has been widely applied to the analysis of of longitudinal data. He et al. (2003) reviewed and compared three estimators of median regression in linear models for longitudinal data. Motivated by the penalized least squares for random effects models, Koenker (2004) proposed a penalized quantile regression method when there were a large number of individual fixed effects that can significantly inflate the variability of the estimates of the main covariate effects. Karlsson (2005) considered the nonlinear quantile regression model for longitudinal data. Fenske et al. (2008) detected the risk factor for obesity in early childhood by using quantile regression methods for longitudinal data. Mu & Wei (2009) studied the dynamic quantile regression transformation model for longitudinal data. Liu & Bottai (2009) studied the mixed-effects models with longitudinal data by employing the quantile regression method. Wang & Fygeson (2009) developed quantile regression inference procedures for longitudinal data when some of the measurements were censored by fixed constants. Wang et al. (2009) developed a quantile estimation method for partially linear varying coefficient models using splines. Wang & Zhu (2011) considered a quantile regression approach for longitudinal data by empirical likelihood method.

However, the existing literature of quantile regression for longitudinal data did not consider the case where observation times are informative. Furthermore, the observations times in the estimating methods are assumed to be independent of the covariates. To relax these limitations, in this paper, we study the quantile regression method for longitudinal data when the response variable depends on the observation times which depend on covariates by following a counting process. To make inference to parameters, estimating equations are constructed. The main difficulties are the Taylor expansion can not be used to derive the asymptotic distribution of the estimators and Newton algorithm can no longer be used to compute the estimators, because the involved estimating equations are non-smooth. In this paper, the key results of empirical process theory, namely the uniform law of large number and the stochastic equicontinuity, are used to derive the asymptotic properties of estimators. This method has been used in the literatures on non-smooth estimating equations, see for example, Pakes & Pollard (1989) and Chen et al. (2003). To overcome the computational difficulty, an iterative method based on MM algorithm of quantile regression(see Hunter and Lange 2000) is proposed. Because it is not easy to estimate the asymptotic variances of quantile regression estimators directly, these asymptotic variances are estimated by using the resampling method proposed by Jin et al. (2001) in this paper.

This paper is organized as follows. In Section 2, we introduce some notations and describe the models we consider in this paper. In Section 3, the inference procedure and the MM algorithm based-iterative method are provided. The consistency, asymptotic normality of the proposed estimators and the asymptotic variance estimate are given in Section 4. Section 5 reports the simulation results and a real data example is given in section 6. The conditions and the proofs are presented in Section 7.

2 Notation and Statistical Models

Suppose that a longitudinal study consists of a random sample of n subjects. For subject

i, let $Y_i(t)$ be the response variable and $X_i(t)$ be a p-dimensional vector of possibly time-dependent covariates, $i = 1, \ldots, n$. The observations of $Y_i(t)$ are taken at time points $t_{i1} < \ldots < t_{in_i}$, where n_i is the total number of observations on the ith subject. The number of observations of the ith subject by time t is $N_i(t) = \sum_{j=1}^{n_i} I(t_{ij} \le t) = N_i^*(\min(t, C_i))$, where C_i is the follow-up time or censoring time for the ith subject and $N_i^*(t)$ is the underlying counting process of sampling times for subject i. We assume that the covariate history $\{X_i(t) : 0 \le t \le C_i\}$ is observed for each individual.

For inference about the response process $Y_i(t)$, if it is completely or conditionally independent of $N_i^*(t)$, then the marginal approach is usually used (Lin & Ying(2001)). Otherwise, as described in Sun et al. (2005), there are three choices: modeling them jointly, modeling $Y_i(t)$ marginally and then $N_i^*(t)$ conditional on it, or modeling $N_i^*(t)$ marginally and then $Y_i(t)$ conditional on it. Our main interest is on the longitudinal process, rather than the observation times. The evaluation of the covariate effects on the longitudinal process is also of interest. Hence, we adopt the third choice, that is, we model $N_i^*(t)$ marginally and then model $Y_i(t)$ conditionally on it.

Define $\mathcal{F}_{it} = \{N_i(s), 0 \le s < t\}$ as the history observed up to time t for the ith subject. Consider following quantile regression model

$$Y_i(t) = \beta(\tau)^\top X_i(t) + \alpha(\tau)^\top H(\mathcal{F}_{it}) + \epsilon_i(t), \tag{1}$$

where ϵ_i is random error with $\Pr(\epsilon_i(t) \le 0 | X_i(t), \mathcal{F}_{it}) = \tau$ and $H(\cdot)$ is a vector of known functions of the counting process $N_i(t)$ up to time $t-$. Here $\beta(\tau)$ is a vector of regression parameters for the covariate, $\alpha(\tau)$ is a q-dimensional vector of coefficients for the counting process. This model implies that the response process depends on both coveriates and the history of the observations. Note that in this model, the parameters β and α generally depend on τ. For notational simplicity and without causing confusion, we drop τ from $\beta(\tau)$ and $\alpha(\tau)$ below.

For the observation process, assume that $N_i^*(t)$ is a nonhomogeneous Poisson process with

$$\mathrm{E}\{dN_i^*(t) | X_i(t)\} = e^{\gamma^\top X_i(t)} d\Lambda_0(t), \quad i = 1, \ldots, n, \tag{2}$$

where γ is a vector of unknown regression parameters and $\Lambda_0(t)$ is the mean cumulative number of observations by time t. In the following, suppose that the censoring time C_i may depend on covariates $X_i(t)$ in an arbitrary fashion, but is independent of $N_i^*(t)$ and $Y_i(t)$ given $X_i(t)$ and $\mathcal{F}_{it}$, in the sense that $\mathrm{E}\{dN_i^*(t)|X_i(t), C_i \ge t\} = \mathrm{E}\{dN_i^*(t)|X_i(t)\}$ and $\mathrm{E}\{Y_i(t)|X_i(t), \mathcal{F}_{it}, C_i \ge t\} = \mathrm{E}\{Y_i(t)|X_i(t), \mathcal{F}_{it}\}$. For simplicity, we restrict inference to the time interval $[0, t_0]$.

There are several possible choices for H in (1). A simple and natural choice for H is $H(\mathcal{F}_{it}) = N_i(t-)$, which means that $\mathcal{F}_{it}$ affects the conditional quantile of the response variable by the total number of observations. Another choice is $H(\mathcal{F}_{it}) = \{N_i(t-) - N_i(t - u)\}/u$, that is, only the average number of the observations in the recent u time units contains information on the response variable. If both the total and recent numbers of observations contain information of the response variable, define H as a vector containing both of them.

It is worth noting that Sun et al. (2005) considered the informative observation times by similar idea on the basis of mean regression method.

3 Estimating Procedure and Computation

3.1 Estimating Equation

We develop an estimation procedure for estimating the unknown parameters under the models (1) and (2). When $\gamma = 0$, namely, the observation times are independent of the covariates, a natural approach for estimating β and α is to use the quantile regression principle of Koenker & Bassett (1978) by minimizing

$$\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{t_0}\rho_\tau(Y_i(t)-\beta^\top X_i(t)-\alpha^\top H(\mathcal{F}_{it}))dN_i(t), \tag{3}$$

where $\rho_\tau(\epsilon)=\epsilon(\tau-I(\epsilon<0))$ is the checking function. Note that minimizing above objective function is equivalent to solving the estimation equations

$$D_n(\theta)=\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{t_0}Z_i(t)\Big\{I(Y_i(t)-\theta^\top Z_i(t)\le 0)-\tau\Big\}dN_i(t)=0, \tag{4}$$

where $Z_i(t)=\{X_i^\top(t),H^\top(\mathcal{F}_{it})\}^\top,\theta=\{\beta^\top,\alpha^\top\}^\top$. Note that, the estimating function on the left hand side of above equation is discontinuous with respect to the parameters, which may lead to multiple solutions or the exact solution may not exist, we follow the methods of Ying et al. (1995) to obtain the quantile estimator. That is, a "root" of θ for (4) may be defined as a minimizer of the function $||D_n(\theta)||$, where $||\cdot||$ is the Euclidean distance. The same convention will be used below.

