IRTG 1792 Discussion Paper 2021-013

Penalized Weigted Competing Risks Models Based on Quantile Regression

Erqian Li [*]
Wolfgang Karl Härdle [*2] [*3] [*4] [*5] [*6]
Xiaowen Dai [*7]
Maozai Tian [*8] [*9]

[*] North China University of Technology, China
[*2] Humboldt-Universität zu Berlin, Germany
[*3] Xiamen University, China
[*4] Singapore Management University, Singapore
[*5] Charles University, Czech Republic
[*6] National Chiao Tung University, Taiwan
[*7] Shanghai Lixin University of Accounting and Finance, China
[*8] Renmin University of China, China
[*9] Xinjiang Medical University, China

This research was supported by the Deutsche Forschungsgesellschaft through the International Research Training Group 1792 "High Dimensional Nonstationary Time Series".

http://irtg1792.hu-berlin.de
ISSN 2568-5619

International Research Training Group 1792

PENALIZED WEIGTED COMPETING RISKS MODELS BASED ON QUANTILE REGRESSION

BY ERQIAN LI[1], WOLFGANG KARL HÄRDLE[2,*] XIAOWEN DAI[3,†] AND MAOZAI TIAN[4,‡]

[1]College of Science, North China University of Technology, Beijing, China li2qian@ncut.edu.cn

[2]Center for Applied Statistics and Economics, Humboldt-Universit?t zu Berlin,10178 Berlin, Germany
[*]wolfgang.k.haerdle@icloud.com

[3] School of Statistics and Mathematics, Shanghai Lixin University of Accounting and Finance, Shanghai, China
[†]daixiaowendaisy@163.com

[4]Center for Applied Statistics, School of Statistics, Renmin University of China, Beijing 100872,
College of Medical Engineering and Technology, Xinjiang Medical University, Ürümqi 830011, China [‡]mztian@ruc.edu.cn

The proportional subdistribution hazards (PSH) model is popularly used to deal with competing risks data. Censored quantile regression provides an important supplement as well as variable selection methods, due to large numbers of irrelevant covariates in practice. In this paper, we study variable s-election procedures based on penalized weighted quantile regression for competing risks models, which is conveniently applied by researchers. Asymptot-ic properties of the proposed estimators including consistency and asymptot-ic normality of non-penalized estimator and consistency of variable selection are established. Monte Carlo simulation studies are conducted, showing that the proposed methods are considerably stable and efficient. A real data about bone marrow transplant (BMT) is also analyzed to illustrate the application of proposed procedure.

1

1. Introduction. In survival analysis, sometimes events fail because of a specific cause and some other causes or competing risks, which were usually treated as censoring. However, when the competing risks are affected by covariates, the usual Cox modeling may be inaccurate. To deal with competing risks data, [3] proposed a novel semiparametric proportional hazards for the subdistribution, or PSH models, which analyze directly the effect of covariates on the marginal probability function, or cumulative incidence function (CIF). The competing risks data often occur in clinical trials containing large numbers of covariates, among which, only few of them have significant or essential influence on the response, generating the variable selection issues, such as the general penalized log-partial likelihood method proposed by [4].

Quantile regression introduced by [8] has been widely known to more comprehensively describe conditional distribution of response on covarites. Existing work about competing risks quantile regression includes [11], which firstly transforms competing risks quantile regression models to accelerated failure model and uses an estimating equation procedure for estimation. Besides, [15] discussed the quantile regression for competing risks data with missing cause of failure. Then [1] and [10] developed variable selection procedures based on unbiased estimating equations with group structures and penalization methods for competing risks quantile regression models.

MSC2020 subject classifications: Primary 62N02, 62P10; secondary 62F12.

Keywords and phrases: Competing risks, Cumulative incidence function, Kaplan-Meier estimator, Redistribution method.

[1]The corresponding author is Professor Maozai Tian.

In the paper, in spite of the estimating equation method, we propose to develop a more general method for competing risks quantile regression, and expand the weighted procedures by considering the re-distribution methods of [17] to the PSH model. By the transformed responses, we can rewrite the competing risks quantiles formulation as the general quantile regression objective function, then apply the constructed weights. With unbiasedness of the subgradient of this weighted objective function at the true cumulative-incidence function and coefficient proved, consistency and asymptotic normality of the penalty-free estimators are established under regularity conditions. Then to realize the variable selection, penalization methods such as the least absolute shrinkage and selection operator (LASSO) proposed by [16] and the adaptive LASSO (ALASSO) developed by [20] are applied to the weighted objective function, which can be easily applied with R package. The consistency of the variable selection procedure is also established, while Monte Carlo simulation is performed to illustrate the efficiency and stability of our proposed procedures. A real data about bone marrow transplant is analyzed using our methods.

