Statistics and Its Interface Volume 10 (2017) 239–254

Power-transformed linear quantile regression estimation for censored competing risks data

Caiyun Fan, Feipeng Zhang*, and Yong Zhou

This paper considers a power-transformed linear quantile regression model for censored competing risks data, based on conditional quantiles defined by using the cumulative incidence function. We propose a two-stage estimating procedure for the regression coefficients and the transformation parameter. In the first step, for a given transformation parameter, we develop an unbiased monotone estimating equation for regression parameters in the quantile model, which can be solved by minimizing a L_1 type convex objective function. In the second step, the transformation parameter can be estimated by constructing the cumulative sum processes. The consistency and asymptotic normality of the regression parameters and transformation parameter are derived. The finite-sample performances of the proposed approach are illustrated by simulation studies and an application to the follicular type lymphoma data set.

AMS 2000 subject classifications: Primary 62G08, 62N01, 62N02, 62P10; secondary 60G07, 62G07, 62G20.
Keywords and phrases: Box-Cox transformation, Censored data, Competing risks, Quantile regression.

1. INTRODUCTION

In time-to-event data, it is often observed that each individual under study may be at risk of at least two mutually exclusive failure types. Such data are commonly referred to as competing risks data. As an example, in the study of follicular cell lymphoma conducted by the Princess Margaret Hospital Pintilie [24], patients with early stage disease (I or II) registered for treatment and then were followed-up until treatment failure, which comprised two different types: relapse (such as local, distant or both) of lymphoma and non-relapse-related deaths. Here, relapse and non-relapse deaths are competing risks events.

Compared with classic survival analysis for only one type of failure, the analysis for competing risks is more challenging. One major difficulty is the identifiability crisis when the censoring time is dependent on the competing risk events. [26] showed that a joint distribution for the time of different failure types cannot be estimated without making strong unverified assumptions. Besides, the marginal distribution of the latent event time has problematic interpretations when hypothesizing the removal of other competing risks events [12]. Another difficulty is that the informative censoring induced by the potential dependence further complicates the analysis of failure times with competing risks.

For competing risks data, it is of great interest to estimate the cumulative incidence function, i.e., the cumulative probability of a specific failure type. Most of the prior work on analyzing the effects of factors for competing risks has focused on examining their effects on the cause-specific hazard functions of the different failure types [16, 25]. However, many authors have noted that the effect of a covariate on the cause-specific hazard function of a particular failure type may be very different from the effect of the covariate on the corresponding cumulative incidence function [11, 22]. As [2, 9, 15] pointed out, the cumulative incidence function provides information secondary to the overall survival function. Consequently, the cumulative incidence function is intuitively appealing, well suited to graphical display and cost-effective in analyzing the absolute risks of the different failure types. [7] introduced a semiparametric proportional hazards model for the cumulative incidence function of a competing risk. Nevertheless, their model not only requires that the regression coefficients are monotonously linked to the cumulative incidence function, but also requires that the occurrence of competing events has an influence on the coefficients.

In many applications, another alternative and flexible model, so-called the quantile regression model, can be chosen to model complicated effect patterns. [20] firstly developed a nonparametric quantile definition based on the cumulative incidence function for competing risks data. Based on the same idea, [21] proposed a formulation of competing risks quantile regression model with a known monotone link function by assuming the independence between the censoring time and covariates. However, the monotone link function is always unknown and the covariate-independent censoring assumption is too stringent in practice. The primary purpose of this paper is to develop a novel and flexible qunatile regression model for competing risks data, which can also be applied to the case that the censoring time depends on the covariates.

As a natural and powerful tool, it is well-known that the transformation model extensively includes a large and important class of modeling structures [4, 5, 13]. Among those,

*Corresponding author.

the parsimonious and interpretable Box-Cox transformation [3] is often used to improve the error normality in linear models. [18] proposed a linear quantile regression model with power transformation for complete data, which may have great potential in various applications. [28] proposed a power-transformed linear quantile regression model for randomly and independently censored survival data. However, to the best of our knowledge, there is no literature on power-transformed linear quantile regression model with the competing risks data. In such situation, the existing approaches may not or not appropriately be used to analyze competing risks data. Moreover, the theoretical and computational developments for quantile regression become much more involved and challenging.

In this paper, we propose a new power-transformed quantile regression model for censored competing risks data. For both covatiate-independent censoring and covariate-dependent censoring, we propose a two-stage procedure for the regression coefficients and the transformation parameter. Specially, in the first step, for a given transformation parameter, we develop an unbiased monotone estimating equation for regression parameters in the quantile model, which can be solved by minimizing a L_1 type convex objective function. In the second step, the transformation parameter can be estimated by constructing the cumulative sum processes. By utilizing the empirical process theory, we can derive that the proposed estimator is uniformly consistent and asymptotically normal. We also provide the explicit form for the variance-covariance matrix of the estimators with the unknown of the error term, and use a bootstrap resampling method for estimating the covariance-matrix to lighten the computational burden. In our empirical studies, we observe that our model has better performances than that of [21], especially when the censoring variable is dependent on the covariates.

The reminder of the paper is organized as follows. Section 2 proposes the estimation procedure for transformation quantile regression model with censored competing risks data. The large sample properties of the proposed estimates are given in Section 3, and simulation studies are presented in Section 4. In Section 5, we illustrate the proposed method with application to a follicular type lymphoma data set. Section 6 concludes the paper with some discussions. The technical proofs are given in the Appendix.

2. TRANSFORMED QUANTILE REGRESSION VIA THE WEIGHTED ESTIMATING EQUATION

We first introduce some necessary notations. Let T and C be the failure and censoring time, respectively, $\widetilde{\boldsymbol{Z}}$ be a $p \times 1$ vector and $\mathbf{Z} = (1, \widetilde{\mathbf{Z}}^T)^T$, and $\epsilon \in \{1, ..., K\}$ be the cause of failure. The observations consist of n independently and identically distributed replicates of $(X, \delta\epsilon, \mathbf{Z})$, denoted by $\{(X_i, \delta_i\epsilon_i, \mathbf{Z}_i), i = 1, \cdots, n\}$, where $X = \min(T, C)$ and $\delta = I(T \le C)$. The cumulative incidence function of cause k, is defined by $F_k(t|\mathbf{Z}) = \Pr(T \le t, \epsilon = k|\mathbf{Z})$, $k = 1, ..., K$. Indeed, $F_k(t|\mathbf{Z})$ represents the probability of observing cause k failure in the presence of other types of events given covariate $\mathbf{Z}$. It is commonly assumed that the censoring variable C is conditionally independent of (T, ϵ) given $\mathbf{Z}$. Let $G_0(t|\mathbf{Z}) = Pr(C > t|\mathbf{Z})$ be the conditional survival function of C given covariates $\mathbf{Z}$.

For $0 < \tau \le P(\epsilon = k)$, the τth conditional quantile of failure time T caused by the kth risk is defined by using cumulative incidence function [21]:

$$Q_k(\tau|\mathbf{Z}) = \inf\{t : F_k(t|\mathbf{Z}) \ge \tau\}, \; k = 1, ..., K.$$

Here, $Q_k(\tau|\mathbf{Z})$ can be easily interpreted as the first time given covariates $\mathbf{Z}$ at which the probability of cause k failure having occurred exceeds τ, in the presence of other events which can exclude occurrence of cause k failure. Without loss of generality, we only focus on $Q_1(\tau|\mathbf{Z})$ in the sequel. If we let $T_1^* = T \times I(\epsilon = 1) + \infty \times I(\epsilon \ne 1)$, it is easy to see that $F_1(t|\mathbf{Z}) = Pr(T_1^* \le t|\mathbf{Z}) = \Pr(T \le t, \epsilon = 1|\mathbf{Z})$. In other words, $F_1(t|\mathbf{Z})$ can be viewed as the cumulative distribution of the variable T_1^*.

[21] formulated a competing risks quantile regression model by assuming

$$Q_1(\tau|\mathbf{Z}) = g\{\boldsymbol{\beta}_0(\tau)^T\mathbf{Z}\}, \quad \text{for } \tau \in [\tau_L, \tau_U],$$

where $\boldsymbol{\beta}_0(\tau)$ is a $(p+1) \times 1$ vector of unknown regression coefficients, $g(\cdot)$ is a known monotone link function, and $0 < \tau_L \le \tau_U \le P(\epsilon = k)$. Despite the model is flexible, it requires specifying the monotone transformation $g(\cdot)$. To relax this stringent assumption, we propose the power-transformed linear quantile regression model

$$Q_1(\tau|\mathbf{Z}) = H_{\gamma_0}^{-1}\left(\boldsymbol{\beta}_0(\tau)^T\mathbf{Z}\right), \tag{1}$$

for $\tau \in [\tau_L, \tau_U]$, where $\boldsymbol{\beta}_0(\cdot)$ is a $(p+1) \times 1$ vector of unknown regression parameters, γ_0 is an unknown transformation parameter. Here, $H_{\gamma_0}^{-1}$ is the inverse of the Box-Cox transformation [3]

$$H_{\gamma_0}(T) = \begin{cases} (T^{\gamma_0} - 1)/\gamma_0, & \text{for} \quad \gamma_0 > 0 \\ \log(T), & \text{for} \quad \gamma_0 = 0. \end{cases}$$

Note that the Box-Cox transformation is often used to improve the error normality in linear models.

In this paper, we propose a two-stage estimation procedure. In the first step, we estimate $\boldsymbol{\beta}_0(\tau)$ for a given γ_0. For uncensored data, the transformed linear quantile regression parameters can be estimated by minimizing

$$\sum_{i=1}^{n} \rho_\tau\left(H_{\gamma_0}(T_{1i}^*) - \boldsymbol{b}^T\mathbf{Z}_i\right), \tag{2}$$

where $\rho_\tau(u) = u\{\tau - I(u \le 0)\}$ is the so-called check function. Note that the minimizer to (2) is equivalent to the root of the following estimating equations:

$$\sum_{i=1}^{n} \mathbf{Z}_i \left\{ I\left(H_{\gamma_0}(T_{1i}^*) \leq \boldsymbol{b}^T \mathbf{Z}_i \right) - \tau \right\} = 0.$$

Recall that when the γ_0 fixed, [21] derived the following estimating equation for censored competing risks data,

$$\begin{aligned} (3) \qquad & S_n(\boldsymbol{b}, \tau) \\ = \ & \sum_{i=1}^{n} \mathbf{Z}_i \left\{ \frac{I\left\{X_i \leq H_{\gamma_0}^{-1}(\boldsymbol{b}^T\mathbf{Z}_i)\right\} I(\delta_i \epsilon_i = 1)}{\widehat{G}(X_i|\mathbf{Z}_i)} - \tau \right\} \\ = \ & 0, \end{aligned}$$

where $\widehat{G}(\cdot|\mathbf{Z}_i)$ is a reasonable estimate for $G_0(\cdot|\mathbf{Z}_i)$. As the authors argued, $\widehat{G}(\cdot|\mathbf{Z}_i)$ may be obtained via ad hoc approaches, but they assumed for simplicity that C is independent of $(T, \epsilon, \mathbf{Z})$, and hence they used the classic Kaplan-Meier estimator for $G_0(\cdot|\mathbf{Z}_i)$.

To relax these restrictions, we propose a new estimator for $\boldsymbol{\beta}_0(\tau)$, which is motivated by the argument of [1] for only right-censored data. The key observation is that the individual has a probability $G_0(T|\mathbf{Z})$ of not being censored for the failure time T. Hence,

$$\begin{aligned} & E\left(\frac{\delta}{G_0(X|\mathbf{Z})} \mathbf{Z} \left[I\left\{X \leq H_{\gamma_0}^{-1}(\boldsymbol{b}^T\mathbf{Z}), \epsilon = 1\right\} - \tau \right] \middle| \mathbf{Z} \right) \\ = \ & E\Bigg\{ E\Bigg(\frac{I(T \leq C)}{G_0(T|\mathbf{Z})} \mathbf{Z} \Big[I\left\{T \leq H_{\gamma_0}^{-1}(\boldsymbol{b}^T\mathbf{Z}), \epsilon = 1\right\} \\ & \qquad - \tau \Big] \Bigg| T, \epsilon, \mathbf{Z} \Bigg) | \mathbf{Z} \Bigg\} \\ = \ & E\Bigg(\mathbf{Z} \Bigg[I\left\{T \leq H_{\gamma_0}^{-1}(\boldsymbol{b}^T\mathbf{Z}), \epsilon = 1\right\} \\ & \qquad - \tau \Bigg] E\left\{ \frac{I(T \leq C)}{G_0(T|\mathbf{Z})} | T, \epsilon, \mathbf{Z} \right\} \Bigg| \mathbf{Z} \Bigg) = 0. \end{aligned}$$

Then, our proposed estimating equation is given by

$$\begin{aligned} (4) \quad & \mathbf{U}_n(\boldsymbol{b}, \tau) \\ = \ & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\delta_i}{\widehat{G}(X_i|\mathbf{Z_i})} \mathbf{Z}_i \left[I\left\{ H_{\gamma_0}(X_i) \leq \boldsymbol{b}^T \mathbf{Z}_i, \epsilon_i = 1 \right\} - \tau \right] \\ \approx \ & 0, \end{aligned}$$

where $\widehat{G}(\cdot|\mathbf{Z}_i)$ is an estimator for $G_0(\cdot|\mathbf{Z}_i)$, which is the conditional survival function of the censoring variable C_i given the covariates. The approximation sign is used here because the function of sum is a discontinuous function of $\boldsymbol{b}$.