When $\gamma\neq 0$, observation times depend on covariates through (2). At this time, if the estimator still were obtained by minimizing (3) or solving (4) directly, it would lead to the information contained in model (2) can not be fully used and the covariate effects can not be accurately evaluated. Thus we consider the following process

$$g_i(\theta,\gamma,\Lambda_0)=\int_{0}^{t_0}Z_i(t)\Big\{I(Y_i(t)-\theta^\top Z_i(t)\le 0)dN_i(t)-\tau\xi_i(t)e^{\gamma^\top X_i(t)}d\Lambda_0(t)\Big\}, \tag{5}$$

where $\xi_i(t)=I(C_i\ge t)$. Under models (1) and (2) and the assumption that C_i is independent of $N_i^*(t)$ and $Y_i(t)$ given $X_i(t)$ and $\mathcal{F}_{it}$, it is easy to show that $\mathrm{E}\{g_i(\theta,\gamma_0,\Lambda_0)\}=0$. Therefore, we can estimate the parameters by constructing the estimating equations

$$G_n(\theta,\gamma_0,\Lambda_0)=\frac{1}{n}\sum_{i=1}^{n}g_i(\theta,\gamma_0,\Lambda_0)=0.$$

In order to estimate θ, we need to get consistent estimators of γ_0 and Λ_0 firstly because both of them are unknown in the above estimating equation. For this purpose, we estimate them by the partial likelihood method in Lin et al. (2000). With this method, $\widehat{\gamma}$ can be obtained by solving

$$\sum_{i=1}^{n} \int_0^{t_0} \{X_i(t) - \bar{X}(t;\gamma)\} dN_i(t) = 0,$$

where $\bar{X}(t;\gamma) = S^{(1)}(t;\gamma)/S^{(0)}(t;\gamma)$, and

$$S^{(k)}(t;\gamma) = \frac{1}{n} \sum_{j=1}^{n} \xi_i(t) X_j(t)^{\otimes k} \exp(\gamma^\top X_j(t)),\ k = 0, 1, 2,$$

where for any vector a, $a^{\otimes 2} = aa^T$. And $\widehat{\Lambda}_0$ is the Nelson-Aalen type estimator of Λ_0, which is given by

$$\widehat{\Lambda}_0(t) = \sum_{i=1}^{n} \int_0^{\top} \frac{dN_i(u)}{\sum_{j=1}^{n} \xi_j(u) \exp(\widehat{\gamma}^\top X_j(t))}.$$

Finally, we estimate θ by solving the following asymptotically unbiased estimating equations

$$G_n(\theta, \widehat{\gamma}, \widehat{\Lambda}_0) = \frac{1}{n} \sum_{i=1}^{n} g_i(\theta, \widehat{\gamma}, \widehat{\Lambda}_0) = 0. \tag{6}$$

We derive an iterative algorithm for solving this estimating equations when there exists a unique solution of (6) in the next subsection. In order to avoid the problem of no solution caused by the nonsmoothness of the estimating equations, one can minimize the norm of estimating function to compute the estimator by the similar methods in Ying et al. (1995) and Pakes & Pollard (1989). Namely,

$$\widehat{\theta} = \arg\inf_{\theta} \|G_n(\theta, \widehat{\gamma}, \widehat{\Lambda}_0)\|.$$

And we can show that $\widehat{\theta}$ is consistent estimator of the true value θ_0 which is

$$\theta_0 = \arg\inf_{\theta} \|G(\theta, \gamma_0, \Lambda_0)\|,$$

where $G(\theta, \gamma_0, \Lambda_0) = \mathrm{E}\{g_i(\theta, \gamma_0, \Lambda_0)\}$. Of course, the true value θ_0 satisfies that $G(\theta_0, \gamma_0, \Lambda_0) = 0$ when $G(\theta, \gamma, \Lambda)$ is continuous about θ given any γ and λ.

3.2 Computation based on MM Algorithm

It is difficult to solve estimating equation (6) directly because the involved estimating function is not smooth. Here we develop an iterative approach based on the MM-algorithm (Hunter & Lange (2000)) for solving this equation.

Before describing the iterative approach, we first introduce the MM algorithm for quantile regression. The key idea of the MM algorithm is searching a surrogate function, which is easy to be optimized, to replace the objective function when it is difficult to optimize directly. The

MM stands for Majorize-Minimize in quantile regression. Suppose the goal is to calculate the minimizer $\widehat{\theta}$ of following objective function

$$L(\theta) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0} \rho_\tau(Y_i(t) - \theta^T Z_i(t))dN_i(t).$$

First, construct the approximating function of $L(\theta)$ as

$$L_\delta(\theta) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0} \rho_\tau^\delta(Y_i(t) - \theta^T Z_i(t))dN_i(t),$$

where

$$\rho_\tau^\delta(\epsilon(t)) = \rho_\tau(\epsilon(t)) - \frac{\delta}{2}\ln(\delta + |\epsilon(t)|),\ \epsilon(t) = Y(t) - \theta^T Z(t),$$

for $\delta > 0$, which is the perturbation constant and determined in Hunter & Lange (2000). Secondly, we minimize the approximate function $L_\delta(\theta)$ by the MM algorithm. Let θ^k denote the minimum point of $L_\delta(\theta)$ by the k-th iteration and the residual value $\epsilon^k = \epsilon^k(t) = Y(t) - (\theta^k)^\top Z(t)$, then $\rho_\tau^\delta(\epsilon)$ is majorized at ϵ^k by the quadratic function $\zeta_\tau^\delta(\epsilon|\epsilon^k)$, which is appropriately chosen chosen so that $\zeta_\tau^\delta(\epsilon^k|\epsilon^k) = \rho_\tau^\delta(\epsilon^k)$. Define the surrogate function

$$Q_\delta(\theta|\theta^k) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0} \zeta_\tau^\delta(\epsilon_i(t)|\epsilon_i^k(t))dN_i(t),$$

and treat the θ^k as a known value, then minimize $Q_\delta(\theta|\theta^k)$ to obtain the θ^{k+1}. Repeat above iterative steps until a convergence criteria is satisfied, $\widehat{\theta}$ can be obtained. This result can be proved along the lines of Proportion A.2 of Hunter & Lange (2000).

After briefly introducing the MM algorithm, we develop the iterative method based on the MM algorithm. By the proof of Theorem 2 given in the Appendix, we can write

$$\begin{aligned} G_n(\theta, \widehat{\gamma}, \widehat{\Lambda}_0) &= \frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0} Z_i(t)\Big\{I(Y_i(t) - \theta^\top Z_i(t) \le 0) - \tau\Big\}dN_i(t) \\ &+\tau\frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0}\Big\{Z_i(t) - \frac{\tilde{S}^{(1)}(t;\widehat{\gamma})}{S^{(0)}(t;\widehat{\gamma})}\Big\}dN_i(t), \end{aligned} \tag{7}$$

where $\tilde{S}^{(1)}(t;\gamma)$ is defined in the next section. The second term in (7) does not involve the parameter θ, hence the difficulty of solving estimating equation (6) is caused by the non-smoothness of the first term of $G_n(\theta, \widehat{\gamma}, \widehat{\Lambda}_0)$. Motivated by the above MM algorithm, we can construct a smooth estimating equation which is easily solved to replace (6).

Note that solving $G_n(\theta, \widehat{\gamma}, \widehat{\Lambda}_0) = 0$ for θ is equivalent to minimizing the following function

$$\begin{aligned} Q_n(\theta, \widehat{\gamma}, \widehat{\Lambda}_0) &= \frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0} \rho_\tau(Y_i(t) - \theta^\top Z_i(t))dN_i(t) \\ &+\tau\frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0}\theta^\top\Big\{Z_i(t) - \frac{\tilde{S}^{(1)}(t;\widehat{\gamma})}{S^{(0)}(t;\widehat{\gamma})}\Big\}dN_i(t), \end{aligned} \tag{8}$$

We can construct a smooth function using the MM algorithm, that is,

$$
\begin{aligned}
Q_n^\delta(\theta,\widehat{\gamma},\widehat{\Lambda}_0) &= \frac{1}{n}\sum_{i=1}^n\int_0^{t_0}\rho_\tau^\delta(Y_i(t)-\theta^\top Z_i(t))dN_i(t)\\
&+\tau\frac{1}{n}\sum_{i=1}^n\int_0^{t_0}\theta^\top\Big\{Z_i(t)-\frac{\tilde{S}^{(1)}(t;\widehat{\gamma})}{S^{(0)}(t;\widehat{\gamma})}\Big\}dN_i(t), \qquad (9)
\end{aligned}
$$

where $\rho_\tau^\delta(\cdot)$ is defined above.

Let θ^k be the solution in kth iteration during minimizing (9) and the residual value $\epsilon^k=\epsilon^k(t)=Y(t)-(\theta^k)^\top Z(t)$, then $\rho_\tau^\delta(\epsilon)$ is majorized at ϵ^k by the quadratic function $\zeta_\tau^\delta(\epsilon|\epsilon^k)$ so that $\zeta_\tau^\delta(\epsilon^k|\epsilon^k)=\rho_\tau^\delta(\epsilon^k)$, where $\zeta_\tau^\delta(\epsilon|\epsilon^k)$ has been defined above. We just minimize the following function

$$
Q_n^\delta(\theta,\widehat{\gamma},\widehat{\Lambda}_0)=Q_\delta(\theta|\theta^k)+\tau\frac{1}{n}\sum_{i=1}^n\int_0^{t_0}\theta^\top\Big\{\{Z_i(t)-\frac{\tilde{S}^{(1)}(t;\widehat{\gamma})}{S^{(0)}(t;\widehat{\gamma})}\Big\}dN_i(t), \qquad (10)
$$

Therefore we can construct the $(k+1)$-step surrogate estimating equation

$$
\bar{G}_n^{(k+1)}(\theta,\widehat{\gamma},\widehat{\Lambda}_0|\theta^k)=\frac{dQ_\delta(\theta|\theta^k)}{d\theta}+\tau\sum_{i=1}^n\int_0^{t_0}\Big\{Z_i(t)-\frac{\tilde{S}^{(1)}(t;\widehat{\gamma})}{S^{(0)}(t;\widehat{\gamma})}\Big\}dN_i(t)=0. \qquad (11)
$$

The θ^{k+1} can be obtained by solving (11). The iteration can be carried out in this way, we stop the iteration when a convergence criteria is satisfied, then the proposed estimators can be obtained.

4 Main Results

4.1 Large Sample Properties

In this section, we provide the asymptotic properties of the proposed quantile estimator. Some useful notations are given firstly.