The remaining part is organized as follows. Our proposed weighted competing risks quantile regression models and its penalized methods are developed in Section 2, with asymptotic properties demonstrated in Section 3. Simulation study as well as the application to the BMT data are performed in Section 4 to illustrate the performance of proposed methods. The technical proofs are provided in the supplementary material.

2. Models. We take the formulation of competing risks quantile regression in [11]. In the setting of competing risk models, assume there exist K causes of failure, denoted by an observable indicator $\epsilon \in \{1, ..., K\}$, the same denotation as [3]. Without loss of generality, we can set $K = 2$. Let T and C denote the failure and censoring time respectively, and we observe $X = \min(T, C)$, and censoring or risk indicator $\delta = I(T \leq C)$, where $I(\cdot)$ is an indicator function. Denote $p \times 1$ bounded time-independent covariate vector as $\tilde{\boldsymbol{Z}}$ and $\boldsymbol{Z} = (1, \tilde{\boldsymbol{Z}}^\top)^\top$. Assume that $\{X_i, \delta_i\epsilon_i, \boldsymbol{Z}_i\}, i = 1, ..., n$ are independent and identically distributed observed samples.

[3] modeled the CIF for failure from cause 1 conditionally on the covariates, $F_1(t|\boldsymbol{Z}) = P(T \leq t, \epsilon = 1|\boldsymbol{Z})$. They proposed the PSH model based on the formula of subdistribution hazard which is defined as

$$\begin{aligned}\lambda_1(t|\boldsymbol{Z}) &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} P\left\{t < T \leq t + \Delta t, \epsilon = 1 \middle| (T \geq t) \cup (T \leq t \cap \epsilon \neq 1), \boldsymbol{Z}\right\} \\ &= \{dF_1(t|\boldsymbol{Z})/dt\}/\{1 - F_1(t|\boldsymbol{Z})\}\end{aligned}$$

in [5]. Analogue to the definition of quantile, define the conditional quantile as $Q_k(\tau|\boldsymbol{Z}) = \inf\{t : F_k(t|\boldsymbol{Z}) \geq \tau\}, k = 1, ..., K$, where $F_k(t|\boldsymbol{Z}) = P\{T \leq t, \epsilon = k\}$ is the CIF for cause k; more details can refer to [11]. For $\tau \in [\tau_L, \tau_U]$, consider $Q_1(\tau|\boldsymbol{Z})$ to be modeled as

$$Q_1(\tau|\boldsymbol{Z}) = g\{\boldsymbol{Z}^\top \boldsymbol{\beta}_0(\tau)\}, \tag{1}$$

where $\boldsymbol{\beta}_0(\tau)$ is a $(p+1) \times 1$ coefficient vector, $g(\cdot)$ is a known monotone increasing and continuously differential bounded link function, $0 < \tau_L \leq \tau_U < 1$. With the statement in Fine & Gray , if we denote $T_1^* = I(\epsilon = 1) \times T + \{1 - I(\epsilon = 1)\} \times \infty$, then T_1^* has distribution function equal to $F_1(t|\boldsymbol{Z})$ when $t < \infty$, and a point mass $P(T^* = \infty|\boldsymbol{Z}) = P(T < \infty, \epsilon \neq 1) = 1 - F_1(\infty|\boldsymbol{Z})$ at $t = \infty$. Then at $\tau < F_1(\infty|\boldsymbol{Z})$, the $\tau-$quantile of T_1^* equals to $F_1^{-1}(\tau|\boldsymbol{Z}) = Q_1(\tau|\boldsymbol{Z}) = g\{\boldsymbol{Z}^\top \boldsymbol{\beta}_0(\tau)\}$ under the formulation of (1).

Remark: According to the formulation of T_1^*, we can see that when $\tau \geq F_1(\infty; \boldsymbol{Z})$, the $\tau-$quantile of T_1^* will become ∞, which is obvious when reviewing the definition that $F_1(t; \boldsymbol{Z}) = P(T \leq t, \epsilon = 1|\boldsymbol{Z}) \leq P(T \leq \infty, \epsilon = 1|\boldsymbol{Z}) = F_1(\infty; \boldsymbol{Z})$ and the fact that $g(\cdot)$ is monotone increasing. This fact provides an thought about the choice of τ_U.