It is worth pointing out that the Equation (4) is monotone but not continuous. The numerical solution to Equation (4) can be obtained by locating the minimizer of the following L_1 type convex function.

$$\begin{aligned} L_n(\boldsymbol{b}, \tau) \ = \ & \sum_{i=1}^{n} I\left(\delta_i \epsilon_i = 1\right) \left| \frac{H_{\gamma_0}(X_i)}{\widehat{G}(X_i|\mathbf{Z}_i)} - \boldsymbol{b}^T \frac{\mathbf{Z}_i}{\widehat{G}(X_i|\mathbf{Z}_i)} \right| \\ & + \left| M - \boldsymbol{b}^T \sum_{l=1}^{n} \frac{-\mathbf{Z}_l I\left(\delta_l \epsilon_l = 1\right)}{\widehat{G}(X_l|\mathbf{Z}_l)} \right| \\ & + \left| M - \boldsymbol{b}^T \sum_{k=1}^{n} \frac{2\tau I\left(\delta_k \neq 0\right) \mathbf{Z}_k}{\widehat{G}(X_k|\mathbf{Z}_k)} \right|, \end{aligned}$$

where M is an extremely large positive number. It should be selected to exceed $\left|\boldsymbol{b}^T \sum_{l=1}^{n} \frac{-\mathbf{Z}_l I(\delta_l \epsilon_l = 1)}{\widehat{G}(X_l|\mathbf{Z}_l)}\right|$ and $\left|\boldsymbol{b}^T \sum_{k=1}^{n} \frac{2\tau I(\delta_k \neq 0) \mathbf{Z}_k}{\widehat{G}(X_k|\mathbf{Z}_k)}\right|$ for all $\boldsymbol{b}$ in the compact parameter space for $\boldsymbol{\beta}_0(\tau)$. The equivalency is shown in the Appendix. Since $L_n(\boldsymbol{b}, \tau)$ is a convex function on b for each τ, its minimizer can be easily found by using standard software, for example, the rq() function in the contributed **R** package *quantreg*.

Recall that [21] proposed a global Kaplan-Meire estimate for $G_0(t|z)$ by independent assumption between the censoring variable C and the covariates $\boldsymbol{Z}$. However, in practice, the distribution of C may depend on the covariates $\mathbf{Z}$. To this end, one may use the local Kaplan-Meier estimator for $G_0(X_i|\mathbf{Z}_i)$ as in [27],

$$\widehat{G}(t|\boldsymbol{z}) = \prod_{i=1}^{n} \left[1 - \frac{B_{ni}(\boldsymbol{z})}{\sum_{j=1}^{n} I(X_j \geq X_i) B_{nj}(\boldsymbol{z})} \right]^{I(X_i \leq t, \delta_i = 0)},$$

where

$$B_{ni}(\boldsymbol{z}) = \mathcal{K}\left(\frac{\boldsymbol{z} - \boldsymbol{z}_i}{h_n}\right) \left[\sum_{j=1}^{n} \mathcal{K}\left(\frac{\boldsymbol{z} - \boldsymbol{z}_j}{h_n}\right) \right]^{-1},$$

and the product kernel function $\mathcal{K}(u_1, ..., u_p) = \prod_{i=1}^{p} K(u_i)$, with $K(\cdot)$ being a univariate kernel function, and $h_n > 0$ is the bandwidth. As suggested by [19, 17], one can choose the biquadratic kernel $K(x) = \frac{15}{16}(1 - x^2)^2 I(|x| \leq 1)$ for the univariate covariate ($p = 1$), and for multiple continuous covariates with $p \geq 2$, one should use a product kernel function with a higher order kernel for each covariate. For example, when $p = 2$, one can choose $K(x) = \frac{15}{32}(3 - 10x^2 + 7x^4) I(|x| \leq 1)$ for each covariate. When $p = 3$, one can choose $K(x) = \frac{35}{256}(15 - 105x^2 + 189x^4 - 99x^6) I(|x| \leq 1)$. Because the estimate of $G_0(X_i|\mathbf{Z}_i)$ may be outside of $[0, 1]$ under these higher order kernels, one can truncate the value to $[0, 1]$ as needed.

In the second step, we propose two estimators based on cusum processes to estimate the transformation parameter γ_0, which is motivated by [18, 28]. Define a discrepancy measure based on the cusum process as

$$\begin{aligned} D_n(\mathbf{z}, \gamma_0) \ = \ & \sum_{i=1}^{n} I(\mathbf{Z}_i \leq \mathbf{z}) \Bigg(\frac{\delta_i}{\widehat{G}(X_i|\boldsymbol{Z}_i)} \\ & \left[\tau - I\left\{ H_{\gamma_0}(X_i) \leq \widehat{\boldsymbol{\beta}}(\gamma_0, \tau)^T \mathbf{Z}_i, \epsilon_i = 1 \right\} \right] \Bigg) \end{aligned}$$

and $I(\mathbf{Z}_i \leq \mathbf{z}) = I(Z_{i1} \leq z_1, Z_{i2} \leq z_2 \cdots, Z_{ip} \leq z_p)$. Note that $D_n(\cdot,\cdot)$ can be used to distinguish the correct transformation from those wrong ones. One estimator for γ_0 can be derived by minimizing the following function,

$$R_n(\gamma_0) = \sum_{i=1}^{n} D_n(\mathbf{Z}_i, \gamma_0)^2.$$

As suggested in [28], for the case with high dimensional covariates, especially many categorical variables, another alternative cusum process is

$$D_n^*(t, \gamma_0) = \sum_{i=1}^{n} I(\widehat{\boldsymbol{\beta}}(\gamma_0, \tau)^T \mathbf{Z}_i \leq t) \left(\frac{\delta_i}{\widehat{G}(X_i|\boldsymbol{Z}_i)} \left[\tau - I\left\{ H_\gamma(X_i) \leq \widehat{\boldsymbol{\beta}}(\gamma_0, \tau)^T \mathbf{Z}_i, \epsilon_i = 1 \right\} \right] \right).$$

Then, another estimator for γ_0 is obtained by minimizing

$$R_n^*(\gamma_0) = \sum_{i=1}^{n} \int_0^\infty D_n^*(t, \gamma_0)^2 dN_i(t),$$

where $N_i(t) = I(H_\gamma(X_i) \leq t)$.

3. LARGE-SAMPLE PROPERTIES OF THE PROPOSED ESTIMATOR

In this section, we establish the consistency and the asymptotic normality of the estimators. To derive the asymptotic properties of the proposed estimator, we require the following technical regularity assumptions.

C1. γ_0 belongs to a compact set Γ.
C2. The support $\mathcal{Z}$ of $\mathbf{Z}$ is uniformly bounded and if $H_\gamma^{-1}(\boldsymbol{\beta}^T \mathbf{Z}) = H_{\gamma_0}^{-1}(\boldsymbol{\beta}_0^T \mathbf{Z})$, then $\boldsymbol{\beta} = \boldsymbol{\beta}_0$, and $\gamma = \gamma_0$. Moreover, the transformation $H_\gamma(.)$ is strictly increasing and twice-continuously differentiable in a neighborhood of γ_0.
C3. There exists constants $\upsilon > 0$ and $\lambda > 0$ such that $\inf_{\boldsymbol{Z} \in \mathcal{Z}} \Pr(\mathrm{C} \geq \lambda | \boldsymbol{Z}) > \upsilon$.
C4. The conditional density functions $f_1(t|\boldsymbol{z})$, and $g_0(t|\boldsymbol{z})$ are uniformly bounded away from infinity and have bounded (uniformly in t) first order partial derivatives with respect to $\boldsymbol{z}$, where $f_1(t|\boldsymbol{z}) = dPr(H_{\gamma_0}(T) \leq t, \epsilon = 1|\boldsymbol{z})/dt$ and $g_0(t|\boldsymbol{z}) = dG_0(t|\boldsymbol{z})/dt$. In addition, $\boldsymbol{\beta}_0(\tau)$ is Lipschitz continuous for $\tau \in [\tau_L, \tau_U]$.
C5. The 100τth of $H_{\gamma_0}(T_1^*)$ given $\mathbf{Z}$ is unique with probability 1 and is strictly less than λ.
C6. The bandwidth h_n satisfies $h_n = O(n^{-\nu})$ with $1/4 < \nu < 1/3$.
C7. The kernel function $K(\cdot) \geq 0$ has a compact support and satisfies the Lipschitz condition of order 1, $\int K(u)du = 1, \int uK(u)du = 0, \int K^2(u)du < \infty$, and $\int |u|^2 K(u)du < \infty$.

Conditions C1 and C3 are standard in the context of survival analysis. Condition C2 ensures the identifiability of the transformation and regression parameters, and the unique parameterization of the transformation. Condition C4 entails the identifiability of $\beta_0(\tau)$ and furthermore, the consistency of $\widehat{\beta}(\widehat{\gamma}, \tau)$. Condition C5 is needed for the estimable problem for the 100τth quantile from the data. Condition C6 is needed to ensure the consistency of the local Kaplan-Meier estimator and Condition C7 is routinely made in nonparametric smoothing.

For brevity, let $\|\cdot\|$ and $\|\cdot\|_\infty$ be Euclidean norm and the supreme norm metric, respectively. Define

$$\mathbf{U}_n^G(\mathbf{b}, \tau, \gamma) = n^{-1/2} \sum_{i=1}^{n} \frac{\delta_i}{G_0(X_i|\mathbf{Z_i})} \mathbf{Z}_i \Big\{ I(H_\gamma(X_i) \leq \mathbf{b}^T \mathbf{Z}_i, \epsilon_i = 1) - \tau \Big\},$$

$$\widetilde{U}_n(\mathbf{b}, \tau, \gamma) = n^{-1/2} \sum_{i=1}^{n} \mathbf{Z}_i \Big\{ \Pr\big(H_\gamma(T_i) \leq \mathbf{b}^T \mathbf{Z}_i, \epsilon_i = 1 | \mathbf{Z}_i\big) - \tau \Big\},$$

and $\Psi(\mathbf{b}, \tau, \gamma) = E\left\{ n^{-1/2} \widetilde{U}_n(\mathbf{b}, \tau, \gamma) \right\}$. We denote $\Psi_{0\gamma}, \Psi_{0\mathbf{b}}$ as the row-vector gradients of Ψ with respect to $(\gamma, \mathbf{b})$ evaluated at the true parameters $(\gamma_0, \boldsymbol{\beta}_0(\tau))$. Condition C2, with assuming $\gamma = \gamma_0$, implies that $\mathbf{Z}$ is linearly independent. Combining with condition C4, we can show that $E(\mathbf{Z}\Psi_{0\mathbf{b}})$ is positive definite. That is, for some $\rho_0 > 0$ and $c_0 > 0$, $\inf_{\mathbf{b} \in \mathcal{B}(\rho_0)} \mathrm{eigminE}(\mathbf{Z}\Psi_{0\mathbf{b}}) \geq c_0$, where $\mathcal{B}(\rho_0) = \{\mathbf{b} : \inf_{\tau \in [\tau_L, \tau_U]} \|\mathbf{b} - \boldsymbol{\beta}_0(\tau)\| \leq \rho_0\}$ and $\mathrm{eigmin}(u)$ is the minimum eigenvalue of u.

Theorem 3.1. *Under conditions C1–C7, with probability 1, it holds*

$$|\widehat{\gamma} - \gamma_0| + \sup_{\tau \in [\tau_L, \tau_U]} \|\widehat{\boldsymbol{\beta}}(\widehat{\gamma}, \tau) - \boldsymbol{\beta}_0(\tau)\| \to 0.$$

Theorem 3.2. *Under conditions C1–C7, we have*

$$\sqrt{n} \begin{pmatrix} \widehat{\gamma} - \gamma_0 \\ \widehat{\boldsymbol{\beta}}(\widehat{\gamma}, \tau) - \boldsymbol{\beta}_\mathbf{0}(\tau) \end{pmatrix} \xrightarrow{\mathcal{D}} N(0, \boldsymbol{\Sigma})$$

for $\tau \in [\tau_L, \tau_U]$, *where* $\boldsymbol{\Sigma}$ *is given at the end of the Appendix.*

It is cumbersome to compute the covariance estimator of the parameter estimates directly by plug-in estimators, because the structure of covariance is quite complicated. To lighten the computational burden, we adopt a nonparametric bootstrap method [6].

Another important issue is the choice of bandwidth. We adopt the L-fold cross-validation method for selecting h_n. Specifically, one can first divide the data set into L parts, which are of almost the same size. For the mth part, we fit the model by using the rest $L-1$ parts of the data, and then evaluate the quantile loss from predicting the τth conditional quantile of T on the uncensored data that are left out. Repeating the above procedure for $m=1,\cdots,L$, the optimal bandwidth h_n is obtained by minimizing the average quantile loss. Fortunately, in our empirical studies, the proposed estimator is not sensitive to the choice of bandwidth.