Define

$$
\begin{aligned}
\Omega &= \frac{1}{n}\sum_{i=1}^n\Big[\int_0^{t_0}\Big\{\frac{s^{(2)}(t;\gamma_0)}{s^{(0)}(t;\gamma_0)}-\bar{x}^{\otimes 2}(t;\gamma_0)\Big\}dN_i(t)\Big], A=\mathrm{E}\Big\{\int_0^{t_0}Z_i(t)Z_i^\top(t)f_\epsilon(0|X_i(t),\mathcal{F}_{it})dN_i(t)\Big\},\\
P &= \frac{1}{n}\sum_{i=1}^n\Big[\int_0^{t_0}\Big\{\frac{\tilde{s}^{(2)}(t;\gamma_0)}{s^{(0)}(t;\gamma_0)}-\frac{\tilde{s}^{(1)}(t;\gamma_0)s^{(1)}(t;\gamma_0)}{(s^{(0)}(t;\gamma_0))^2}\Big\}dN_i(t)\Big], \bar{Z}(t;\gamma)=\frac{\tilde{S}^{(1)}(t;\gamma)}{S^{(0)}(t;\gamma)}, \bar{z}(t;\gamma)=\frac{\tilde{s}^{(1)}(t;\gamma)}{s^{(0)}(t;\gamma)},
\end{aligned}
$$

and

$$
h_i(\theta)=\int_0^{t_0}Z_i(t)\Big\{I(Y_i(t)-\theta^\top Z_i(t)\le 0)-\tau\Big\}dN_i(t), M_i(t)=N_i(t)-\int_0^\top\xi_i(s)\exp(\gamma^\top X_i(s))d\Lambda_0(s),
$$

where

$$
\begin{aligned}
s^{(k)}(t;\gamma) &= \mathrm{E}\Big\{\xi_i(t)X_j(t)^{\otimes k}\exp(\gamma^\top X_j(t))\Big\},\ k=0,1,2,\ \ \bar{x}(t;\gamma)=\frac{s^{(1)}(t;\gamma)}{s^{(0)}(t;\gamma)},\\
\tilde{S}^{(1)}(t;\gamma) &= \frac{1}{n}\sum_{j=1}^{n}\xi_j(t)Z_j(t)\exp(\gamma^\top X_j(t)),\tilde{S}^{(2)}(t;\gamma)=\frac{1}{n}\sum_{j=1}^{n}\xi_j(t)Z_j(t)X_j^\top(t)\exp(\gamma^\top X_j(t)),\\
\tilde{s}^{(1)}(t;\gamma) &= \mathrm{E}\Big\{\xi_j(t)Z_j(t)\exp(\gamma^\top X_j(t))\Big\},\tilde{s}^{(2)}(t;\gamma)=\mathrm{E}\Big\{\xi_j(t)Z_j(t)X_j^\top(t)\exp(\gamma^\top X_j(t))\Big\}.
\end{aligned}
$$

Based on the model assumption, it is always true that $G(\theta_0,\gamma_0,\Lambda_0)=0$ for true θ_0,γ_0 and Λ_0. We assume that there exists a unique θ_0 such that $G(\theta_0,\gamma_0,\Lambda_0)=0$. Generally speaking, condition (A_4) given in Appendix ensures that $G(\theta_0,\gamma_0,\Lambda_0)=0$ has an unique solution.

The following theorem provides sufficient conditions for the consistency of $\widehat{\theta}$.

Theorem 1. *Assume that conditions $(A_1),(A_4),(A_6)-(A_9)$ in Section 6 are satisfied, then $\widehat{\theta}\xrightarrow{\Pr}\theta_0$.*

The asymptotic distribution of $\widehat{\theta}$ is given by the following theorem.

Theorem 2. *Assume that conditions $(A_1)-(A_9)$ in Section 6 are satisfied, then*

$$\sqrt{n}(\widehat{\theta}-\theta_0)\xrightarrow{D}N(0,\Sigma),$$

where $\Sigma=A^{-1}VA^{-1}$ and

$$V=\mathrm{E}\Big[h_i(\theta_0)+\tau\int_0^{t_0}\Big\{Z_i(t)-\bar{z}(t;\gamma_0)\Big\}dM_i(t)-\tau P\Omega^{-1}\int_0^{t_0}\Big\{X_i(t)-\bar{x}(t;\gamma_0)\Big\}dM_i(t)\Big]^{\otimes 2}.$$

4.2 Asymptotic Variance

A direct estimator of Σ can be constructed by the plug-in method. However, it is difficult to estimate the variance of proposed estimator because it involves bandwidth selection in the kernel density estimation of $f_\epsilon(0|X_i(t),\mathcal{F}_{it})$ which is included in A.

An alternative approach to estimating the asymptotic variance is the resampling method of Jin et al. (2001). This method involves the stochastically perturbed estimating equations

$$\tilde{G_n}(\theta,\widehat{\gamma},\widehat{\Lambda}_0(t)) =: \frac{1}{n}\sum_{i=1}^{n}g_i(\theta,\widehat{\gamma},\widehat{\Lambda}_0(t))U_i\approx 0, \tag{12}$$

where U_i $(i=1,\ldots,n)$ which satisfies $\mathrm{var}(U_i)=\{\mathrm{E}(U_i)\}^2$ are i.i.d nonnegative random variables from a known distribution function. Let θ^* be the solution of the perturbation estimating equation. It can be shown that $n^{\frac{1}{2}}(\theta^*-\widehat{\theta})$ has the same limiting distribution as $n^{\frac{1}{2}}(\widehat{\theta}-\theta_0)$ using a method similar to that given in Jin et al. (2001). Thus we can estimate the variance of $\widehat{\theta}$ by the empirical variance of θ^*, which can be obtained by resampling.

5 Simulation Study

In this section, we conducted simulation study to evaluate the performance of proposed quantile estimator. We also demonstrated the advantage of the proposed quantile regression over the existing mean regression when distribution of errors has heavy tails or when the error variances are heteroscedastic.

We considered generating models with two covariates. Firstly, we generated X_{i1} and X_{i2} from the uniform distribution over interval (0,1) and the standard normal distribution for $i = 1, \ldots, n$, respectively. Then, given the values of X_{i1} and X_{i2}, the observation times were generated from the Poisson process $N_i^*(t)$ with intensity rate

$$\mathrm{E}\{dN_i^*(t)|X_i\} = \lambda_0 e^{\gamma_1 X_{i1} + \gamma_2 X_{i2}} dt, \tag{13}$$

where $\gamma_1 = -0.25, \gamma_2 = 0.5, \lambda_0 = 2$. The censoring time C was assumed to follow the uniform distribution over interval $(\eta/2, \eta)$, where η was selected to give the desired censoring rate. Let $\eta = 6$, and the average observation times for different subjects are about 12. The response variable was generated from the regression model

$$Y_i(t) = \beta_1 X_{i1} + \beta_2 X_{i2} + \alpha N_i(t-) + (1 + \sigma X_{i1})\epsilon_i(t), \quad i = 1, \ldots, n,$$

where $\beta_1 = -1, \beta_2 = 1, \alpha = 1.5$. Under above model, the τth conditional quantile of $Y_i(t)$ is

$$Q_\tau(Y_i(t)|X_i, \mathcal{F}_{it}) = \beta_1(\tau)X_{i1} + \beta_2(\tau)X_{i2} + \alpha(\tau)N_i(t-) + Q(\tau), \quad i = 1, \ldots, n, \tag{14}$$

where $\beta_1(\tau) = \beta_1 + \sigma Q(\tau), \beta_2(\tau) = \beta_2, \alpha(\tau) = \alpha$ and $Q(\tau)$ is the τth quantile of ϵ_i.

We considered three cases:

- S_1: homoscedastic normal error model with $\sigma = 0$, $\epsilon_i \sim \mathcal{N}(0, 1)$.
- S_2: homoscedastic heavy tailed model with $\sigma = 0$ and $\epsilon_i \sim \mathcal{C}(0, 1)$.
- S_3: heteroscedastic errors with $\sigma = 1, \epsilon_i \sim \mathcal{N}(0, 0.25)$.

Under S_3, $\beta_1(\tau) = \beta_1 + Q(\tau), \beta_2(\tau) = \beta_2, \alpha(\tau) = \alpha$, where $Q(\tau)$ is the τth quantile of normal distribution $\mathcal{N}(0, 0.25)$. The sample size is $n = 100$. The resampling size for variance estimation is 500.

The estimating equations were solved using the iterative algorithm described in Section 3.2. We also could solve the perturbated estimating equation which defined in Section 4.2 by using the similar iterative method. We used random variables generated from the exponential distribution with mean 1 in the perturbed estimating equations.

Tables 1-3 present the results from the three generating models in cases S_1-S_3 for 0.25th, median and 0.75th. "true" gives the true parameters and "est" reports the estimated parameters. Empirical comparisons are in terms of the magnitude of bias ("bias"), standard deviations of the estimates ("sd"), standard error of estimates ("se"), and the actual confidence interval coverage when the nominal target coverage probability is 0.95 ("cp").