With reference to [9], for proper τ, $\boldsymbol{\beta}_0(\tau)$ is supposed to be the minimizer of the following expected loss function with respect to $\boldsymbol{\beta}(\tau)$,

$$\boldsymbol{\beta}_0(\tau) = \arg\min_{\boldsymbol{\beta}(\tau)} E\rho_\tau(g^{-1}(T_1^*) - \boldsymbol{Z}^\top \boldsymbol{\beta}(\tau)), \tag{2}$$

where $\rho_\tau(u) = u\{\tau - I(u \leq 0)\}$ is called "check" function.

In sample scenario, we can obtain the estimator $\hat{\boldsymbol{\beta}}(\tau)$ of $\boldsymbol{\beta}_0(\tau)$ via minimizing the following objective function:

$$\min_{\boldsymbol{\beta}(\tau)} \sum_{i=1}^{n} \rho_\tau \left(g^{-1}(T_{1,i}^*) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}(\tau) \right). \tag{3}$$

2.1. *Weighted Competing Risks Quantile Regression.* Similar to [11], our paper first considers the case in which there is no missing data (i.e., there is no censoring). As a result, $X = T$ and $\delta = 1, \delta\epsilon = \epsilon$. As aforementioned, we can estimate $\boldsymbol{\beta}_0(\tau)$ via the minimization problem (3). Since $T_{1,i}^*$ is not observed, we modify (3) to

$$\sum_{i=1}^{n} I(\epsilon_i = 1)\rho_\tau(g^{-1}(X_i) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}(\tau)) + I(\epsilon_i \neq 1)\rho_\tau(g^{-1}(X^\infty) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}(\tau)), \tag{4}$$

where X^∞ is any value sufficiently large to exceed all $Z_i^\top \boldsymbol{\beta}(\tau)$. Then it's not difficult to derive the negative subgradient of (4) with respect to $\boldsymbol{\beta}(\tau)$.

For censoring case, we aim to construct such a weighted quantile objective function to estimate $\boldsymbol{\beta}_0(\tau)$ as follows,

$$\begin{aligned} Q(\boldsymbol{\beta}(\tau), w_0) = \sum_{i=1}^{n} \Big\{ & w_{0i}\rho_\tau(g^{-1}(X_i) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}(\tau)) \\ & + (1 - w_{0i})\rho_\tau(g^{-1}(X^{+\infty}) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}(\tau)) \Big\}. \end{aligned} \tag{5}$$

The weight function is re-constructed based on competing risks analogy to [18], as follows,

$$w_{0i} = \begin{cases} 1, & \delta_i\epsilon_i = 1, \\ 0, & \delta_i\epsilon_i \neq 1, F_1(C_i|\boldsymbol{Z}_i) > \tau, \\ \frac{\tau - F_1(C_i|\boldsymbol{Z}_i)}{1 - F_1(C_i|\boldsymbol{Z}_i)}, & \delta_i\epsilon_i \neq 1, F_1(C_i|\boldsymbol{Z}_i) \leq \tau. \end{cases} \tag{6}$$

The understanding of the weighting scheme can be referred to Remark 1 of [17] with slight modifications. In our case of competing risks quantile regression , each point contributes to the subgradient condition only via the sign of $g^{-1}(T_{1,i}^*) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau)$. For data with $\delta_i\epsilon_i = 1$, we know $X_i = T_i \leq C_i, \epsilon_i = 1, i.e., X_i = T_{1,i}^*$, and $I(g^{-1}(T_{1,i}^*) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau) < 0)$ can be observed, thus we assign a weight of 1 for this case. For data with $\delta_i\epsilon_i \neq 1$ and $F_1(C_i|\boldsymbol{Z}_i) > \tau$, then $T_i > C_i, F_1(C_i|\boldsymbol{Z}_i) > \tau$ or $T_i \leq C_i, \epsilon_i = 2, F_1(C_i|\boldsymbol{Z}_i) > \tau$; in the first scenario, $T_{1,i}^* \geq T_i \geq X_i = C_i > g(\boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau)), I(g^{-1}(T_{1,i}^*) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau) < 0) = 0$; in the second scenario, $T_i \leq C_i, \epsilon = 2, I(g^{-1}(T_{1,i}^*) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau) < 0) = 0$, where we assign a weight of 0. The ambiguous situation is $\delta_i\epsilon_i \neq 1$ and $F_1(C_i|\boldsymbol{Z}_i) < \tau$,i.e., $C_i \leq F_1^{-1}(\tau|\boldsymbol{Z}_i) = g(\boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau))$. If $\delta_i = 1, \epsilon_i = 2, X_i = T_i < C_i < g(\boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau))$, or $I\{g^{-1}(X_i) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau) < 0\} = 1$; if $\delta_i = 0, X_i = C_i < \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau)$, i.e. $I\{g^{-1}(X_i) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau) < 0\} = 1$. But the $I(T_{1,i}^* - g(\boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau)) < 0)$ cannot be observed.