4. SIMULATION STUDIES

We conduct simulation studies to assess the finite-sample performance of the proposed methods. We consider $p=2$, and the covariate vector $\widetilde{\boldsymbol{Z}} = (Z^{(1)}, Z^{(2)})^T$ is generated as $Z^{(1)} \sim$ Uniform(0,1), $Z^{(2)} \sim$ Bernoulli(0.5). The distribution of (T,ϵ) satisfies $\Pr(\epsilon=1|\mathbf{Z}) = p_0 I\left(Z^{(2)}=0\right) + p_1 I\left(Z^{(2)}=1\right)$,

$$\Pr(T \le t|\epsilon=1,\mathbf{Z}) = \Phi\left(\frac{H_{\gamma_0}(t)-\boldsymbol{\theta}_0^T\mathbf{Z}}{\text{sd}}\right),$$

$$\Pr(T \le t|\epsilon=2,\mathbf{Z}) = \Phi\left(\frac{H_{\gamma_0}(t)-\boldsymbol{\alpha}_0^T\mathbf{Z}}{\text{sd}}\right),$$

where $\Phi(\cdot)$ is the standard normal distribution function, $(p_0,p_1)=(0.9,0.7)$, $\boldsymbol{\alpha}_0=(0,0.2)^T$, $\boldsymbol{\theta}_0=(1,0.5)^T$ and sd $=0.5$. Hence, the underlying quantile regression model takes the form

$$H_{\gamma_0}\{Q_1(\tau|\mathbf{Z})\} = \Phi^{-1}(\tau/p_0)*\text{sd} + \theta_0^{(1)}Z^{(1)} + \left\{\theta_0^{(2)} + \Phi^{-1}(\tau/p_1)*\text{sd} - \Phi^{-1}(\tau/p_0)*\text{sd}\right\}Z^{(2)}.$$

We consider two types of censorships: covariate-independent censoring and covariate-dependent censoring.

(1) Covariate-independent censoring:
The censoring variable C is generated independently form a mixture of Uniform $(0,d_u)$ and a point mass at d_u, that is,

$$\Pr(C \le x) = 0.8(x/d_u)I(0 \le x < d_u) + I(x \ge d_u).$$

(2) Covariate-dependent censoring:
C is generated from Uniform$(0, c_u Z^{(1)})$.

In each scenario, we consider two transformation parameters $\gamma_0=0$ and $\gamma_0=0.5$, and two different censoring rates $C\%=10\%$ and $C\%=30\%$ by controlling the constants d_u or c_u. For each configuration, we replicate 500 data sets with sample size $n=200$. We set $\tau=0.1,0.2,0.3,0.4$, and $M=10^5$, which is proposed in [21]. For simplicity, we use $\widehat{\beta}^{(\cdot)}$ to represent the estimated $\widehat{\beta}^{(\cdot)}(\widehat{\gamma},\tau)$ in the below tables. To obtain the standard errors (SE) of the parameter estimates, we use the nonparametric bootstrap method with 500 resampled data sets.

We report the empirical bias (Bias), the empirical standard deviation (SD), the average of estimated standard errors (SE) and the 95% coverage probability (CP) in each table. Tables 1 and 2 show the results for covariate-independent censoring for $\gamma_0=0$ and $\gamma_0=0.5$, respectively. From Tables 1–2, all the estimators are unbiased and have small mean square errors, the estimated standard errors are close to the empirical standard deviations, and the coverage probabilities are close to the nominal levels.

For the covariate-dependent censoring case, the results are summarized in Tables 3–4. The results are similar to those for covariate-indpendent censoring case. In summary, the results from our simulation indicate that the propsoed estimators have good performance in finite samples. Here, motivated by the asymptotic result, the bandwidth is set as $h_n = n^{-1/3+0.001}$. Additional simulations show that the results are not very sensitive to the bandwidth and the kernel function.

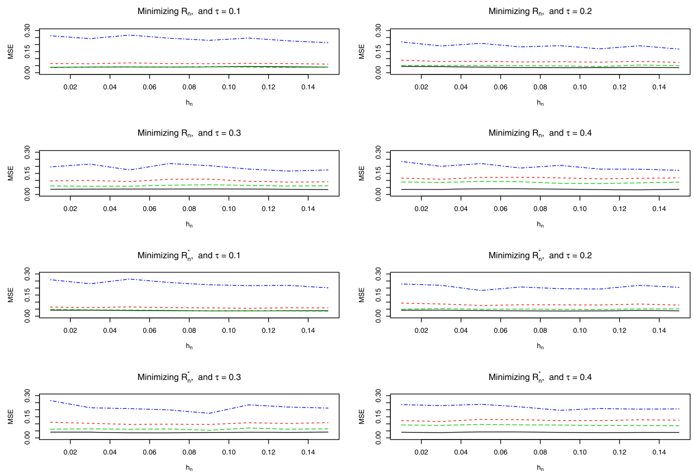

Figure 1. Plots of mean square errors versus h_n for $n=200$ observations with $\gamma=0.5$ and censoring rate of 30%. The MSEs of each regression coefficient are presented: $\beta^{(1)}$ (black solid '—'), $\beta^{(2)}$ (red dashed '-'), $\beta^{(3)}$ (green longdash '–') and γ (blue twodash '-.-').

To investigate the sensitivity of the proposed method to h_n, we investigate the mean square errors of the proposed estimators with $h_n \in [0.01, 0.15]$. Figure 1 plots the MSE of the proposed estimates versus h_n in covariate-dependent censoring case with $\gamma_0=0.5$ and $n=200$ with the censoring rate of about 30%. One can observe that the performance of the proposed method is not sensitive to the choice of h_n.

We also conduct the comparison between the proposed estimator and the Peng and Fine's estimator by [21] in both the covariate-independent censoring and covariate-dependent censoring. The results are tabulated in Table 5 for γ_0 fixed under censoring rate 30%. From Table 5, the proposed estimator is comparable with the Peng and Fine's estimator when covariate-independent censoring, whereas the proposed estimator has smaller mean square errors (RMSE) than the Peng and Fine's estimator when covariate-dependent censoring.

Table 1. Simulation results for covariate-independent censoring with unknown transformation parameter $\gamma_0 = 0$

			minimizing $R_n(\gamma)$				minimizing $R_n^*(\gamma)$			
τ	$C\%$		$\widehat{\beta}^{(1)}(\tau)$	$\widehat{\beta}^{(2)}(\tau)$	$\widehat{\beta}^{(3)}(\tau)$	$\widehat{\gamma}$	$\widehat{\beta}^{(1)}(\tau)$	$\widehat{\beta}^{(2)}(\tau)$	$\widehat{\beta}^{(3)}(\tau)$	$\widehat{\gamma}$
0.1	10	Bias	0.014	-0.001	-0.005	-0.033	-0.012	0.041	0.018	0.036
		SD	0.206	0.270	0.191	0.367	0.208	0.274	0.200	0.372
		SE	0.230	0.310	0.221	0.419	0.235	0.314	0.229	0.421
		CP	0.960	0.970	0.974	0.968	0.956	0.964	0.978	0.970
	30	Bias	0.041	-0.013	-0.007	0.017	0.016	0.027	0.021	0.099
		SD	0.215	0.295	0.206	0.422	0.213	0.291	0.214	0.428
		SE	0.233	0.323	0.230	0.471	0.234	0.325	0.238	0.474
		CP	0.948	0.968	0.960	0.980	0.950	0.970	0.970	0.964
0.2	10	Bias	-0.003	0.015	0.011	-0.018	-0.016	0.046	0.026	0.023
		SD	0.193	0.290	0.197	0.307	0.199	0.300	0.197	0.311
		SE	0.223	0.344	0.238	0.364	0.223	0.344	0.242	0.365
		CP	0.964	0.978	0.976	0.970	0.952	0.970	0.978	0.968
	30	Bias	-0.000	0.050	0.023	0.074	-0.010	0.070	0.035	0.101
		SD	0.210	0.332	0.226	0.367	0.206	0.322	0.231	0.373
		SE	0.229	0.372	0.260	0.433	0.227	0.374	0.264	0.435
		CP	0.954	0.972	0.978	0.970	0.950	0.972	0.972	0.960
0.3	10	Bias	-0.008	0.022	0.016	-0.019	-0.004	0.018	0.014	-0.015
		SD	0.192	0.310	0.220	0.287	0.186	0.296	0.212	0.277
		SE	0.224	0.388	0.280	0.350	0.221	0.385	0.282	0.353
		CP	0.964	0.970	0.974	0.972	0.964	0.984	0.984	0.984
	30	Bias	0.004	0.043	0.023	0.039	-0.013	0.090	0.047	0.098
		SD	0.219	0.390	0.273	0.384	0.215	0.386	0.270	0.370
		SE	0.235	0.434	0.319	0.436	0.228	0.429	0.317	0.435
		CP	0.952	0.962	0.960	0.958	0.944	0.956	0.956	0.958
0.4	10	Bias	-0.010	0.032	0.031	-0.015	-0.014	0.041	0.037	-0.008
		SD	0.193	0.347	0.272	0.289	0.193	0.349	0.266	0.290
		SE	0.222	0.426	0.348	0.351	0.219	0.420	0.350	0.354
		CP	0.952	0.972	0.970	0.976	0.948	0.974	0.964	0.970
	30	Bias	-0.027	0.122	0.084	0.086	-0.021	0.120	0.089	0.098
		SD	0.234	0.480	0.339	0.416	0.222	0.465	0.348	0.406
		SE	0.238	0.493	0.416	0.452	0.228	0.482	0.416	0.451
		CP	0.942	0.928	0.954	0.934	0.958	0.940	0.964	0.950

† Bias, empirical bias; SD, the empirical standard deviation; SE, the average of estimated standard errors; CP, the 95% coverage probability.

As suggested by the Editor, to assess the effect of dimensionality, we add $p-2$ additional independent covariates $Z^{(3)}, \cdots, Z^{(p)}$ and set $n = 200$. Each of the added variable is generated from Uniform(0,2). The censoring time C is generated as $C = \varsigma_0 + \varsigma_1 Z^{(1)} + \cdots + \varsigma_p Z^{(p)}$, where $\varsigma_i \geq 0$, $i = 0, \cdots, p$, and the $\varsigma_i s$ are taken to have about 10% or 30% censoring rate. For simplicity, we took the bandwidth $h_n = 0.5, 0.7$ for p = 3 and 4, respectively. The transformation parameter is $\gamma = 0$. The results are summarized in Table 6. The results for $\gamma = 0.5$ are similar, we omit them to save space. Overall, the proposed method performs moderately well.

5. A REAL EXAMPLE

We analyze the data set involving patients with follicular type lymphoma [24, 23], conducted by the Princess Margaret Hospital between 1967 and 1996, in which the failure responses to the treatment were mainly presented by two competing end points: first relapse (local, distant or both) and nonrecurrence-related death. The goal of the study was to investigate the survival time of patients after follicular type lymphoma. In this trial, total of 541 patients were diagnosed with follicular cell lymphoma, out of which 272 individuals experienced lymphoma relapse which was the interested event, 76 individuals were observed with nonrecurrence-related death and the rest were lost to follow-up. The covariates were age, haemoglobin (HGB), clinical stage (CS, 0 for stage I; 1 for stage II), and treatment (0 for radiation and chemotherapy (RCMT); 1 for radiation alone (RT)). We first plot the Kaplan-Meier survival curves of the censoring time separately for patients groups stratified by clinical stage, treatment, the median age (58 years), or median haemoglobin (140 g/l).From Figure 2, we observe that there seems to be survival difference of the censoring time across each covariate, which indicates the censoring

Table 2. Simulation results for covariate-independent censoring with with unknown transformation parameter $\gamma_0 = 0.5$

			minimizing $R_n(\gamma)$				minimizing $R_n^*(\gamma)$			
τ	$C\%$		$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$	$\widehat{\gamma}$	$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$	$\widehat{\gamma}$
0.1	10	Bias	0.038	-0.034	-0.011	-0.089	0.017	-0.000	0.009	-0.019
		SD	0.194	0.235	0.182	0.412	0.198	0.237	0.190	0.414
		SE	0.198	0.242	0.181	0.445	0.199	0.244	0.187	0.451
		CP	0.924	0.960	0.922	0.984	0.922	0.962	0.926	0.976
	30	Bias	0.029	-0.033	-0.017	-0.116	0.021	-0.013	-0.006	-0.065
		SD	0.200	0.239	0.187	0.424	0.194	0.232	0.189	0.420
		SE	0.213	0.261	0.199	0.480	0.214	0.264	0.205	0.487
		CP	0.946	0.950	0.958	0.972	0.948	0.966	0.972	0.972
0.2	10	Bias	0.011	-0.007	0.009	-0.038	0.008	0.001	0.016	-0.026
		SD	0.179	0.250	0.189	0.356	0.183	0.257	0.202	0.377
		SE	0.189	0.267	0.192	0.412	0.191	0.272	0.197	0.423
		CP	0.938	0.940	0.930	0.958	0.934	0.950	0.940	0.956
	30	Bias	0.025	-0.028	-0.013	-0.100	0.019	-0.008	-0.004	-0.062
		SD	0.187	0.273	0.198	0.414	0.189	0.277	0.202	0.423
		SE	0.205	0.289	0.209	0.457	0.205	0.292	0.215	0.464
		CP	0.956	0.952	0.942	0.970	0.952	0.960	0.954	0.980
0.3	10	Bias	0.012	-0.011	0.005	-0.056	0.007	-0.000	0.016	-0.042
		SD	0.173	0.285	0.215	0.375	0.181	0.296	0.231	0.394
		SE	0.186	0.298	0.217	0.416	0.189	0.305	0.225	0.434
		CP	0.942	0.942	0.930	0.958	0.948	0.956	0.930	0.976
	30	Bias	0.029	-0.032	-0.023	-0.102	0.024	-0.018	-0.015	-0.088
		SD	0.186	0.308	0.227	0.414	0.191	0.322	0.237	0.445
		SE	0.199	0.318	0.234	0.465	0.200	0.323	0.240	0.476
		CP	0.940	0.948	0.940	0.978	0.926	0.940	0.942	0.966
0.4	10	Bias	0.005	-0.011	0.011	-0.074	0.006	-0.007	0.017	-0.070
		SD	0.182	0.341	0.273	0.417	0.182	0.345	0.283	0.435
		SE	0.183	0.328	0.252	0.433	0.184	0.334	0.260	0.451
		CP	0.948	0.938	0.926	0.956	0.924	0.928	0.916	0.942
	30	Bias	0.032	-0.045	-0.028	-0.122	0.030	-0.034	-0.020	-0.113
		SD	0.183	0.347	0.268	0.442	0.191	0.366	0.278	0.469
		SE	0.191	0.345	0.274	0.488	0.191	0.351	0.282	0.498
		CP	0.940	0.944	0.926	0.958	0.924	0.930	0.918	0.946

† Bias, empirical bias; SD, the empirical standard deviation; SE, the average of estimated standard errors; CP, the 95% coverage probability.