Table 1: Simulation results of S_1 based on 500 replications

	$\tau = 0.25$			$\tau = 0.5$			$\tau = 0.75$		
	$\beta_1(\tau)$	$\beta_2(\tau)$	$\alpha(\tau)$	$\beta_1(\tau)$	$\beta_2(\tau)$	$\alpha(\tau)$	$\beta_1(\tau)$	$\beta_2(\tau)$	$\alpha(\tau)$
true	-1.000	1.000	1.500	-1.0000	1.0000	1.5000	-1.000	1.0000	1.5000
est	-0.9936	0.9765	1.5072	-1.0004	1.0006	1.5006	-1.0129	1.0131	1.4998
bias	-0.0064	-0.0235	0.0072	-0.0004	0.0006	0.0006	-0.0129	0.0131	-0.0002
sd	0.2614	0.1569	0.0148	0.2349	0.1195	0.0121	0.2508	0.1380	0.0135
se	0.2698	0.1547	0.0140	0.2383	0.1222	0.0129	0.2604	0.1486	0.0137
cp	0.9400	0.9380	0.9340	0.9640	0.9560	0.9540	0.9440	0.9720	0.9400

Table 2: Simulation results of S_2 based on 500 replications

	$\tau = 0.25$			$\tau = 0.5$			$\tau = 0.75$		
	$\beta_1(\tau)$	$\beta_2(\tau)$	$\alpha(\tau)$	$\beta_1(\tau)$	$\beta_2(\tau)$	$\alpha(\tau)$	$\beta_1(\tau)$	$\beta_2(\tau)$	$\alpha(\tau)$
true	-1.000	1.000	1.500	-1.0000	1.0000	1.5000	-1.000	1.0000	1.5000
est	-1.0338	1.0013	1.4983	-0.9628	1.0059	1.4982	-0.8702	1.0075	1.5002
bias	-0.0338	0.0013	-0.0017	0.0372	0.0059	-0.0018	0.1298	0.0075	0.0002
sd	0.5170	0.2710	0.0308	0.2984	0.1764	0.0176	0.5550	0.2889	0.0284
se	0.5398	0.2776	0.0306	0.3126	0.1835	0.0187	0.6050	0.3158	0.0331
cp	0.9560	0.9580	0.9480	0.9560	0.9520	0.9580	0.9520	0.9480	0.9680

Table 1 presents simulation results of S_1. From Table 1, we can see that the results of median are better than those of other quantiles. This is not surprising because for the normal error, the effective sample size is larger for the median regression than those for the 0.25 or 0.75 quantile regressions. Table 2 presents simulation results of S_2. Results from the median regression are also better than those from other quantile regressions. The reason is similar to that of S_1. Table 3 reports the simulation results of S_3. Performance of $\beta_1(\tau)$ is worse than those of $\beta_2(\tau)$ and $\alpha(\tau)$. This mainly due to the X_{1i}-related heteroscedastic structure of error.

Table 3: Simulation results of S_3 based on 500 replications

	$\tau = 0.25$			$\tau = 0.5$			$\tau = 0.75$		
	$\beta_1(\tau)$	$\beta_2(\tau)$	$\alpha(\tau)$	$\beta_1(\tau)$	$\beta_2(\tau)$	$\alpha(\tau)$	$\beta_1(\tau)$	$\beta_2(\tau)$	$\alpha(\tau)$
true	-1.3372	1.000	1.500	-1.0000	1.0000	1.5000	-0.6628	1.0000	1.5000
est	-1.3426	0.9828	1.5060	-0.9936	1.0016	1.4995	-0.7111	1.0088	1.4998
bias	-0.0054	-0.0172	0.0060	0.0064	0.0016	-0.0005	-0.0483	0.0088	-0.0002
sd	0.2289	0.1157	0.0109	0.1971	0.0988	0.0097	0.2214	0.1009	0.0097
se	0.2375	0.1104	0.0106	0.1940	0.0934	0.0094	0.2620	0.1134	0.0107
cp	0.9400	0.9280	0.9400	0.9300	0.9400	0.9420	0.9740	0.9780	0.9600

Table 4: Estimated conditional quantile function of response based on 500 replications

	$\tau = 0.25$			$\tau = 0.5$			$\tau = 0.75$		
t	S_1	S_2	S_3	S_1	S_2	S_3	S_1	S_2	S_3
0.5	0.8343	0.5131	1.0199	1.5077	1.5236	1.5071	2.1898	2.5549	2.0040
1.0	2.8554	2.5313	3.0404	3.5277	3.5437	3.5272	4.2087	4.5754	4.0231
1.5	4.8592	4.5502	5.0436	5.5303	5.5641	5.5299	6.2103	6.5961	6.0249
2.0	6.8762	6.5719	7.0601	7.5462	7.5877	7.5458	8.2251	8.6201	8.0400
2.5	8.8899	8.5657	9.0732	9.5587	9.5835	9.5585	10.2365	10.6161	10.0517
3.0	10.8996	10.5691	11.0823	11.5672	11.5885	11.5671	12.2439	12.6213	12.0594
3.5	12.7430	12.4134	12.9252	13.4095	13.4344	13.4095	14.0852	14.4675	13.9009
4.0	14.2516	13.9207	14.4335	14.9173	14.9430	14.9174	15.5921	15.9763	15.4080

Table 5: MSE of mean regression and median regression based on 1000 replications

	S_1		S_2		S_3	
	mean	median	mean	median	mean	median
β_1	0.0515	0.0538	–	0.0924	0.0411	0.0405
β_2	0.0144	0.0158	–	0.0316	0.0100	0.0086
α	0.0002	0.0002	–	0.0003	0.0001	0.0001

From Tables 1 to 3, parameter estimators are very close to true value, none of those estimators exhibit bias of any substantial magnitude, and the actual coverage probabilities are close to the 0.95 nominal level. In addition, the estimated standard errors obtained by resmapling provide good estimates of the sampling variabilities, that is, they are close to the standard deviations calculated based on the replications.

We also evaluated the performance of the estimated conditional quantiles for response process. For each replication, For each replication, we took the average of n conditional quantiles of the response variable at a particular time t as the estimated quantile at t for this replication, and took the average of estimated quantiles for all replications as the finally estimated conditional quantiles at t. Table 4 shows these estimated conditional quantiles of $Y(t)$ at $t = 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4$. From Table 4 again, we can see that, for the same t, quantiles increase as τ becoming larger. For the same τ, quantiles increase as t increasing, it is reasonable because $N(t-)$ is monotonically increase respect to t, and model (14) indicates that larger $N(t-)$ lead to larger quantile.

We also compared the proposed quantile regression estimators with mean regression estimators which are estimated from following marginal conditional mean regression model

$$\mathrm{E}\{Y_i(t)|X_i, \mathcal{F}_{it}\} = \beta_1 X_{i1} + \beta_2 X_{i2} + \alpha N_i(t-), \quad i = 1, \ldots, n,$$

and the observation times were also generated from (13). Table 5 exhibits the mean square errors of the mean regression and median regression for all three cases. In Table 5, “mean” represents the mean regression and "median" represents median regression. For S_1 with normal error terms, the MSEs based on the mean regression is slightly smaller than those based on the median regression. Under S_2, mean regression can not work because the moment of Cauchy distribution does not exist, but median regression still works well. Under S_3, the MSEs of median regression are smaller than those of the mean regression.

On the whole, our quantile estimators are perform very well and quantile regression indeed more robust and efficient than mean regression when the errors are heavy tailed or have heteroscedasticity.

6 Real Example

We apply the proposed method to analyze the bladder cancer data introduced earlier. This data set have been analyzed by Sun & Wei (2000), Sun et al. (2005), among others. There were 85 patients who participated in this study, among them, 47 patients were randomly allocated to the placebo group and the remained 38 patients were in the thiotepa treatment group. During the study, the number of initial bladder tumors, the size of the largest initial tumor, the number of tumors that occurred between clinical visits and the clinical visit times (observation times) were recorded. The unit of the observation times was in months, and the largest observation time was 53 months. The purpose of this study is to evaluate the effect of the thiotepa treatment and clinical visits on the tumor recurrence rates under different quantiles.

In the analysis, for $i = 1, \ldots, 85$, define $Y_i(t)$ as the natural logarithm of the number of observed tumors plus 1 at time t. Here we add 1 to the number of observed tumors to avoid taking logarithm of 0. Let $X_{i1} = 1$ if subject i was in the thiotepa group and 0 if subject i was in the placebo group. Let X_{2i} be the number of initial tumors. Let t_0 be the longest observation time, namely, 53 months. We considered two choices of the function H: $H_1(\mathcal{F}_{it}) = N_i(t-)$ and $H_2(\mathcal{F}_{it}) = N_i(t-) - N_i(t-6)$. Under H_1 and H_2, we evaluated the effect of all clinical visit times to t and the observation times in most recent 6 months on the tumor recurrence rate respectively.

For observation process, we obtained $\widehat{\gamma}_1 = 0.4952, \widehat{\gamma}_2 = -0.0098$, with estimated standard errors of 0.1291 and 0.0346. That means patients in the thiotepa treatment group tend to visit clinic significantly more often than those in the placebo group, and the clinical visit process seems to be unrelated to the initial number of tumors.

Table 6 reports analysis results for $\tau = 0.25, 0.5, 0.75$ and 0.95. In the table, "est" gives the parameter estimates and "se" presents the standard errors. "NIO" shows the results of the case ignoring informative observation times. Note that, both treatment effect and effect of the number of initial tumors significantly affect tumor recurrence rate, when informative observation times are ignored. Under both H_1 and H_2, the treatment does not significantly affect the tumor recurrence rate, under 25th and 95th quantiles, while it is significantly negative related to tumor recurrence rate, under median and 75th quantile. This indicates

Table 6: Bladder cancer data result

τ		NIO $\beta_1(\tau)$	NIO $\beta_2(\tau)$	H_1 $\beta_1(\tau)$	H_1 $\beta_2(\tau)$	H_1 $\alpha(\tau)$	H_2 $\beta_1(\tau)$	H_2 $\beta_2(\tau)$	H_2 $\alpha(\tau)$
0.25	est	0.6003	0.0013	0.2664	0.1959	-0.0369	-0.0228	0.0668	0.1155
	se	0.0671	0.0181	0.1527	0.0080	0.0064	0.1028	0.0172	0.0106
0.5	est	-0.0787	-0.0000	-0.1038	0.1502	-0.0153	-0.6618	0.1128	0.0946
	se	0.0063	0.0029	0.0194	0.0147	0.0025	0.0669	0.0194	0.0165
0.75	est	-1.5689	-0.0004	-1.6473	-0.0027	-0.0007	-0.9846	-0.0047	-0.0804
	se	0.3056	0.0476	0.3709	0.1716	0.0388	0.1683	0.0077	0.0288
0.95	est	-3.6570	-0.0410	-2.9296	0.0176	-0.0441	-0.5483	-0.0866	-0.6471
	se	1.7861	0.3277	2.4870	0.5478	0.2600	0.3961	0.0561	0.0299

Table 7: Estimated conditional quantiles of Y

treatment group $X_1 = 1$

τ	H_1	H_1: $X_2=1$	$X_2=2$	$X_2=3$	$X_2=4$	H_2	H_2: $X_2=1$	$X_2=2$	$X_2=3$	$X_2=4$
0.25	1	-0.2825	-0.2002	-0.1179	-0.0356	1	-0.5775	-0.5107	-0.4439	-0.3771
	2	-0.2728	-0.1905	-0.1082	-0.0259	2	-0.4620	-0.3952	-0.3284	-0.2616
0.5	1	0.1593	0.3095	0.4597	0.6099	1	-0.1847	-0.0719	0.0409	0.1537
	2	0.1440	0.2942	0.4444	0.5946	2	-0.0901	0.0227	0.1355	0.2483
0.75	1	0.7229	0.7202	0.7175	0.7148	1	0.4022	0.3975	0.3928	0.3881
	2	0.7222	0.7195	0.7168	0.7141	2	0.3218	0.3171	0.3124	0.3077
0.95	1	1.7098	1.7274	1.7450	1.7626	1	3.6576	3.5710	3.4844	3.3978
	2	1.6657	1.6833	1.7009	1.7185	2	3.0105	2.9239	2.8373	2.7507

placebo group $X_1 = 0$

τ	H_1	H_1: $X_2=1$	$X_2=2$	$X_2=3$	$X_2=4$	H_2	H_2: $X_2=1$	$X_2=2$	$X_2=3$	$X_2=4$
0.25	1	-0.7637	-0.6814	-0.5991	-0.5168	1	-0.5547	-0.4879	-0.4211	-0.3543
	2	-0.7540	-0.6717	-0.5894	-0.5071	2	-0.4392	-0.3724	-0.3056	-0.2388
0.5	1	0.4046	0.5548	0.7050	0.8552	1	0.4771	0.5899	0.7027	0.8155
	2	0.3893	0.5395	0.6897	0.8399	2	0.5717	0.6845	0.7973	0.9101
0.75	1	2.3702	2.3675	2.3648	2.3621	1	1.3868	1.3821	1.3774	1.3727
	2	2.3695	2.3668	2.3641	2.3614	2	1.3064	1.3017	1.2970	1.2923
0.95	1	4.6394	4.6570	4.6746	4.6922	1	4.2059	4.1193	4.0327	3.9461
	2	4.5953	4.6129	4.6305	4.6481	2	3.5588	3.4722	3.3856	3.2990

that the treatment effect is not significant for the patients, who with lower or higher tumor recurrence rate, and the therapy works well for these patients, who have moderate tumor recurrence rate. It may because patients who with lower recurrence rate, not only tend to have strong resistance to tumor recurrence, but also tend to have strong resistance to therapy. Therapy also does not work well for the patients, with higher recurrence rate, because their weaker resistance to tumor recurrence. For patients with moderate tumor recurrence rate, thiotepa treatment tends to decrease their tumor occurrence rate. Different from the treatment effect, the effect of the number for initial tumors on tumor recurrence rate, is significant under 25th quantile and median, and becomes less significant, when τ increase to 0.75 and 0.95. It implies that, among the patients whose recurrence rate is lower than median, patients with the more number of initial tumors tend to have a higher tumor occurrence rate. For patients whose recurrence rate is higher than median, their numbers of initial tumors seem to be unrelated to their recurrence rate. Note that, the tumor recurrence rate and observation times are negatively correlated, under H_1 and all quantiles, but the effect of observation times on tumor recurrence rate, is significantly only under lower quantiles $\tau = 0.25, 0.5$. This means that the more often the patients visit the clinic, the lower tumor recurrence rate they tend to has on the whole. It is interesting to note that under H_2, the observation times and recurrence rate are positively correlated under 25th quantile and median, but are negatively correlated for $\tau = 0.75$ and 0.95. In other words, among the patients with lower recurrence rate, the more observation times within past six months, the higher tumor recurrence rate they tend to have. Meanwhile, among the patients with higher recurrence rate, increasing of clinical visit times within past six months tends to decrease their tumor recurrence rate. It may because for these lower recurrence rate-patients, they do not need to visit clinic much often in general, but once they found their tumor recurrence increasing, they would visit hospital frequently, within a period of time such as six months. Table 7 presents the fitted conditional quantiles of Y for both treatment and placebo groups, under different numbers of initial tumors and H_1, H_2. From this table, we can see that the fitted quantiles increase monotonically increasing for the increasing sequence of τ.

Different from the mean regression method in Sun et al. (2005), the proposed method could evaluate the effect of treatment, number of initial tumors and observation times under different quantiles, and obtained some interesting phenomenon which could not be found by mean regression.

7 Conditions and Proofs

Throughout this paper, suppose that the true value θ_0 is an interior point of parameter space Θ. Furthermore, suppose that γ belongs to a compact nuisance parameter space $\tilde{\Theta}$. Let $f_{\epsilon_i}(0|X_i(t), \mathcal{F}_{it})$ be the conditional density function of random error ϵ_i. We also require following conditions.

(A_1) $X(t)$ and $H(.)$ are continuous and right continuous, respectively, and have bounded total variation on $[0, t_0]$.

(A_2) The density function $f_\epsilon(0|X_i(t), \mathcal{F}_{it})$ is uniformly bounded away from 0 and infinity.

(A_3) $\mathrm{E}\{\|Z(t)\|^4\} < \infty$ for $t \in [0, t_0]$.

(A_4) $\mathrm{E}\big\{\int_0^{t_0} Z_i(t)Z_i^\top(t) f_\epsilon(Y_i(t) - \theta^\top Z_i(t)|X_i(t), \mathcal{F}_{it}) dN_i(t)\big\}$ is nonsingular in a neighborhood of θ_0.

(A_5) $\mathrm{E}\{h_i(\theta_0)\}^{\otimes 2} < \infty$ and

$$\mathrm{E}\bigg[\int_0^{t_0} \Big\{Z_i(t) - \bar{z}(t;\gamma_0)\Big\}^{\otimes 2} \xi_i(t)\exp(\gamma^\top X_i(t)) d\Lambda_0(t)\bigg] < \infty.$$

(A_6) $\{N_i(.), X_i(.), \xi_i(.)\}, i = 1, \ldots, n$ are i.i.d..

(A_7) $\Pr(C_i \geq t_0) > 0, i = 1, \ldots, n$.

(A_8) $N_i(t_0), i = 1, \ldots, n$ are bounded by a constant.

(A_9) The following matrix is positive definite

$$\mathrm{E}\bigg[\int_0^{t_0} \Big\{X_i(t) - \bar{x}(t;\gamma_0)\Big\}^{\otimes 2} \xi_i(t)\exp(\gamma^\top X_i(t)) d\Lambda_0(t)\bigg].$$

Condition (A_1) is common in the literature on time-varying covariate effect models. Conditions (A_2) and (A_4) are trivial for quantile regression. (A_3) states moment conditions on the covariate $Z(t)$. Condition (A_4) ensures that $G(\theta, \gamma, \Lambda_0(t))$ has the unique zero at the true value θ_0 which is important to identify the unknown parameters. (A_5) and (A_9) are needed to show that the covariance of the estimator $\widehat{\theta}$ is finite. $(A_6) - (A_9)$ are needed to derive the asymptotic normality and weak convergence of $\widehat{\gamma}$ and $\widehat{\Lambda}_0(t)$, respectively. See also Lin et al. (2000) for a discussion on these conditions.

We first present a lemma that is important to the proof of the main theorem.