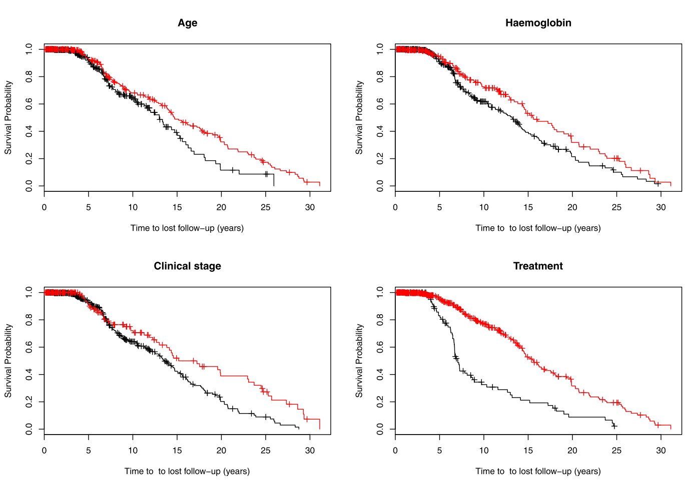

Figure 2. The estimated Kaplan-Meier survival curves for the follicular type lymphoma study.

time is significantly dependent on the covariates. For choosing the bandwidth h_n, we adopt the 10-fold cross-validation method.

Applying the proposed power-transformed quantile regression model to the lymphoma study, we estimate γ by minimizing $R_n^*(\gamma)$ and take 500 bootstrap samples for the variance estimation. Figure 3 presents that all the relapse cumulative incidence probabilities exceed 0.45 in the right tail, but not attain 0.2 for the death. Hence, we postulate the value of τ from 0.1 up to 0.45 in steps of 0.05 for the relapse, and from 0.05 to 0.1 in steps of 0.01 for the competing risk. Based on the numerical experience, we find $M = 10^5$ is large enough. We also use it in this real dataset. Due to the transformation on the failure time, there are distinct scales of the covariate effects obtained from each quantile regression. It is more useful to derive the marginal covariate effects in the original scale of the outcome [14, 18, 28]. For a given covariate $\boldsymbol{z}_0$, we assess the marginal effects in the following way. For the jth covariate, we consider two different cases:

Table 3. Simulation results for covariate-dependent censoring with with unknown transformation parameter $\gamma_0 = 0$

			minimizing $R_n(\gamma)$				minimizing $R_n^*(\gamma)$			
τ	$C\%$		$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$	$\widehat{\gamma}$	$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$	$\widehat{\gamma}$
0.1	10	Bias	0.012	0.002	-0.004	-0.023	-0.011	0.036	0.018	0.039
		SD	0.210	0.276	0.199	0.383	0.202	0.270	0.196	0.373
		SE	0.229	0.310	0.221	0.422	0.234	0.314	0.229	0.425
		CP	0.958	0.966	0.972	0.972	0.956	0.976	0.966	0.970
	30	Bias	0.036	-0.024	-0.003	0.019	0.009	0.024	0.021	0.109
		SD	0.216	0.290	0.206	0.435	0.209	0.289	0.209	0.443
		SE	0.232	0.320	0.229	0.479	0.232	0.322	0.236	0.484
		CP	0.944	0.964	0.966	0.982	0.956	0.974	0.966	0.970
0.2	10	Bias	-0.009	0.021	0.019	-0.004	-0.014	0.037	0.028	0.022
		SD	0.192	0.291	0.200	0.320	0.196	0.293	0.200	0.316
		SE	0.223	0.346	0.240	0.369	0.222	0.345	0.242	0.369
		CP	0.964	0.966	0.976	0.964	0.962	0.974	0.980	0.974
	30	Bias	0.001	0.033	0.013	0.062	-0.000	0.043	0.022	0.085
		SD	0.213	0.336	0.220	0.382	0.201	0.320	0.224	0.385
		SE	0.227	0.367	0.257	0.442	0.225	0.368	0.260	0.445
		CP	0.958	0.970	0.978	0.974	0.958	0.976	0.984	0.970
0.3	10	Bias	-0.006	0.016	0.012	-0.016	-0.005	0.019	0.016	-0.011
		SD	0.194	0.309	0.215	0.286	0.194	0.310	0.218	0.293
		SE	0.224	0.389	0.281	0.354	0.221	0.387	0.283	0.357
		CP	0.968	0.974	0.978	0.976	0.960	0.976	0.984	0.974
	30	Bias	-0.007	0.040	0.022	0.047	-0.018	0.074	0.040	0.091
		SD	0.218	0.379	0.254	0.375	0.208	0.375	0.253	0.376
		SE	0.232	0.425	0.311	0.444	0.226	0.420	0.309	0.443
		CP	0.946	0.964	0.972	0.964	0.954	0.964	0.966	0.958
0.4	10	Bias	-0.015	0.038	0.035	-0.008	-0.017	0.049	0.042	0.002
		SD	0.199	0.359	0.270	0.300	0.196	0.359	0.273	0.300
		SE	0.223	0.428	0.348	0.355	0.219	0.421	0.351	0.359
		CP	0.960	0.972	0.968	0.978	0.964	0.958	0.966	0.966
	30	Bias	-0.012	0.062	0.051	0.048	-0.007	0.062	0.057	0.061
		SD	0.222	0.458	0.327	0.423	0.220	0.435	0.338	0.409
		SE	0.233	0.475	0.403	0.456	0.225	0.467	0.402	0.457
		CP	0.940	0.922	0.950	0.932	0.948	0.946	0.964	0.938

† Bias, empirical bias; SD, the empirical standard deviation; SE, the average of estimated standard errors; CP, the 95% coverage probability.

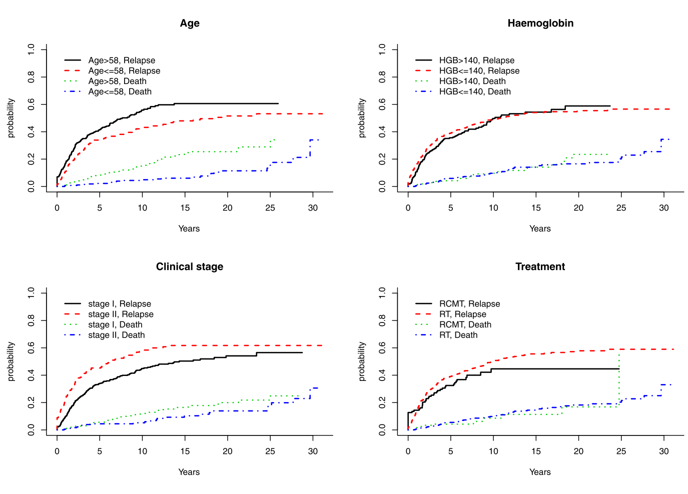

Figure 3. Cause-specific cumulative incidence estimates of relapse and nonrecurrence-related death for the follicular type lymphoma data.

(i) When $\boldsymbol{z}$ is discrete,

$$Q_1(\tau|\boldsymbol{z}_{0(-j)}, Z_j = 1) - Q_1(\tau|\boldsymbol{z}_{0(-j)}, Z_j = 0),$$

where $\boldsymbol{z}_{0(-j)}$ is all the components of z_0 with deleting the jth component.

(ii) When $\boldsymbol{z}$ is continuous, then

$$\frac{\partial Q_1(\tau|\boldsymbol{Z})}{\partial Z_j}|_{\boldsymbol{z}_0} = \begin{cases} \beta_{\tau,j}(\gamma_\tau \boldsymbol{\beta}_\tau^T \boldsymbol{z}_0 + 1)^{1/\gamma_\tau - 1}, & \gamma_\tau \neq 0 \\ \beta_{\tau,j}\exp(\boldsymbol{\beta}_\tau^T \boldsymbol{z}_0), & \gamma_\tau = 0 \end{cases}.$$

Here we take $\boldsymbol{z}_0$ as particular covariates from a 58-year-old patient with HGB = 140 g/l, CS = 1, treatment = 1. The estimates (in bold solid lines) and the pointwise 95% confidence bands for the marginal covariate effects and the transformation parameter γ are displayed in Figure 4. We can see that the patient age and haemoglobin do not appear to significantly affect the recurrence. But the patients in clinical stage II seem to experience recurrence sooner than those in stage I. The estimate of the Box-Cox transformation

Table 4. Simulation results for covariate-dependent censoring with with unknown transformation parameter $\gamma_0 = 0.5$

			minimizing $R_n(\gamma)$				minimizing $R_n^*(\gamma)$			
τ	$C\%$		$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$	$\widehat{\gamma}$	$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$	$\widehat{\gamma}$
0.1	10	Bias	0.036	-0.039	-0.021	-0.108	0.026	-0.019	-0.009	-0.066
		SD	0.192	0.223	0.180	0.384	0.200	0.237	0.187	0.415
		SE	0.198	0.242	0.181	0.447	0.199	0.244	0.187	0.452
		CP	0.926	0.958	0.920	0.982	0.918	0.952	0.928	0.972
	30	Bias	0.026	-0.052	-0.017	-0.148	0.015	-0.028	-0.010	-0.091
		SD	0.209	0.251	0.192	0.451	0.205	0.244	0.196	0.448
		SE	0.213	0.260	0.200	0.492	0.213	0.262	0.206	0.501
		CP	0.954	0.950	0.950	0.958	0.948	0.944	0.958	0.964
0.2	10	Bias	0.021	-0.022	-0.013	-0.074	0.022	-0.019	-0.012	-0.068
		SD	0.184	0.260	0.182	0.375	0.191	0.274	0.188	0.404
		SE	0.189	0.267	0.191	0.413	0.191	0.271	0.197	0.423
		CP	0.948	0.950	0.938	0.964	0.946	0.942	0.934	0.960
	30	Bias	0.018	-0.038	-0.022	-0.110	0.012	-0.025	-0.011	-0.084
		SD	0.186	0.272	0.187	0.416	0.194	0.278	0.199	0.436
		SE	0.205	0.288	0.210	0.470	0.204	0.289	0.214	0.476
		CP	0.950	0.958	0.952	0.972	0.942	0.946	0.940	0.972
0.3	10	Bias	0.011	-0.016	-0.010	-0.072	0.018	-0.024	-0.012	-0.080
		SD	0.181	0.295	0.215	0.386	0.183	0.304	0.226	0.402
		SE	0.186	0.298	0.217	0.417	0.188	0.304	0.224	0.433
		CP	0.940	0.938	0.946	0.968	0.938	0.948	0.924	0.970
	30	Bias	0.019	-0.045	-0.027	-0.115	0.022	-0.044	-0.030	-0.118
		SD	0.181	0.300	0.225	0.421	0.185	0.314	0.229	0.452
		SE	0.198	0.315	0.236	0.477	0.198	0.320	0.243	0.489
		CP	0.954	0.948	0.940	0.976	0.944	0.936	0.938	0.964
0.4	10	Bias	0.012	-0.028	-0.010	-0.090	0.009	-0.013	-0.003	-0.083
		SD	0.173	0.316	0.247	0.397	0.183	0.345	0.263	0.444
		SE	0.183	0.328	0.251	0.434	0.184	0.335	0.259	0.452
		CP	0.940	0.950	0.928	0.956	0.924	0.920	0.918	0.932
	30	Bias	0.019	-0.059	-0.035	-0.140	0.020	-0.060	-0.031	-0.155
		SD	0.181	0.340	0.256	0.445	0.190	0.364	0.278	0.499
		SE	0.191	0.344	0.292	0.501	0.192	0.350	0.300	0.513
		CP	0.944	0.932	0.932	0.954	0.928	0.904	0.912	0.924

† Bias, empirical bias; SD, the empirical standard deviation; SE, the average of estimated standard errors; CP, the 95% coverage probability.