Lemma 1. *Assume that $(A_1) - (A_3)$ hold, then for all positive values $\varepsilon_n = o(1)$, we have*

$$I_1 =: \sup_{\|\theta-\theta_0\|\leq\varepsilon_n} \|G_n(\theta, \widehat{\gamma}, \widehat{\Lambda}_0(t)) - G(\theta, \gamma_0, \Lambda_0(t)) - G_n(\theta_0, \widehat{\gamma}, \widehat{\Lambda}_0(t))\| = o_p(n^{-1/2}).$$

Proof. It is easy to see that

$$\begin{aligned} I_1 &\leq \sup_{\|\theta-\theta_0\|\leq\varepsilon_n} \|G_n(\theta, \widehat{\gamma}, \widehat{\Lambda}_0(t)) - G_n(\theta, \gamma_0, \Lambda_0(t)) + G_n(\theta_0, \gamma_0, \Lambda_0(t)) - G_n(\theta_0, \widehat{\gamma}, \widehat{\Lambda}_0(t))\| \\ &\quad + \sup_{\|\theta-\theta_0\|\leq\varepsilon_n} \|G_n(\theta, \gamma_0, \Lambda_0(t)) - G(\theta, \gamma_0, \Lambda_0(t)) - G_n(\theta_0, \gamma_0, \Lambda_0(t))\| \\ &=: I_2 + I_3. \end{aligned}$$

For I_2, we have

$$
\begin{aligned}
I_2 &= \sup_{\|\theta-\theta_0\|\le\varepsilon_n} \bigg\| \frac{1}{n}\sum_{i=1}^n \bigg[\int_0^{t_0} Z_i(t)\Big\{ I(Y_i(t)-\theta^\top Z_i(t)<0)dN_i(t) - \tau\xi_i(t)e^{\widehat{\gamma}^\top X_i(t)} d\widehat{\Lambda}_0(t)\Big\} \\
&\quad - \int_0^{t_0} Z_i(t)\Big\{ I(Y_i(t)-\theta^\top Z_i(t)<0)dN_i(t) - \tau\xi_i(t)e^{\gamma_0^\top X_i(t)} d\Lambda_0(t)\Big\} \\
&\quad + \int_0^{t_0} Z_i(t)\Big\{ I(Y_i(t)-\theta_0^\top Z_i(t)<0)dN_i(t) - \tau\xi_i(t)e^{\gamma_0^\top X_i(t)} d\Lambda_0(t)\Big\} \\
&\quad - \int_0^{t_0} Z_i(t)\Big\{ I(Y_i(t)-\theta_0^\top Z_i(t)<0)dN_i(t) - \tau\xi_i(t)e^{\widehat{\gamma}^\top X_i(t)} d\widehat{\Lambda}_0(t)\Big\}\bigg]\bigg\| \\
&= 0.
\end{aligned}
$$

Hence, it is sufficient to prove that $I_3 = o_p(n^{-1/2})$. For simplicity, denote

$$
m_i(\theta) =: \int_0^{t_0} Z_i(t) I(Y_i(t)-\theta^\top Z_i(t)<0)dN_i(t).
$$

Note that

$$
I_3 = \sup_{\|\theta-\theta_0\|\le\varepsilon_n} \bigg\| \frac{1}{n}\sum_{i=1}^n m_i(\theta) - \mathrm{E}m_i(\theta) - \frac{1}{n}\sum_{i=1}^n m_i(\theta_0) + \mathrm{E}m_i(\theta_0)\bigg\|,
$$

by the Lemma 2.17 of Pakes & Pollard (1989), it is sufficient to prove that $\{m_i(\theta), \theta\in\Theta\}$ is an Euclidean class and $m_i(\theta)$ is $L_2(P)$-continuous at θ_0. By Lemma 22 (ii) in Nolan & Pollard (1987) and (A_1), the Euclidean properties of the two classes $\{I(Y_i(t) < \theta^\top Z_i(t)), \theta\in\Theta\}$, $\{Z_i(t), t\in[0,t_0]\}$ with the envelope $F_1 = 1$ and bounded of $Z(t)$ hold because the function $I(\cdot)$ and $Z(t)$ are of bounded variation. Hence, $\{m_i(\theta), \theta\in\Theta\}$ is an Euclidean class for a constant envelope by Lemma 5 in Sherman (1994).

In the following, we verify that $m_i(\theta)$ is $L_2(P)$-continuous at θ_0. Let Z_{ij}, and m_{ij} denote the jth coordinates of Z_i and m_i, respectively. For simplicity, we omit the subscript i in these expressions such as $Z_{ij}, m_{ij}, Y_i, N_i(t), \mathcal{F}_{it}$ and the observation points $t_{ij}, j = 1,\ldots,n_i$. We just need to prove that for all $\theta\in\Theta$, there exist constants $K_j > 0$ such that

$$
I_4 =: \sup_{\|\theta-\theta_0\|\le\varepsilon_n} \mathrm{E}\{m_j(\theta) - m_j(\theta_0)\}^2 \le K_j\varepsilon_n. \tag{15}
$$

By simple calculation, we obtain

$$
\begin{aligned}
I_4 &= \sup_{\|\theta-\theta_0\|\le\varepsilon_n} \mathrm{E}\bigg[\int_0^{t_0} Z_j(t)\Big\{ I(Y(t)-\theta^\top Z(t)<0) - I(Y(t)-\theta_0^\top Z(t)<0)\Big\} dN(t)\bigg]^2 \\
&= \sup_{\|\theta-\theta_0\|\le\varepsilon_n} \mathrm{E}\bigg[\mathrm{E}\bigg[\bigg[\int_0^{t_0} Z_j(t)\Big\{ I(Y(t)-\theta^\top Z(t)<0) \\
&\quad -I(Y(t)-\theta_0^\top Z(t)<0)\Big\} dN(t)\bigg]^2 \bigg| X_i(t), \mathcal{F}_t\bigg]\bigg]
\end{aligned}
$$

by Holder inequality, (A_2) and (A_3), we have

$$\begin{aligned}
I_4 &\leq \sup_{\|\theta-\theta_0\|\leq\varepsilon_n} \mathrm{E}\Big[\mathrm{E}\Big[\Big[\int_0^{t_0} Z_j^2(t)\Big\{I(Y(t)-\theta^\top Z(t)<0) \\
&\quad -I(Y(t)-\theta_0^\top Z(t)<0)\Big\}dN(t)\Big]\Big|X_i(t),\mathcal{F}_t\Big]\Big] \\
&\leq \sup_{\|\theta-\theta_0\|\leq\varepsilon_n} \int_0^{t_0}\Big[\{\mathrm{E}Z_j^4(t)\}^{\frac{1}{2}}\mathrm{E}^{\frac{1}{2}}\{\big|F_{Y|X,\mathcal{F}_t}(\theta^\top Z(t)+\varepsilon_n)-F_{Y|X,\mathcal{F}_t}(\theta_0^\top Z(t)-\varepsilon_n)\big|\}\Big]dN(t) \\
&\leq K_j\varepsilon_n,
\end{aligned}$$

where $K_j<\infty$. Thus, the proof of this lemma is completed. $\square$

Proof of Theorem 1. By Corollary 3.2 of Pakes & Pollard (1989) and the definition of the estimator $\widehat{\theta}$, we just need to verify that conditions (ii) and (iii) in that corollary hold. In fact, by (A_4), we have

$$\begin{aligned}
\inf_{\|\theta-\theta_0\|>\eta}\|G(\theta,\gamma_0,\Lambda_0(t))\| &= \inf_{\|\theta-\theta_0\|>\eta}\|G(\theta,\gamma_0,\Lambda_0(t))-G(\theta_0,\gamma_0,\Lambda_0(t))\| \\
&= \inf_{\|\theta-\theta_0\|>\eta}\Big\|\mathrm{E}\Big[\int_0^{t_0} Z_i(t)\Big\{I(Y_i(t)-\theta^\top Z_i(t)<0) \\
&\quad -I(Y_i(t)-\theta_0^\top Z_i(t)<0)\Big\}dN_i(t)\Big]\Big\| \\
&= \mathrm{E}\Big\{\int_0^{t_0} Z_i(t)Z_i^\top(t)f_\epsilon(Y_i(t)-\tilde{\theta}^T Z_i(t)|X_i(t),\mathcal{F}_{it})dN_i(t)\Big\} \\
&\quad \times \inf_{\|\theta-\theta_0\|>\eta}\|\theta-\theta_0\|>0,
\end{aligned}$$

where $\tilde{\theta}$ lies between θ and θ_0. Hence, condition (ii) of Corollary 3.2 of Pakes & Pollard (1989) is satisfied. Let's verify the last condition, namely

$$I_5 =: \sup_{\theta\in\Theta}\|G_n(\theta,\widehat{\gamma},\widehat{\Lambda}_0(t))-G(\theta,\gamma_0,\Lambda_0(t))\| = o_p(1).$$

Note that

$$\begin{aligned}
I_5 &\leq \sup_{\theta\in\Theta}\|G_n(\theta,\widehat{\gamma},\widehat{\Lambda}_0(t))-G_n(\theta,\gamma_0,\Lambda_0(t))\|+\sup_{\theta\in\Theta}\|G_n(\theta,\gamma_0,\Lambda_0(t))-G(\theta,\gamma_0,\Lambda_0(t))\| \\
&=: I_6+I_7.
\end{aligned}$$

For I_6, we have

$$\begin{aligned}
I_6 &= \tau\Big\|\frac{1}{n}\sum_{i=1}^n\Big\{\int_0^{t_0} Z_i(t)\xi_i(t)e^{\widehat{\gamma}^\top X_i(t)}d\widehat{\Lambda}_0(t)-\int_0^{t_0} Z_i(t)\xi_i(t)e^{\gamma_0^\top X_i(t)}d\Lambda_0(t)\Big\}\Big\| \\
&\leq \tau\Big\|\frac{1}{n}\sum_{i=1}^n\int_0^{t_0} Z_i(t)\xi_i(t)e^{\gamma_0^\top X_i(t)}d\{\widehat{\Lambda}_0(t)-\Lambda_0(t)\}\Big\| \\
&\quad +\tau\Big\|\frac{1}{n}\sum_{i=1}^n\int_0^{t_0} Z_i(t)\xi_i(t)\{e^{\widehat{\gamma}^\top X_i(t)}-e^{\gamma_0^\top X_i(t)}\}d\Lambda_0(t)\Big\| \\
&\quad +\tau\Big\|\frac{1}{n}\sum_{i=1}^n\int_0^{t_0} Z_i(t)\xi_i(t)\{e^{\widehat{\gamma}^\top X_i(t)}-e^{\gamma_0^\top X_i(t)}\}d\{\widehat{\Lambda}_0(t)-\Lambda_0(t)\}\Big\| \\
&=: \tau(I_{61}+I_{62}+I_{63}).
\end{aligned}$$

By the Taylor expansion of $\widehat{\Lambda}_0(t)$ at γ_0 (see Sun & Wu (2005)), we obtain

$$\widehat{\Lambda}_0(t)-\Lambda_0(t)=\frac{1}{n}\sum_{i=1}^{n}\int_0^{\top}\frac{dM_i(u)}{s^{(0)}(u;\gamma_0)}-\int_0^{\top}\bar{x}(u;\gamma_0)d\Lambda_0(u)(\widehat{\gamma}-\gamma_0)+o_p(1),$$

hence, we have

$$\begin{aligned}
I_{61} &= \left\|\frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0}\frac{\tilde{S}^{(1)}(t;\gamma_0)}{s^{(0)}(t;\gamma_0)}dM_i(t)-(\widehat{\gamma}-\gamma_0)\int_0^{t_0}\tilde{S}^{(1)}(t;\gamma_0)\bar{x}(u;\gamma_0)d\Lambda_0(t)+o_p(1)\right\| \\
&\le \left\|\frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0}\frac{\tilde{S}^{(1)}(t;\gamma_0)}{s^{(0)}(t;\gamma_0)}dM_i(t)\right\|+\|\widehat{\gamma}-\gamma_0\|\left\|\int_0^{t_0}\tilde{S}^{(1)}(t;\gamma_0)\bar{x}(u;\gamma_0)d\Lambda_0(t)\right\|+o_p(1),
\end{aligned}$$

by Lin et al. (2000), we know that $\|\widehat{\gamma}-\gamma_0\|=O_p(n^{-\frac{1}{2}})$, hence the second term of I_{61} is $O_p(n^{-\frac{1}{2}})$. By Martingale Central Limit Theorem and the Continuous Mapping Theorem, it is easy to get that the first term of I_{61} is $O_p(n^{-\frac{1}{2}})$, hence, we can get that $I_{61}=o_p(1)$.

By Taylor expansion and $\widehat{\gamma}-\gamma_0=O_p(n^{-\frac{1}{2}})$, it easy to get that $I_{62}=o_p(1)$.

For I_{63}, by Taylor expansion of $e^{\widehat{\gamma}^{\top}X_i(t)}$ at γ_0, we have

$$\begin{aligned}
I_{63} &= \|(\widehat{\gamma}-\gamma_0)+o_p(\widehat{\gamma}-\gamma_0))\|\left\|\frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0}Z_i(t)X_i^{\top}(t)\xi_i(t_0)e^{\gamma_0^{\top}X_i(t)}d\{\widehat{\Lambda}_0(t)-\Lambda_0(t)\}\right\| \\
&= \|(\widehat{\gamma}-\gamma_0)+o_p(\widehat{\gamma}-\gamma_0))\|\left\|\int_0^{t_0}\left\{\frac{1}{n}\sum_{i=1}^{n}Z_i(t)X_i^{\top}(t)\xi_i(t_0)e^{\gamma_0^{\top}X_i(t)}\right\}d\{\widehat{\Lambda}_0(t)-\Lambda_0(t)\}\right\| \\
&= \|(\widehat{\gamma}-\gamma_0)+o_p(\widehat{\gamma}-\gamma_0))\|\left\|\int_0^{t_0}\tilde{S}^{(2)}(t;\gamma_0)d\{\widehat{\Lambda}_0(t)-\Lambda_0(t)\}\right\| \\
&= \|(\widehat{\gamma}-\gamma_0)+o_p(\widehat{\gamma}-\gamma_0))\|\left\|\frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0}\frac{\tilde{S}^{(2)}(t;\gamma_0)}{s^{(0)}(t;\gamma_0)}dM_i(t)\right. \\
&\quad \left.-(\widehat{\gamma}-\gamma_0)\int_0^{t_0}\tilde{S}^{(2)}(t;\gamma_0)\bar{x}(u;\gamma_0)d\Lambda_0(t)+o_p(1)\right\| \\
&\le \|(\widehat{\gamma}-\gamma_0)+o_p(\widehat{\gamma}-\gamma_0))\|\left\{\left\|\frac{1}{n}\sum_{i=1}^{n}\int_0^{t_0}\frac{\tilde{S}^{(2)}(t;\gamma_0)}{s^{(0)}(t;\gamma_0)}dM_i(t)\right\|\right. \\
&\quad \left.+\|(\widehat{\gamma}-\gamma_0)\|\left\|\int_0^{t_0}\tilde{S}^{(2)}(t;\gamma_0)\bar{x}(u;\gamma_0)d\Lambda_0(t)\right\|+o_p(1)\right\},
\end{aligned}$$

Again using $\|\widehat{\gamma}-\gamma_0\|=O_p(n^{-\frac{1}{2}})$, the Martingale Central Limit Theorem and the Continuous Mapping Theorem, it is easy to obtain $I_{63}=o_p(1)$. This means that $I_6=o_p(1)$. At last, we just need to verify that $I_7=o_p(1)$.

For simplicity, denote

$$\tilde{g}_i(\theta,\gamma,\Lambda_0,t)=\int_0^{t}Z_i(s)\Big\{I(Y_i(s)-\theta^{\top}Z_i(s)<0)dN_i(s)-\tau\xi_i(s)e^{\gamma^{\top}X_i(s)}d\Lambda_0(s)\Big\},$$

to prove $I_7=o_p(1)$, according to the Lemma (2.8) of Packs & Pollard (1989), it is sufficient to prove that the class $\{\tilde{g}_i(\theta,\gamma,\Lambda_0,t),t\in[0,t_0]\}$ is the Euclidean class with envelop F which has finite expectation. Similar to the proof of Lemma 1, by Lemma 22 (ii) in Nolan & Pollard (1987), it can be shown that $\{\xi_i(t), t \in [0, t_0]\}$ and $\{e^{\gamma^\top X_i(t)}, \gamma \in \tilde{\Theta}, t \in [0, t_0]\}$ are Euclidean classes for envelope $F = 1$ and envelope $\{\sup_{\gamma\in\tilde{\Theta},t\in[0,t_0]} e^{\gamma^\top X_i(t)}\}$ respectively because indicator function has bounded variation and $e^{\gamma^\top X_i(t)}$ is Lipschitz continuous when γ belongs to a compact parameter space $\tilde{\Theta}$. Therefore, by Lemma 5 in Sherman (1994), $\{\tilde{g}_i(\theta, \gamma, \Lambda_0, t), t \in [0, t_0]\}$ is a Euclidean class.

Therefore, all of conditions in Corollary 3.2 of Pakes & Pollard (1989) hold, and we have completed the proof of Theorem 1. □

Proof of Theorem 2. We verify the conditions of Theorem 3.3 in Pakes & Pollard (1989). By (A_4), Lemma 1 and definition of the estimator, it is easy to see that conditions (i)-(iii) and (v) of Theorem 3.3 of Pakes & Pollard (1989) hold. Here, it is sufficient to prove that condition (iv) is satisfied.

Note that

$$
\begin{aligned}
\sqrt{n}G_n(\theta_0, \widehat{\gamma}, \widehat{\Lambda}_0(t)) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \int_0^{t_0} Z_i(t)\Big\{I(Y_i(t) - \theta_0^\top Z_i(t) < 0)dN_i(t) - \tau\xi_i(t)e^{\widehat{\gamma}^\top X_i(t)}d\widehat{\Lambda}_0(t)\Big\} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n \int_0^{t_0} Z_i(t)\Big\{I(Y_i(t) - \theta_0^\top Z_i(t) < 0) - \tau\Big\}dN_i(t) \\
&\quad +\tau\frac{1}{\sqrt{n}}\sum_{i=1}^n \Big\{\int_0^{t_0} Z_i(t)dN_i(t) - \int_0^{t_0} \xi_i(t)Z_i(t)e^{\widehat{\gamma}^\top X_i(t)}d\widehat{\Lambda}_0(t)\Big\} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n h_i(\theta_0) + \tau\frac{1}{\sqrt{n}}\sum_{i=1}^n \int_0^{t_0} \Big\{Z_i(t) - \frac{\tilde{S}^{(1)}(t;\widehat{\gamma})}{S^{(0)}(t;\widehat{\gamma})}\Big\}dN_i(t) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n h_i(\theta_0) + \tau\frac{1}{\sqrt{n}}\sum_{i=1}^n \int_0^{t_0} \Big\{Z_i(t) - \frac{\tilde{S}^{(1)}(t;\gamma_0)}{S^{(0)}(t;\gamma_0)}\Big\}dN_i(t) \\
&\quad -\tau P\sqrt{n}(\widehat{\gamma} - \gamma_0) + o_p(1) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n h_i(\theta_0) + \tau\frac{1}{\sqrt{n}}\sum_{i=1}^n \int_0^{t_0} \Big\{Z_i(t) - \bar{Z}(t;\gamma_0)\Big\}dM_i(t) \\
&\quad -\tau P\Omega^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n \int_0^{t_0} \Big\{X_i(t) - \bar{X}(t;\gamma_0)\Big\}dM_i(t) + o_p(1),
\end{aligned}
$$

hence, by central limit theorems, we have

$$
\sqrt{n}G_n(\theta_0, \widehat{\gamma}, \widehat{\Lambda}_0(t)) \xrightarrow{D} N(0, V),
$$

where

$$
V = \mathrm{E}\Big[h_i(\theta_0) + \tau\int_0^{t_0} \Big\{Z_i(t) - \bar{z}(t;\gamma_0)\Big\}dM_i(t) - \tau P\Omega^{-1}\int_0^{t_0} \Big\{X_i(t) - \bar{x}(t;\gamma_0)\Big\}dM_i(t)\Big]^{\otimes 2}.
$$