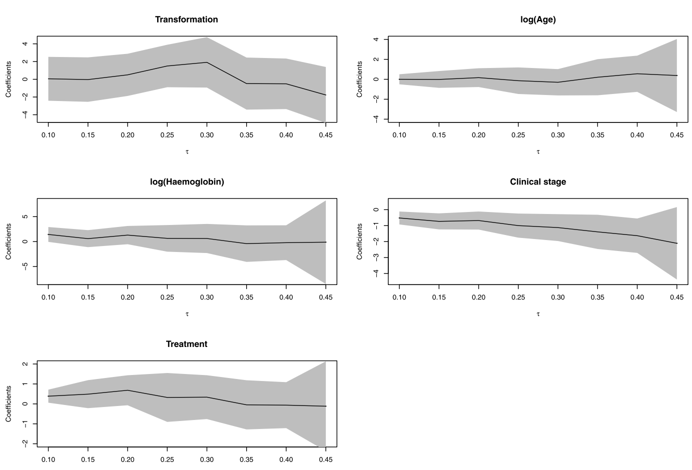

Figure 4. Marginal covariate effects for a 58-year-old patient with hgb = 140 g/l, clinstg = II, treatment = RT based on the quantile regression analysis of the follicular type lymphoma relapse study with a parametric Box-Cox transformation.

parameter γ fluctuates around the horizonal zero axis but tends to be negative for the latter quantiles ($\tau > 0.3$).

Furthermore, Figures 5–6 describe the estimation results for relapse and nonrecurrence-related death, respectively. The results suggest that the patients in clinical stage II tend to experience relapse earlier. The treatment significantly affects the relapse time for the first few early quantiles ($\tau < 0.25$). Evaluating the covariates of age and HGB, we find no significant effects on relapse for all of the regression quantiles. Figure 6 shows that all of the covariates are not significant except the age. That is, older patients tend to experience nonrecurrence-related death sooner.

6. CONCLUDING REMARKS

In this paper, we develop a power-transformed linear quantile regression model for competing risks data under both the covariate-independent censoring and covariate-dependent censoring. Since the estimating functions are not smooth with respect to regression parameters, we solve the

Table 5. *Comparison of the proposed estimator and Peng and Fine's estimator when γ is fixed under censoring rate of 30% for $\tau = 0.1, 0.2, 0.3, 0.4$*

		Proposed						Peng and Fine's estimator					
		$\gamma = 0$			$\gamma = 0.5$			$\gamma = 0$			$\gamma = 0.5$		
τ		$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$	$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$	$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$	$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$
Covariate-independent censoring													
0.1	Bias	0.017	-0.037	-0.013	-0.006	-0.003	0.003	-0.008	0.015	0.007	-0.003	0.009	0.008
	SD	0.154	0.130	0.155	0.159	0.132	0.161	0.155	0.130	0.148	0.154	0.129	0.154
	RMSE	0.155	0.135	0.156	0.159	0.132	0.161	0.155	0.130	0.148	0.154	0.129	0.154
0.2	Bias	0.030	-0.058	-0.024	-0.004	-0.009	0.001	-0.007	0.011	0.011	-0.003	0.008	0.006
	SD	0.133	0.115	0.132	0.137	0.118	0.131	0.136	0.114	0.132	0.137	0.115	0.133
	RMSE	0.136	0.129	0.134	0.137	0.118	0.131	0.136	0.115	0.132	0.137	0.115	0.133
0.3	Bias	0.034	-0.069	-0.028	-0.001	-0.012	-0.004	-0.020	0.025	0.022	-0.001	0.008	0.005
	SD	0.128	0.110	0.129	0.137	0.117	0.133	0.150	0.132	0.152	0.137	0.119	0.139
	RMSE	0.132	0.130	0.132	0.137	0.117	0.132	0.151	0.134	0.154	0.137	0.119	0.139
0.4	Bias	0.034	-0.072	-0.027	-0.004	-0.013	0.005	-0.050	0.056	0.063	0.001	0.005	0.012
	SD	0.133	0.120	0.138	0.136	0.120	0.141	0.186	0.176	0.217	0.148	0.131	0.152
	RMSE	0.137	0.140	0.141	0.136	0.121	0.141	0.192	0.184	0.226	0.148	0.131	0.152
Covariate-dependent censoring													
0.1	Bias	-0.053	0.033	-0.008	-0.071	0.052	0.006	0.085	-0.092	0.014	0.110	-0.115	0.019
	SD	0.177	0.140	0.148	0.179	0.141	0.152	0.172	0.141	0.150	0.183	0.148	0.154
	RMSE	0.185	0.144	0.148	0.192	0.150	0.152	0.192	0.168	0.151	0.213	0.187	0.155
0.2	Bias	-0.062	0.037	-0.011	-0.076	0.053	0.008	0.149	-0.155	0.020	0.192	-0.192	0.027
	SD	0.159	0.127	0.134	0.162	0.128	0.138	0.170	0.139	0.136	0.184	0.150	0.143
	RMSE	0.171	0.132	0.135	0.179	0.138	0.138	0.226	0.208	0.137	0.266	0.243	0.145
0.3	Bias	-0.059	0.030	-0.010	-0.076	0.052	0.007	0.226	-0.226	0.035	0.294	-0.283	0.046
	SD	0.154	0.124	0.133	0.169	0.133	0.140	0.184	0.152	0.143	0.208	0.170	0.151
	RMSE	0.165	0.128	0.133	0.185	0.142	0.140	0.292	0.273	0.147	0.360	0.330	0.158
0.4	Bias	-0.063	0.030	-0.010	-0.080	0.051	0.016	0.357	-0.340	0.070	0.471	-0.430	0.081
	SD	0.164	0.137	0.144	0.181	0.145	0.146	0.242	0.202	0.175	0.308	0.249	0.188
	RMSE	0.176	0.140	0.144	0.197	0.154	0.146	0.431	0.395	0.188	0.563	0.497	0.204

† Bias, empirical bias; SD, the empirical standard deviation; RMSE, the root mean square error.

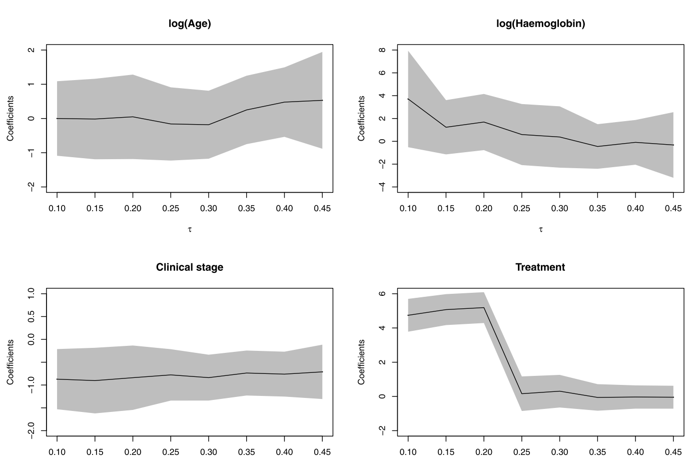

Figure 5. Quantile regression covariate effects for the relapse when fixing the Box-Cox transformation parameter $\gamma = 0$.

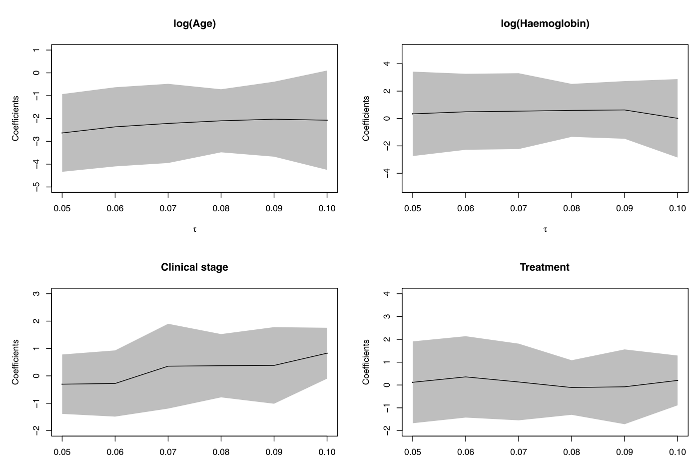

Figure 6. Quantile regression covariate effects for the competing risks when fixing the Box-Cox transformation parameter $\gamma = 0$.

Table 6. The effect of dimensionality when transformation paramater is unfixed, here p is the number of non-intercept covariates in the model ($\gamma_0 = 0$)

			minimizing $R_n(\gamma)$				minimizing $R_n^*(\gamma)$			
$C\%$	τ		$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$	$\widehat{\gamma}$	$\widehat{\beta}^{(1)}$	$\widehat{\beta}^{(2)}$	$\widehat{\beta}^{(3)}$	$\widehat{\gamma}$
$p = 3$										
10	0.1	Bias	0.025	-0.009	0.006	0.053	0.006	0.023	0.020	0.096
		SD	0.201	0.267	0.202	0.387	0.212	0.284	0.214	0.406
	0.2	Bias	0.004	0.019	0.030	0.065	-0.002	0.029	0.038	0.078
		SD	0.196	0.313	0.216	0.361	0.202	0.325	0.220	0.362
	0.3	Bias	-0.004	0.033	0.049	0.074	0.001	0.026	0.045	0.075
		SD	0.194	0.338	0.246	0.327	0.191	0.323	0.245	0.320
	0.4	Bias	-0.010	0.054	0.082	0.075	-0.001	0.037	0.069	0.063
		SD	0.200	0.410	0.312	0.355	0.200	0.398	0.310	0.358
30	0.1	Bias	0.042	-0.095	-0.007	0.016	0.035	-0.076	0.004	0.079
		SD	0.191	0.255	0.197	0.478	0.195	0.265	0.200	0.502
	0.2	Bias	0.016	-0.085	0.019	0.031	0.013	-0.070	0.026	0.053
		SD	0.191	0.299	0.219	0.446	0.198	0.320	0.229	0.480
	0.3	Bias	-0.002	-0.065	0.049	0.043	0.001	-0.069	0.044	0.024
		SD	0.200	0.365	0.275	0.484	0.203	0.382	0.289	0.511
	0.4	Bias	-0.009	-0.102	0.047	-0.051	-0.005	-0.109	0.039	-0.073
		SD	0.188	0.395	0.339	0.527	0.199	0.429	0.341	0.561
$p = 4$										
10	0.1	Bias	0.025	-0.011	0.009	0.055	0.005	0.023	0.023	0.111
		SD	0.188	0.260	0.193	0.400	0.193	0.267	0.196	0.398
	0.2	Bias	0.003	0.014	0.030	0.062	-0.010	0.040	0.041	0.098
		SD	0.196	0.303	0.216	0.350	0.198	0.305	0.214	0.346
	0.3	Bias	-0.013	0.043	0.059	0.084	-0.013	0.046	0.059	0.090
		SD	0.198	0.341	0.250	0.326	0.195	0.338	0.246	0.326
	0.4	Bias	-0.022	0.075	0.097	0.106	-0.024	0.079	0.099	0.111
		SD	0.200	0.388	0.304	0.334	0.201	0.388	0.301	0.334
30	0.1	Bias	0.032	-0.088	0.010	0.051	0.019	-0.054	0.023	0.148
		SD	0.187	0.259	0.193	0.477	0.184	0.260	0.200	0.476
	0.2	Bias	0.016	-0.095	0.022	0.023	0.011	-0.080	0.035	0.056
		SD	0.196	0.301	0.218	0.462	0.197	0.305	0.226	0.467
	0.3	Bias	0.001	-0.094	0.042	0.009	0.007	-0.111	0.033	-0.003
		SD	0.197	0.348	0.270	0.486	0.195	0.325	0.260	0.468
	0.4	Bias	-0.015	-0.118	0.042	-0.070	-0.013	-0.126	0.041	-0.079
		SD	0.203	0.407	0.336	0.526	0.207	0.410	0.346	0.539

† Bias, empirical bias; SD, the empirical standard deviation.

corresponding estimating equations in general sense by involving minimizations of a series of L_1-type convex objective functions. It can be quickly and stably implemented, utilizing existing functions in software **R**. We also establish that the parameter estimates are strongly consistent and asymptotically normal. Both simulation studies and the analysis of follicular type lymphoma data illustrate that the proposed model is flexible and its estimator has good performance.

For standard survival setting with randomly censored data, quantile regression involves a serious issue of identifiability. To make sure the inference is valid, τ_U must satisfy certain intrinsic constraints imposed by model assumption (1) and regularity conditions C4 and C5. In practice, however, we recommend to conduct some tentative analysis which may shed light on appropriate choices for τ_U and τ_L in a τ-range of interest. Specifically, we can plot the cumulative incidence functions, or calculate the proportion of the events. Furthermore, it is worth pointing out that the proposed method needs to estimate $G_0(\cdot|\boldsymbol{Z})$, which inevitably suffers from the curse of dimensionality if $\boldsymbol{Z}$ is multi-dimensional. If the dimension of $\mathbf{Z}$ is high, we can firstly consider model-based dimension reduction on estimating $G_0(\cdot|\boldsymbol{Z})$, for example using Cox proportional hazard model between the censoring variable and covariates or developing variable selection method. That is our intention in further research to develop methods for such case with high dimensional covariates. In addition, the Kalpan-Meier estimates may be unstable at the right tails. This line of research provides the further investigation by the technique proposed in Zhou [29].

ACKNOWLEDGMENT

The authors are grateful to the Editor, Associate Editor and referees for their useful comments. Fan's work is partially sponsored by Shanghai Pujiang Program (16PJC041), Shanghai Young Teacher Training Scheme of Universities in 2016 (ZZSWM15019), Shanghai Summit and Plateau Discipline. Zhang's work is partially supported by National Natural Science Foundation of China (NSFC) (11401194) and the Fundamental Research Funds for the Central Universities (531107050739). Zhou's work is partially supported by National Natural Science Foundation of China (NSFC) (71271128), the State Key Program of National Natural Science Foundation of China (71331006 and 91546202), National Center for Mathematics and Interdisciplinary Sciences (NCMIS), Key Laboratory of RCSDS, Academy of Mathematics and Systems Science (AMSS), Chinese Academy of Sciences (CAS)(2008DP173182) and Shanghai First-class Discipline A, Program for Changjiang Scholars (PCSIRT) and Innovative Research Team in Shanghai University of Finance and Economics (SUFE)(IRT13077).