It means that the condition (iv) of Theorem 3.3 of Pakes & Pollard (1989) hold, hence, by Theorem 3.3 in Pakes & Pollard (1989), we can derive the conclusion. This completes the proof of the theorem. □

Reference

Andersen, P. K., Gill, R. D. (1982), Cox's Regression Model for Counting Processes: A Large Sample Study. *Ann. Statist.*, **10**, 1100-1120.

Chen, X. H., Linton, O., Keilegom, I. V. (2003), Estimation of Semiparametric Models When Criterion Function is Not Smooth. *Econometrica*, **71**, 1591-1608.

Cox D. R.(1972), Regression Models and Life-Tables, *J. Roy. Stat. Soc. B.* ,**34**, 187-220.

Diggle, P. J., Liang, K. Y., Zeger, S. L. (1994). *Analysis of Longitudinal Data.* Oxford University Press, New York.

Fleming, T.R., Harrington, D.P. (1991) *Counting Processes and Survival Analysis.* Wiley, New York.

Fenske, N., Fahrmeir, L., Rzehak, P., Höhle, M. (2008). Detection of Risk Factors for Obesity in Early Childhood with Quantile Regression Methods for Longitudinal Data. Technical Report Number 038, Department of Statistics University of Munich.

Hansen, L. P.(1982). Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica*, **50**, 1029-1054.

He, X., Fu, B., Fung, W. K.(2003). Median Regression for Longitudinal Data, *Statist. Med.*, **22**, 3655-3669.

Honoré, B., Khan, S., Powell, J. L.(2002). Quantile Regression under Random Censoring, *J. Econometrics*, **109**, 67-105.

Hoover, D. R., Rice, J. A., Wu, C. O., Yang, L. P. (1998). Nonparametric Smoothing Estimates of Time-Varying Coefficient Models with Longitudinal data. Biometrika 85, 809-822.

Hunter D. R., Lange, K.(2000). Quantile Regression via an MM Algorithm. *J. Comput. Graph. Stat.*, **9**, 60-77.

Jing, Z. Z., Ying, Z. L., Wei, L. J.(2001). A Simple Resampling Method by Perturbing the Minimand. *Biometrika*, **88**, 381-390.

Karlsson, A.(2005). Nonlinear Quantile Regression Estimation of Longitudinal Data. *Commun. Stat. Simulat.*, **37**, 114-131.

Koenker, R., Bassett, G.S, (1978). Regression Quantiles. *Econometrica*, **46**, 33-50.

Koenker, R.(2004). Quantile Regression for Longitudinal Data. *J. Multivariate. Anal.*, **91**, 74-89.

Koenker, R.(2005). *Quantile Regression.* Cambbridge University Press, New York.

Kosorok, M. R.(2006). *Introduction to Empirical Processes and Semiparametric Inference.* Springer Press, New York.

Laird, N. M., Ware, J. H.(1982). Random-Effects Models for Longitudinal Data. *Biometrics*, **38**, 963-974.

Liang, H., Wu, H., Carroll, R. J. (2003). The Relationship Between Virologic and Immunologic Responses in AIDS Clinical Research Using Mixed-Effects Varying-Coefficient Semiparametric Models with Measurement Error. *Biostatistics*, **4**, 297-312.

Lin, D. Y., Wei, L. J., Yang, I., Ying, Z. (2000). Semiparametric Regression for the Mean and Rate Functions of Recurrent Events. *J. Roy. Stat. Soc. B.*, **62**, 711-730.

Lin, D. Y., Wei, L. J., Ying, Z. (2001). Semiparametric Transformation Models for Point Processes. *J. Amer. Statist. Assoc.*, **96**, 620-628.

Lin, D. Y., Ying, Z. (2001). Semiparametric and Nonparametric Regression Analysis of Longitudinal Data. *J. Amer. Statist. Assoc.*, **96**, 103-126.

Liu, Y. Bottai, M.(2009). Mixed-Effects Models for Conditional Quantiles with Longitudinal Data. *Int. J. Biostat.*, **5**, Art. 28.

Martinussen, T., Scheike, T. H. (1999). A Semiparametric Additive Regression Model for Longitudinal Data. *Biometrika*, **86**, 691-702.

Martinussen, T., Scheike, T. H. (2000). A Nonparametric Dynamic Additive Regression Model for Longitudinal Data. *Ann. Statist.*, **28**, 1000-1025.

Martinussen, T., Scheike, T. H. (2001). Sampling Adjusted Analysis of Dynamic Additive Regression Models for Longitudinal Data. *Scand. J. Statist.*, **28**, 303-323.

Moyeed, R. A., Diggle, P. J. (1994). Rates of Convergence in Semiparametric Modelling of Longitudinal Data. *Aust. J. Statist.*, **36**, 75-93.

Mu, Y., Wei, Y. (2009). A Dynamic Quqntile Regression Transformation Model for Longitudinal Data. *Stat. Sinica*, **19**, 1137-1153.

Nolan, D. and Pollard, D., (1987). U-Processes: Rates of Convergence. *Ann. Statist.*, **15**, 780-799.

Pakes, A., Pollard, D.(1989). Simulation and the Asymptotics of Optimization Estimators. *Econometrica*, **57**, 1027-1057.

Scheike, T. H., Zhang, M. J. (1998). Cumulative Regression Function Tests for Longitudinal Data. *Ann. Statist.*, **26**, 1328-1355.

Sherman, R. P.(1994). Maximal Inequalities for Degenerate U-Processes with Applications to Optimization Estimators. *Ann. Statist.*, **22**, 439-459.

Sun, Y., Wu, H. (2003). AUC-Based Tests for Nonparametric Functions with Longitudinal Data. *Statist. Sinica*, **13**, 593-612.

Sun, Y., Wu, H. (2005). Semiparametric Time-Varying Coefficients Regression Model for Longitudinal Data. *Scand. J. Statist.*, **32**, 21-47.

Sun, J., Park, D. Sun, L., Zhao, X. (2005). Semiparametric Regression Analysis of Longitudinal Aata with Informative Observation Times. *J. Amer. Statist. Assoc.*, **100**, 882-889.

Sun, J., Sun, L., Liu, D. (2007). Regression Analysis of Longitudinal Data in the Presence of Informative Observation and Censoring Times. *J. Amer. Statist. Assoc.*, **102**, 1397-1406.

Sun, J., Wei, L. J.(2000). Regression Analysis fo Panel Count Data with Covariate-Dependent Observation and Censoring Times. *J. Roy. Stat. Soc. B.*, **62**, 293-302.

Van der Vaart, A. W., Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer, New York.

Van der Vaart, A. W.(1998). *Asymptotic Statistics.* Cambridge University Press, New York.

Wang, H. X., Fygenson, M.(2009). Inference for Censored Quantile Regrssion Models in Longitudinal Studies. *Ann. Statist.*, **37**, 756-781.

Wang, H. X., Zhu, Z. Y., and Zhou, J. H.(2009). Quantile Regression in Partially Linear Varying Coefficient Models. *Ann. Statist.*, **37**, 3841-3866.

Wang, H. X., Zhu, Z. Y. (2011). Empirical Likelihood for Quantile Regression Models with Longitudinal Data. *J. Stat. Plan. Infer.*, **141**, 1603-1615.

Wei, Y., He, X. (2006). Disscussion Paper Conditional Growth Charts. *Ann. Statist.*, **34**, 2069-2097.

Wu, C. O., Chiang, C. T., Hoover, D. R. (1998). Asymptotic Confidence Regions for Kernel Smoothing of a Time-Varying Coefficient Model with Longitudinal Data. *J. Amer. Statist. Assoc.*, **88**, 1388-1402.

Wu, H., Liang, H. (2004). Random Varying-Coefficient Models with Smoothing Covariates, Applications to an AIDS Clinical Study. *Scand. J. Statist.*, **31**, 3-19.

Wu, H., Zhang, J. T. (2002a). The Study of Long-Term HIV Dynamics using Semiparametric Nonlinear Mixed-Effects Models. *Statist. Med.*, **21**, 3655-3675.

Wu, H., Zhang, J. T. (2002b). Local Polynomial Mixed-Effiects Models for Longitudinal Data. *J. Amer. Statist. Assoc.*, **97**, 883-897.

Ying, Z., Jung, S. H. and Wei, L. J.(1995). Survival Analysis with Median Regression Models, *J. Amer. Statist. Assoc.*, **90**, 174-184.