APPENDIX

Connection between $L_n(\boldsymbol{b},\tau)$ and $U_n(\boldsymbol{b},\tau)$.

Firstly we observe that the derivative of the first term of $L_n(\boldsymbol{b},\tau)$ wrt $\boldsymbol{b}$ equals

$$\sum_{i=1}^{n} I\left(\delta_i\epsilon_i=1\right)\frac{\boldsymbol{Z}_i}{\widehat{G}(X_i|\mathbf{Z}_i)}\Big[I\{H_\gamma(X_i)\le \boldsymbol{b}^T\mathbf{Z}_i\} - I\{H_\gamma(X_i)\ge \boldsymbol{b}^T\mathbf{Z}_i\}\Big],$$

which, after some simple calculation, can be written as

$$2\sum_{i=1}^{n} I\left(\delta_i\epsilon_i=1\right)\frac{\boldsymbol{Z}_i I\{H_\gamma(X_i)\le \boldsymbol{b}^T\mathbf{Z}_i\}}{\widehat{G}(X_i|\mathbf{Z}_i)} - \sum_{i=1}^{n} I\left(\delta_i\epsilon_i=1\right)\frac{\boldsymbol{Z}_i}{\widehat{G}(X_i|\mathbf{Z}_i)}.$$

Suppose that M bounds $|\boldsymbol{b}^T\sum_{l=1}^{n}\frac{-\mathbf{Z}_l I(\delta_l\epsilon_l=1)}{\widehat{G}(X_l|\mathbf{Z}_l)}|$ and $\left|\boldsymbol{b}^T\sum_{k=1}^{n}\frac{2\tau I(\delta_k=1)\mathbf{Z}_k}{\widehat{G}(X_i|\mathbf{Z}_i)}\right|$ from above for all $\boldsymbol{b}$'s in the compact parameter space for $\boldsymbol{\beta}_0$, the derivative of the rest two terms of $L_n(\boldsymbol{b},\tau)$ wrt $\boldsymbol{b}$ becomes

$$\sum_{i=1}^{n} I\left(\delta_i\epsilon_i=1\right)\frac{\boldsymbol{Z}_i}{\widehat{G}(X_i|\mathbf{Z}_i)} - \sum_{k=1}^{n}\frac{2\tau I(\delta_k=1)\boldsymbol{Z}_k}{\widehat{G}(X_k|\mathbf{Z}_k)}.$$

Then it follows that

$$\begin{aligned}\frac{\partial L_n(\boldsymbol{b},\tau)}{\partial \boldsymbol{b}} &= 2\sum_{i=1}^{n} I\left(\delta_i\epsilon_i=1\right)\frac{\boldsymbol{Z}_i I\{H_\gamma(X_i)\le \boldsymbol{b}^T\mathbf{Z}_i\}}{\widehat{G}(X_i|\mathbf{Z}_i)}\\ &- \sum_{i=1}^{n} I\left(\delta_i\epsilon_i=1\right)\frac{\boldsymbol{Z}_i}{\widehat{G}(X_i|\mathbf{Z}_i)}\\ &+ \sum_{i=1}^{n} I\left(\delta_i\epsilon_i=1\right)\frac{\boldsymbol{Z}_i}{\widehat{G}(X_i|\mathbf{Z}_i)}\\ &- \sum_{k=1}^{n}\frac{2\tau I(\delta_k=1)\boldsymbol{Z}_k}{\widehat{G}(X_k|\mathbf{Z}_k)}\\ &= 2\sqrt{n}U_n(\boldsymbol{b},\tau).\end{aligned}$$

Such a connection, combined with the fact that $L_n(\boldsymbol{b},\tau)$ is a convex function of $\boldsymbol{b}$, illustrates that the minimizer of $L_n(\boldsymbol{b},\tau)$ is equivalent to the solution of Equation (4).

Lemma 1. *Assume that conditions C4, C6 and C7 hold, then*

$$\begin{aligned}\|\widehat{G}-G_0\|_\infty &= \sup_t\sup_{\boldsymbol{z}}|\widehat{G}(t|\boldsymbol{z})-G_0(t|\boldsymbol{z})|\\ &= O_p\left((\log n/n)^{1/2}n^{\nu/2}+n^{-2\nu}\right).\end{aligned}$$

Proof. This follows directly from Theorem 2.1 of [10], and we omit the details for saving space. □

Proof of Theorem 1. Because Γ is compact, we suppose that $\widehat{\gamma}\to\gamma^*$. We now show that $\widehat{\boldsymbol{\beta}}(\widehat{\gamma},\tau)$ is bounded. Otherwise, for a subsequence, denoted by n, $\|\widehat{\boldsymbol{\beta}}(\widehat{\gamma},\tau)\|\to\infty$. Define

$$\begin{aligned}\widehat{\boldsymbol{\beta}}^*(\tau) &= \left\{1-\frac{1}{\|\widehat{\beta}(\widehat{\gamma},\tau)-\boldsymbol{\beta}(\gamma^*,\tau)\|}\right\}\boldsymbol{\beta}(\gamma^*,\tau)\\ &+ \frac{1}{\|\widehat{\boldsymbol{\beta}}(\widehat{\gamma},\tau)-\beta(\gamma^*,\tau)\|}\widehat{\boldsymbol{\beta}}(\widehat{\gamma},\tau).\end{aligned}$$

Note that $\widehat{\boldsymbol{\beta}}^*(\tau)$ is bounded and that its distance from $\boldsymbol{\beta}(\gamma^*,\tau)$ is 1. We can further choose a subsequence to assume that $\widehat{\boldsymbol{\beta}}^*(\tau)$ has a limit $\boldsymbol{\beta}^*(\tau)$

Define

$$\begin{aligned}\mathcal{F} &= \Bigg(\frac{\delta_i}{G_0(X_i|\mathbf{Z_i})}\mathbf{Z}_i\left[I\left\{H_\gamma(X_i)\le \mathbf{b}^T\mathbf{Z}_i,\epsilon_i=1\right\}-\tau\right],\\ &\qquad \tau\in[\tau_L,\tau_U],\gamma\in\Gamma,\|\mathbf{b}-\beta(\gamma^*,\tau)\|\le 1\Bigg).\end{aligned}$$

The function class $\mathcal{F}$ is Donsker and thus Glivenko-Cantelli class, and $\widehat{G}(t|\boldsymbol{z})\to G_0(t|\boldsymbol{z})$ uniformly in $t\in[0,\lambda]$. Then,

$$\begin{aligned}&\sup_{\tau\in[\tau_L,\tau_U],\mathbf{b}\in\mathcal{B}}\|n^{-1/2}\mathbf{U}_n^G(\mathbf{b},\tau,\gamma)-\Psi(\mathbf{b},\tau,\gamma)\|\\ &= o(1),a.s.\end{aligned}\tag{5}$$

where $\mathcal{B}=\{\mathbf{b}:\|\mathbf{b}-\boldsymbol{\beta}(\gamma^*,\tau)\|\le 1\}$.

Secondly, we note that for any $\boldsymbol{u}\in R^{p+1}$ satisfying $\|\boldsymbol{u}\|^2=1$, $\boldsymbol{u}^T\Psi(\boldsymbol{\beta}(\gamma^*,\tau)+\boldsymbol{u}\eta,\tau,\gamma^*)$ is increasing in η. Then for $\eta\ge\rho_0$,

$$\begin{aligned}&\boldsymbol{u}^T\left[\Psi(\boldsymbol{\beta}(\gamma^*,\tau)+\boldsymbol{u}\eta,\tau,\gamma^*)-\Psi(\boldsymbol{\beta}(\gamma^*,\tau),\tau,\gamma^*)\right]\\ &\ge \boldsymbol{u}^T\left[\Psi(\boldsymbol{\beta}(\gamma^*,\tau)+\boldsymbol{u}\rho_0,\tau,\gamma^*)-\Psi(\boldsymbol{\beta}(\gamma^*,\tau),\tau,\gamma^*)\right]\ge 0.\end{aligned}$$

By the Cauch-Schawarz inequality,

$$\begin{aligned} & \|\Psi(\boldsymbol{\beta}(\gamma^*,\tau)+\boldsymbol{u}\eta,\tau,\gamma^*)-\Psi(\boldsymbol{\beta}(\gamma^*,\tau),\tau,\gamma^*)\|^2.\|\boldsymbol{u}\|^2 \\ \geq\ & \left(\boldsymbol{u}^T\left[\Psi(\boldsymbol{\beta}(\gamma^*,\tau)+\boldsymbol{u}\eta,\tau,\gamma^*)-\Psi(\boldsymbol{\beta}(\gamma^*,\tau),\tau,\gamma^*)\right]\right)^2 \\ \geq\ & \left(\boldsymbol{u}^T\left[\Psi(\boldsymbol{\beta}(\gamma^*,\tau)+\boldsymbol{u}\rho,\tau,\gamma^*)-\Psi(\boldsymbol{\beta}(\gamma^*,\tau),\tau,\gamma^*)\right]\right)^2 \\ \geq\ & c_0^2\rho_0^2. \end{aligned}$$

Therefore, we have

$$\inf_{\boldsymbol{b}\overline{\in}\mathcal{B}(\rho_0)}\|\Psi(\boldsymbol{b},\tau,\gamma^*)-\Psi(\boldsymbol{\beta}(\gamma^*,\tau),\tau,\gamma^*)\|\geq c_0\rho_0.$$

Some simple algebraic manipulation shows that

$$\begin{aligned} & \Psi(\widehat{\boldsymbol{\beta}}^*(\tau),\tau,\gamma^*)-\Psi(\boldsymbol{\beta}(\gamma^*,\tau),\tau,\gamma^*) \\ & = n^{-1/2}\mathbf{U}_n^G(\widehat{\boldsymbol{\beta}}^*(\tau),\tau,\gamma^*)-\Psi(\boldsymbol{\beta}(\gamma^*,\tau),\tau,\gamma^*) \\ & -\left[n^{-1/2}\mathbf{U}_n^G(\widehat{\boldsymbol{\beta}}^*(\tau),\tau,\gamma^*)-\Psi(\widehat{\boldsymbol{\beta}}^*(\tau),\tau,\gamma^*)\right]. \end{aligned}$$

Then we can obtain that
$\Psi(\widehat{\boldsymbol{\beta}}^*(\tau),\tau,\gamma^*)-\Psi(\boldsymbol{\beta}(\gamma^*,\tau),\tau,\gamma^*)=o(1)$ from (5). This is a contradiction, because $\|\widehat{\boldsymbol{\beta}}^*(\tau)-\boldsymbol{\beta}(\gamma^*,\tau)\|=1$. Hence we can show that $\widehat{\boldsymbol{\beta}}(\widehat{\gamma},\tau)$ is bounded, and

$$\sup_{\tau\in[\tau_L,\tau_U]}\|\widehat{\boldsymbol{\beta}}(\widehat{\gamma},\tau)-\boldsymbol{\beta}(\gamma^*,\tau)\|\to 0.$$

Next, if we can show that $\gamma^*=\gamma_0$, and $\boldsymbol{\beta}(\gamma^*,\tau)=\boldsymbol{\beta}_0(\tau)$, the result is true. By the Glivenko-Cantelli theorem, for any compact set $\mathcal{A}$

$$\begin{aligned} & \sup_{\boldsymbol{\beta}\in\mathcal{A},\gamma\in\Gamma}\Bigg|n^{-1}\sum_{i=1}^n I(\mathbf{Z}_i\leq\mathbf{z})\Big(\frac{\delta_i}{\widehat{G}(X_i|\mathbf{Z_i})}[I\{H_\gamma(X_i)\leq\boldsymbol{\beta}^T\mathbf{Z}_i, \\ & \epsilon_i=1\}-\tau]\Big)-E\left\{I(\mathbf{Z}\leq\mathbf{z})[I\{H_\gamma(T)\leq\boldsymbol{\beta}^T\mathbf{Z},\epsilon=1\}-\tau]\right\}\Bigg| \\ & \xrightarrow{a.s.} 0. \end{aligned}$$

Thus

$$\begin{aligned} n^{-3}R_n(\widehat{\gamma},\tau)\to E\Big(E\big[I(\mathbf{Z}\leq\mathbf{z})[I\{H_{\gamma^*}(T) \\ \leq\boldsymbol{\beta}(\gamma^*,\tau)^T\mathbf{Z},\epsilon=1\}-\tau]\big]^2|\mathbf{z}=\mathbf{Z}\Big). \end{aligned}$$

Because $n^{-3}R_n(\gamma_0,\tau)\to 0$ and $n^{-3}R_n(\gamma_0,\tau)\geq n^{-3}R_n(\widehat{\gamma},\tau)$, we conclude that

$$\begin{aligned} & E\big\{I(\mathbf{Z}\leq\mathbf{z})\{I(H_{\gamma^*}(T) \\ & \quad\leq\boldsymbol{\beta}(\gamma^*,\tau)^T\mathbf{Z},\epsilon=1)-\tau\}|\mathbf{z}=\mathbf{Z}\big\}=0, \end{aligned}$$

that is $\Pr\left[\{H_{\gamma^*}(T)\leq\boldsymbol{\beta}(\gamma^*,\tau)^T\mathbf{Z},\epsilon=1\}|\mathbf{Z}\right]=\tau$. From condition(C5.), this gives

$$\boldsymbol{\beta}(\gamma^*,\tau)^T\mathbf{Z}=H_{\gamma^*}[H_{\gamma_0}^{-1}\{\boldsymbol{\beta}_0(\tau)^T\mathbf{Z}\}].$$

From condition (C2.), we get that $\gamma^*=\gamma_0$ and $\boldsymbol{\beta}(\gamma^*,\tau)=\boldsymbol{\beta}_0(\tau)$. □

Proof of Theorem 2. We know that $\widehat{\boldsymbol{\beta}}(\widehat{\gamma},\tau)$ is the solution to the equation of

$$n^{-1}\sum_{i=1}^n\frac{\delta_i}{\widehat{G}(X_i|\mathbf{Z_i})}\mathbf{Z}_i\left[I\left\{H_{\widehat{\gamma}}(X_i)\leq\beta^T\mathbf{Z}_i,\epsilon_i=1\right\}-\tau\right]=0.$$

From the uniform expansion ([8]),

$$\sqrt{n}\{\widehat{G}(t)-G(t)\}=n^{-1/2}\sum_{i=1}^n V(X_i,\delta_i;t)+o_p(1),$$

where $V(X_i,\delta_i;t)$ is the influence function for the Kaplan-Meier estimator for the censoring distribution, that is,

$$V(X_i,\delta_i;t)=-G_0(t)\int_0^t\frac{\widehat{G}(u-)dM_i(u)}{G_0(u)\sum_{i=1}^n I(X_i\geq u)},$$

where $M_i(t)$ is the martingale for the censoring time. If C is dependent on the covariates, from Theorem 2.3 of [10] and the proof of Theorem 2 in [27], with assumptions C3, C4, C6, and C7, we have that

$$\begin{aligned} (6)\qquad & \widehat{G}(t|\boldsymbol{z})-G_0(t|\boldsymbol{z}) \\ & = (nh_n)^{-1}\sum_{i=1}^n K\left((z-Z_i)/h_n\right)G_0(t|\boldsymbol{z})\cdot \\ & \quad \varphi(X_i,\delta_i\epsilon_i,t,z)+O_p\left((\log n/nh_n)^{3/4}+h_n^2\right), \end{aligned}$$

where

$$\begin{aligned} \varphi(X_i,\delta_i\epsilon_i,t,z)= & \int_0^{X_i\wedge t}\frac{-g_0(s|\boldsymbol{z})ds}{\{G_0(s|\boldsymbol{z})\}^2\{1-F_1(s|\boldsymbol{z})\}} \\ & +\frac{I(\delta_i\epsilon_i=0))I(X_i\leq t)}{G_0(X_i|\boldsymbol{z})\{1-F_1(X_i|\boldsymbol{z})\}}. \end{aligned}$$

Noting the Donsker property of the class

$$\Bigg\{\frac{\delta}{G_0(X|\boldsymbol{Z})}\left[I\left\{H_\gamma(X)\leq\beta^T\mathbf{Z},\epsilon=1\right\}-\tau\right]: \text{G}_0 \text{ is decreasing bounded away from 0, } \gamma \text{ and } \boldsymbol{\beta} \text{ are in the neighborhood of } \gamma_0 \text{ and } \boldsymbol{\beta}_0, \tau\in[\tau_L,\tau_U]\Bigg\},$$

we obtain that

$$\begin{aligned} & -\sqrt{n}E\left(\boldsymbol{Z}[I\{H_{\widehat{\gamma}}(T)-\widehat{\boldsymbol{\beta}}(\widehat{\gamma},\tau)^T\boldsymbol{Z}\leq 0,\epsilon=1\}-\tau]\right) \\ = & \sqrt{n}(\mathbf{P}_n-\mathbf{P})Q_0(X,Z,\delta,\epsilon)+o_p(1), \end{aligned}$$

where $\mathbf{P}_n$ denotes the empirical measure, $\mathbf{P}$ is the expectation, and

$$(7)\qquad Q_0(X,Z,\delta,\epsilon)$$

$$= \frac{\delta}{G_0(X|\boldsymbol{Z})}\boldsymbol{Z}\left[I\{H_{\gamma_0}(X)-\boldsymbol{\beta}_0(\tau)^T\boldsymbol{Z}\le 0,\epsilon=1\}-\tau\right] -\widetilde{E}\Bigg(\frac{\widetilde{\delta}}{G_0^2(\widetilde{X}|\boldsymbol{Z})}\widetilde{\boldsymbol{Z}}\big[I\left\{H_{\gamma_0}(\widetilde{X})-\boldsymbol{\beta}_0(\tau)^T\widetilde{\boldsymbol{Z}}\le 0,\ \widetilde{\epsilon}=1\right\} - \ \tau\big]\left\{\widehat{G}(\widetilde{X}|\boldsymbol{Z})-G_0(\widetilde{X}|\boldsymbol{Z})\right\}\Bigg).$$

Plugging (6) into (7), we can get that

$$\begin{aligned} &Q_0(X,Z,\delta,\epsilon)\\ =\ & \frac{\delta}{G_0(X|\boldsymbol{Z})}\boldsymbol{Z}\left[I\{H_{\gamma_0}(X)-\boldsymbol{\beta}_0(\tau)^T\boldsymbol{Z}\le 0,\epsilon=1\}-\tau\right]\\ -\ & \widetilde{E}\Bigg\{\frac{\widetilde{\delta}}{G_0^2(\widetilde{X}|\boldsymbol{Z})}\widetilde{\boldsymbol{Z}}\big[I\left\{H_{\gamma_0}(\widetilde{X})-\boldsymbol{\beta}_0(\tau)^T\widetilde{\boldsymbol{Z}}\le 0,\ \widetilde{\epsilon}=1\right\}\\ &-\tau\big]f_{\boldsymbol{z}}(\widetilde{\boldsymbol{Z}})G_0(\widetilde{X}|\boldsymbol{Z})\varphi(X,\delta\epsilon,\widetilde{X},Z)\Bigg\}, \end{aligned}$$

where $(\widetilde{X},\widetilde{\boldsymbol{Z}},\widetilde{\delta},\widetilde{\epsilon})$ are iid copies of $(X,\mathbf{Z},\delta,\epsilon)$ and $\widetilde{E}(\cdot)$ takes the expectation with respect to $(\widetilde{X},\widetilde{\boldsymbol{Z}},\widetilde{\delta},\widetilde{\epsilon})$ only. Therefore, after Taylor expansion of the left side, we have that

$$\begin{aligned} &-\sqrt{n}\Big[E\left\{\boldsymbol{Z}\Psi_{0\gamma}(\mathbf{Z})\right\}(\widehat{\gamma}-\gamma_0)\\ &\qquad + E\left\{\boldsymbol{Z}\Psi_{0\boldsymbol{b}}(\mathbf{Z})\right\}\{\widehat{\boldsymbol{\beta}}(\widehat{\gamma},\tau)-\beta_0(\tau)\}\Big]\\ =\ & \sqrt{n}(\mathbf{P}_n-\mathbf{P})Q_0(X,Z,\delta,\epsilon)+o_p(1+\sqrt{n}|\widehat{\gamma}-\gamma_0|\\ &+\sqrt{n}\|\widehat{\boldsymbol{\beta}}(\widehat{\gamma},\tau)-\boldsymbol{\beta}_0(\tau)\|). \end{aligned}$$

Then we obtain that

$$\begin{aligned} (8)\qquad & \sqrt{n}\left[-E\{\boldsymbol{Z}\Psi_{0\boldsymbol{b}}(\mathbf{Z})\}\right]A_1(\widehat{\gamma}-\gamma_0)\\ &+\sqrt{n}E\left\{\boldsymbol{Z}\Psi_{0\boldsymbol{b}}(\mathbf{Z})\right\}\left\{\widehat{\boldsymbol{\beta}}(\widehat{\gamma},\tau)-\beta_0(\tau)\right\}\\ &=-\sqrt{n}(\mathbf{P}_n-\mathbf{P})Q_0(X,Z,\delta,\epsilon)+o_p(1), \end{aligned}$$

where $A_1=-[E\{\boldsymbol{Z}\Psi_{0\boldsymbol{b}}(\mathbf{Z})\}]^{-1}E\{\boldsymbol{Z}\Psi_{0\gamma}(\mathbf{Z})\}$.

On the other hand, we note that the class of functions

$$\Bigg\{I(\mathbf{Z}\le\mathbf{z})\frac{\delta}{G(X|\boldsymbol{Z})}\left[I\left\{H_\gamma(X)\le\beta^T\mathbf{Z},\epsilon=1\right\}-\tau\right]: \text{G is decreasing bounded away from 0, } \gamma \text{ and } \boldsymbol{\beta} \text{ are in the neighborhood of } \gamma_0 \text{ and } \boldsymbol{\beta}_0,\ \tau\in[\tau_L,\tau_U]\Bigg\}$$

is a Donsker class. Thus

$$\begin{aligned} \sup_{\mathbf{z},\gamma\in\Gamma}\Bigg| & n^{-1}D_n(\mathbf{z},\gamma)-E\Bigg\{I(\mathbf{Z}\le\mathbf{z})\frac{\delta}{\widehat{G}(X|\boldsymbol{Z})}\\ &\left[I\left\{H_\gamma(X)\le\widehat{\boldsymbol{\beta}}(\gamma,\tau)^T\boldsymbol{Z},\epsilon=1\right\}-\tau\right]\Bigg\}\\ &-(\mathbf{P_n}-\mathbf{P})\Big[I(\mathbf{Z}\le\mathbf{z})\frac{\delta}{G_0(X|\boldsymbol{Z})}\\ &\left\{I\left(H_\gamma(X)\le\beta(\gamma,\tau)^T\boldsymbol{Z},\epsilon=1\right)-\tau\right\}\Big]\Bigg|\\ &=o_p(n^{-1/2}). \end{aligned}$$

In addition,

$$\begin{aligned} &E\Big[I(\mathbf{Z}\le\mathbf{z})\frac{\delta}{\widehat{G}(X|\boldsymbol{Z})}\{I\left(H_\gamma(X)\le\widehat{\boldsymbol{\beta}}(\gamma,\tau)^T\boldsymbol{Z},\epsilon=1\right)\\ &\qquad -\tau\}\Big]\\ &=E\Bigg\{I(\mathbf{Z}\le\mathbf{z})\frac{\delta}{G_0(X|\boldsymbol{Z})}\big[I\{H_\gamma(X_i)\le\widehat{\boldsymbol{\beta}}(\gamma,\tau)^T\boldsymbol{Z},\\ &\epsilon=1\}-\tau\big]\Bigg\}-\{1+o_p(1)\}E\Bigg[I(\mathbf{Z}\le\mathbf{z})\frac{\delta}{G_0(X|\boldsymbol{Z})^2}\\ &\left\{I\left(H_\gamma(X)\le\widehat{\boldsymbol{\beta}}(\gamma,\tau)^T\boldsymbol{Z},\epsilon=1\right)\right\}\\ &\qquad\left(\widehat{G}(X|\boldsymbol{Z})-G_0(X|\boldsymbol{Z})\right)\Bigg]. \end{aligned}$$

After Taylor expansion of the first term at $\boldsymbol{\beta}(\gamma,\tau)$ and using the expansion of $\widehat{G}$ in (6), we obtain that, uniformly in $\mathbf{z}$ and γ in a neighborhood of γ_0,

$$\begin{aligned} & n^{-1}D_n(\mathbf{z},\gamma)\\ =\ & E\left\{I(\mathbf{Z}\le\mathbf{z})\left[I\left\{H_\gamma(T)\le\boldsymbol{\beta}(\gamma,\tau)^T\boldsymbol{Z},\epsilon=1\right\}-\tau\right]\right\}\\ &+(\mathbf{P_n}-\mathbf{P})Q_1(X,Z,\delta:\gamma,\mathbf{z})+E\left\{I(\mathbf{Z}\le\mathbf{z})\Psi_{0b}(Z)\right\}\\ &\qquad(\widehat{\boldsymbol{\beta}}(\gamma,\tau)-\boldsymbol{\beta}(\gamma,\tau))\\ &+o_p(|\gamma-\gamma_0|+\|\widehat{\boldsymbol{\beta}}(\gamma,\tau)-\boldsymbol{\beta}(\gamma,\tau)\|)+o_p(n^{-1/2}), \end{aligned}$$

where

$$\begin{aligned} & Q_1(X,Z,\delta,\epsilon:\gamma,\mathbf{z})\\ =\ & I(\mathbf{Z}\le\mathbf{z})\frac{\delta}{G_0(X|\boldsymbol{Z})}\big[I\left\{H_\gamma(X)\le\boldsymbol{\beta}(\gamma,\tau)^T\boldsymbol{Z},\epsilon=1\right\}\\ &-\tau\big]-\widetilde{E}\Bigg\{I(\widetilde{\boldsymbol{Z}}\le\mathbf{z})\frac{\widetilde{\delta}}{G_0(\widetilde{X}|\boldsymbol{Z})^2}\\ &\left[I\left\{H_\gamma(\widetilde{X})-\boldsymbol{\beta}(\gamma,\tau)^T\widetilde{\boldsymbol{Z}}\le 0,\widetilde{\epsilon}=1\right\}-\tau\right]\\ &f_{\boldsymbol{z}}(\widetilde{\boldsymbol{Z}})G_0(\widetilde{X}|\boldsymbol{Z})\varphi(X,\delta\epsilon,\widetilde{X},Z)\Bigg\}. \end{aligned}$$

Hence,

$$\begin{aligned} & n^{-3}R_n(\gamma)\\ =\ & E\Big(\Big[E\{I(\mathbf{Z}\le\mathbf{z})\big[I\left\{H_\gamma(T)\le\boldsymbol{\beta}(\gamma,\tau)^T\boldsymbol{Z},\epsilon=1\right\} \end{aligned}$$

$$- \tau]\}\Big]^2 |\mathbf{z} = \mathbf{Z}\Big)$$

$$+ 2E\Big(E\Big[I(\mathbf{Z} \le \mathbf{z})\{I\left(H_\gamma(T) \le \boldsymbol{\beta}(\gamma, \tau)^T \boldsymbol{Z}, \epsilon = 1\right) - \tau\}\Big] \times (\mathbf{P_n} - \mathbf{P})Q_1(X, Z, \delta; \gamma, \mathbf{z})|\mathbf{z} = \mathbf{Z}\Big)$$

$$+ 2E\Big(E\Big[I(\mathbf{Z} \le \mathbf{z})\{I\left(H_\gamma(T) \le \boldsymbol{\beta}(\gamma, \tau)^T \boldsymbol{Z}, \epsilon = 1\right) - \tau\}\Big] \times E\left\{I(\mathbf{Z} \le \mathbf{z})\Psi_{0\mathbf{b}}(Z)\right\}\{\widehat{\boldsymbol{\beta}}(\gamma, \tau) - \boldsymbol{\beta}(\gamma, \tau)\}|\mathbf{z} = \mathbf{Z}\Big)$$

$$+ o_p(|\gamma - \gamma_0|^2 + \|\widehat{\boldsymbol{\beta}}(\gamma, \tau) - \boldsymbol{\beta}(\gamma, \tau)\|^2) + o_p(n^{-1}).$$

Note that in a neighborhood of γ_0,

$$\begin{aligned} & E\left\{I(\mathbf{Z} \le \mathbf{z})\left[I\left\{H_\gamma(T) \le \boldsymbol{\beta}(\gamma, \tau)^T \boldsymbol{Z}, \epsilon = 1\right\} - \tau\right]\right\} \\ = \ & B(\mathbf{z})(\gamma - \gamma_0) + o(|\gamma - \gamma_0|), \end{aligned}$$

where

$$\begin{aligned} B(\mathbf{z}) &= E\left\{I(\mathbf{Z} \le \mathbf{z})\Psi_{0\gamma}(Z)\right\} + E\left\{I(\mathbf{Z} \le \mathbf{z})\Psi_{0\mathbf{b}}(Z)\right\}^T A_1 \\ &= A_3(\mathbf{z}) + A_2(\mathbf{z})^T A_1, \end{aligned}$$

with $A_3(\mathbf{z}) = E\left\{I(\mathbf{Z} \le \mathbf{z})\Psi_{0\gamma}(Z)\right\}$, $A_2(\mathbf{z}) = E\left\{I(\mathbf{Z} \le \mathbf{z})\Psi_{0\mathbf{b}}(Z)\right\}$. Therefore $n^{-3}R_n(\gamma)$ has a quadratic expansion near γ_0 as

$$\begin{aligned} & E\{B(\mathbf{Z})^2\}(\gamma - \gamma_0)^2 \\ & + 2(\gamma - \gamma_0)\Big[(\mathbf{P_n} - \mathbf{P})\widetilde{E}\{B(\widetilde{\mathbf{Z}})Q_1(X, Z, \delta, \epsilon; \gamma_0, \widetilde{\mathbf{z}})\} \\ & + E\{B(Z)A_2(\mathbf{Z})^T\}\{\widehat{\boldsymbol{\beta}}(\widehat{\gamma}, \tau) - \boldsymbol{\beta}(\gamma, \tau)\}\Big] \\ & + o_p(|\gamma - \gamma_0|^2 + \|\widehat{\boldsymbol{\beta}}(\gamma, \tau) - \boldsymbol{\beta}(\gamma, \tau)\|^2) + o_p(n^{-1}). \end{aligned}$$

Because $\widehat{\gamma}$ minimize $R_n(\gamma)$, we obtain that

$$\begin{aligned} & E\{B(\mathbf{Z})^2\}(\gamma - \gamma_0) + E\{B(Z)A_2(\mathbf{Z})^T\}\{\widehat{\boldsymbol{\beta}}(\widehat{\gamma}, \tau) - \boldsymbol{\beta}(\widehat{\gamma}, \tau)\} \\ & = -(\mathbf{P_n} - \mathbf{P})\widetilde{E}\{B(\widetilde{\mathbf{Z}})Q_1(X, Z, \delta, \epsilon; \gamma_0, \widetilde{\mathbf{z}})\} \\ & \quad + o_p(|\widehat{\gamma} - \gamma_0| + \|\widehat{\boldsymbol{\beta}}(\widehat{\gamma}, \tau) - \boldsymbol{\beta}(\gamma, \tau)\|) + o_p(n^{-1/2}). \end{aligned}$$

Using the expansion of $\boldsymbol{\beta}(\widehat{\gamma}, \tau) = \boldsymbol{\beta}_0(\tau) + A_1(\widehat{\gamma} - \gamma_0) + o_p(|\widehat{\gamma} - \gamma_0|)$, we have that

$$\begin{aligned} (9) \quad & \sqrt{n}\Big\{E\left[A_3(\mathbf{Z})^2\right] + A_1^T E\left[A_2(\mathbf{Z})A_3(\mathbf{Z})\right]\Big\}(\gamma - \gamma_0) \\ & + \sqrt{n}\Big[E\left\{A_3(\mathbf{Z})A_2(\mathbf{Z})^T\right\} + A_1^T E\left\{A_2(\mathbf{Z})A_2(\mathbf{Z})^T\right\}\Big] \Big\{\widehat{\boldsymbol{\beta}}(\widehat{\gamma}, \tau) - \boldsymbol{\beta}_0(\tau)\Big\} \\ = \ & -\sqrt{n}(\mathbf{P_n} - \mathbf{P})\widetilde{E}\{B(\widetilde{\mathbf{Z}})Q_1(X, Z, \delta, \epsilon; \gamma_0, \widetilde{\mathbf{z}})\} \\ & + o_p(\sqrt{n}|\widehat{\gamma} - \gamma_0| + \sqrt{n}\|\widehat{\boldsymbol{\beta}}(\widehat{\gamma}, \tau) - \boldsymbol{\beta}(\gamma, \tau)\| + 1). \end{aligned}$$

Combining (9) with (8), we note that the coefficient matrix

$$\Lambda = \begin{pmatrix} -E\{\boldsymbol{Z}\Psi_{0\boldsymbol{b}}(\mathbf{Z})\}A_1 & E[\boldsymbol{Z}\Psi_{0\boldsymbol{b}}(\mathbf{Z})] \\ E\{A_3(\mathbf{Z})^2\} + A_1^T E\{A_2(\mathbf{Z})A_3(\mathbf{Z})\} & E\{A_3(\mathbf{Z})A_2(\mathbf{Z})^T\} + A_1^T E\{A_2(\mathbf{Z})A_2(\mathbf{Z})^T\} \end{pmatrix}$$

is nonsingular, because it has same rank as

$$\begin{pmatrix} 0 & E\{\boldsymbol{Z}\Psi_{0\boldsymbol{b}}(\mathbf{Z})\} \\ E\{B(\mathbf{Z})^2\} & E\{A_3(\mathbf{Z})A_2(\mathbf{Z})^T\} + A_1^T E\{A_2(\mathbf{Z})A_2(\mathbf{Z})^T\} \end{pmatrix},$$

which has full rank. Thus the asymptotic covariance is given by

$$\Sigma = \Lambda^{-1} E\left(\begin{bmatrix} Q_0(X, Z, \delta, \epsilon) \\ \widetilde{E}\{B(\widetilde{\mathbf{Z}})Q_1(X, Z, \delta, \epsilon; \gamma_0, \widetilde{\mathbf{z}})\} \end{bmatrix}^{\otimes 2}\right)(\Lambda^{-1})^T,$$

where$\mathbf{x}^{\otimes 2} = \mathbf{x}\mathbf{x}^T$. We just consider the case in which γ is estimated by minimizing $R_n(\gamma)$, the same arguments can be applied to the situation on $R_n^*(\gamma)$. □

Received 27 February 2014

REFERENCES

[1] Bang, H. and Tsiatis, A. A. (2002). Median regression with censored cost data. *Biometrics* **58** 643–649. MR1926117

[2] Benichou, J. and Gail, M. H. (1990). Estimates of absolute cause-specific risk in cohort studies. *Biometrics* **46** 813–826.

[3] Box, G. E. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)* **26** 211–252. MR0192611

[4] Cai, T., Tian, L. and Wei, L. (2005). Semiparametric Box-Cox power transformation models for censored survival observations. *Biometrika* **92** 619–632. MR2202650

[5] Cheng, S., Wei, L. and Ying, Z. (1995). Analysis of transformation models with censored data. *Biometrika* **82** 835–845. MR1380818

[6] Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. New York, Chapman & Hall. MR1270903

[7] Fine, J. P. and Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association* **94** 496–509. MR1702320

[8] Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis*. New York, Wiley. MR1100924

[9] Gaynor, J. J., Feuer, E. J., Tan, C. C., Wu, D. H., Little, C. R., Straus, D. J., Clarkson, B. D. and Brennan, M. F. (1993). On the use of cause-specific failure and conditional failure probabilities: examples from clinical oncology data. *Journal of the American Statistical Association* **88** 400–409.

[10] Gonzalez-Manteiga, W. and Cadarso-Suarez, C. (1994). Asymptotic properties of a generalized Kaplan-Meier estimator with some applications. *Communications in Statistics-Theory and Methods* **4** 65–78. MR1366364

[11] Gray, R. J. (1988). A class of K-sample tests for comparing the cumulative incidence of a competing risk. *The Annals of Statistics* **16** 1141–1154. MR0959192

[12] Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. 2nd ed. New York, John Wiley & Sons. MR1924807

[13] Khan, S. and Tamer, E. (2007). Partial rank estimation of duration models with general forms of censoring. *Journal of Econometrics* **136** 251–280. MR2328593
[14] Koenker, R. and Geling, O. (2001). Reappraising medfly longevity: a quantile regression survival analysis. *Journal of the American Statistical Association* **96** 458–468. MR1939348
[15] Korn, E. L. and Dorey, F. J. (1992). Applications of crude incidence curves. *Statistics in Medicine* **11** 813–829.
[16] Larson, M. (1984). Covariate analysis of competing-risks data with log-linear models. *Biometrics* **40** 459–469.
[17] Leng, C. and Tong, X. W. (2014). Censored quantile regression via Box-Cox transformation under conditional independence. *Statistica Sinica* **24** 221–249. MR3183682
[18] Mu, Y. and He, X. (2007). Power transformation toward a linear regression quantile. *Journal of the American Statistical Association* **102** 269–279. MR2293308
[19] Muller, H. G. (1988). *Nonparametric Regression Analysis of Longitudinal Data*. Springer, New York. MR0960887
[20] Peng, L. and Fine, J. (2007). Nonparametric quantile inference with competing–risks data. *Biometrika* **94** 735–744. MR2410020
[21] Peng, L. and Fine, J. P. (2009). Competing risks quantile regression. *Journal of the American Statistical Association* **104** 1440–1453. MR2596999
[22] Pepe, M. S. (1991). Inference for events with dependent risks in multiple endpoint studies. *Journal of the American Statistical Association* **86** 770–778. MR1147104
[23] Petersen, P., Gospodarowicz, M., Tsang, R., Pintilie, M., Wells, W., Hodgson, D., Sun, A., Crump, M., Patterson, B. and Bailey, D. (2004). Long-term outcome in stage I and II follicular lymphoma following treatment with involved field radiation therapy alone. *Journal of Clinical Oncology* **22** 6521.
[24] Pintilie, M. (2006). *Competing Risks: A Practical Perspective*. New York, John Wiley & Sons.
[25] Prentice, R. L., Kalbfleisch, J. D., Peterson Jr, A. V., Flournoy, N., Farewell, V. and Breslow, N. (1978). The analysis of failure times in the presence of competing risks. *Biometrics* **34** 541–554.
[26] Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences* **72** 20–22. MR0356425
[27] Wang, H. J. and Wang, L. (2009). Locally weighted censored quantile regression. *Journal of the American Statistical Association* **104** 1117–1128. MR2562007
[28] Yin, G., Zeng, D. and Li, H. (2008). Power-transformed linear quantile regression with censored data. *Journal of the American Statistical Association* **103** 1214–1224. MR2462894
[29] Zhou, L. (2006). A simple censored median regression estimator. *Statistica Sinica* **16** 1043–1058. MR2281314

Caiyun Fan
School of Statistics and Information
Shanghai University of International Business and Economics
Shanghai 201620
P.R. China
E-mail address: fancaiyun2010@163.com

Feipeng Zhang
School of Finance and Statistics
Hunan University
Changsha 410082
P.R. China
E-mail address: zhangfp108@163.com

Yong Zhou
School of Statistics and Management
Shanghai University of Finance and Economics
Shanghai 200433
P.R. China

Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Beijing 100190
P.R. China
E-mail address: yzhou@amss.ac.cn