Thus, we assign the weight $w_i(F_0) = \frac{\tau - F_1(C_i|\boldsymbol{Z}_i)}{1-F_1(C_i|\boldsymbol{Z}_i)}$ for this case, where given $(\boldsymbol{Z}_i, C_i)$,

$$
\begin{aligned}
&E\left\{I(g^{-1}(T^*_{1,i}) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau) < 0)|\delta_i\epsilon_i \neq 1, \boldsymbol{Z}_i\right\}\\
&\quad = \frac{P\left\{\epsilon_i = 1, T_i < g(\boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau))|\boldsymbol{Z}_i\right\} - P\left\{\epsilon_i = 1, T_i < C_i\right\}}{1 - P\left(T_i \le C_i, \epsilon_i = 1|\boldsymbol{Z}_i\right)}\\
&\quad = \frac{\tau - F_1(C_i|\boldsymbol{Z}_i)}{1 - F_1(C_i|\boldsymbol{Z}_i)}. \qquad (7)
\end{aligned}
$$

Actually, we can show that a subgradient of the weighted quantile objective function (5) with respect to $\boldsymbol{\beta}(\tau)$

$$
\mathbf{M}_n(\boldsymbol{\beta}(\tau), w_0) = \sum_{i=1}^n \boldsymbol{Z}_i \left\{\tau - w_{0i} I(g^{-1}(X_i) < \boldsymbol{Z}_i^\top \boldsymbol{\beta}(\tau))\right\} \tag{8}
$$

is an unbiased estimating function of $\boldsymbol{\beta}_0(\tau)$.

$$
\begin{aligned}
&E\left[w_{0i} I\{g^{-1}(X_i) < \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau)\}|\boldsymbol{Z}_i\right]\\
&= E\left(I\{\delta_i\epsilon_i = 1\} w_{0i} I\{g^{-1}(X_i) < \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau)\}|\boldsymbol{Z}_i\right)\\
&\quad + E\left(I\{\delta_i\epsilon_i \neq 1, F_1(C_i) > \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau)\} w_{0i} I\{g^{-1}(X_i) < \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau)\}|\boldsymbol{Z}_i\right)\\
&\quad + E\left(I\{\delta_i\epsilon_i \neq 1, F_1(C_i) \le \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau)\} w_{0i} I\{g^{-1}(X_i) < \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau)\}|\boldsymbol{Z}_i\right)\\
&= P\left(\epsilon_i = 1, g^{-1}(T_i) < \boldsymbol{Z}_i^\top \boldsymbol{\beta}_0(\tau)|\boldsymbol{Z}_i\right) = \tau.
\end{aligned}
$$

Although unbiasedness of (8) is proved with $F_1(C_i|\boldsymbol{Z}_i)$ in w_{0i}, the underlying distribution $F_1(t|\boldsymbol{Z})$ or w_{0i} is unknown in practice. Here we use the IPCW ([13]) estimator proposed by [11] to estimate $F_1(t|\boldsymbol{Z})$,

$$
\hat{F}_1(x|\boldsymbol{Z}) = \frac{1}{n}\sum_{i=1}^n \left\{\frac{I(X_i \le x, \delta_i\epsilon_i = 1)}{1 - \hat{G}(X_i)}\right\}, \tag{9}
$$

where $1 - \hat{G}(\cdot t)$ is the the Kaplan-Meier estimator in [7] of the survival function of C .

By plugging (9) in the expression of w_{0i}, we can get the estimated weights $w_i(\hat{F}_1)$,

$$
w_i(\hat{F}_1) = \begin{cases} 1, & \delta_i\epsilon_i = 1,\\ 0, & \delta_i\epsilon_i \neq 1, \hat{F}_1(C_i|\boldsymbol{Z}_i) > \tau,\\ \frac{\tau - \hat{F}_1(C_i|\boldsymbol{Z}_i)}{1 - \hat{F}_1(C_i|\boldsymbol{Z}_i)}, & \delta_i\epsilon_i \neq 1, \hat{F}_1(C_i|\boldsymbol{Z}_i) \le \tau, \end{cases} \tag{10}
$$

where $\hat{F}_1$ is as (9). Then we obtain the weighted censoring quantile regression estimator $\hat{\boldsymbol{\beta}}(\tau)$ by minimizing the weighted objective function,

$$
Q(\boldsymbol{\beta}(\tau), w_i(\hat{F}_1)) = \sum_{i=1}^n \left\{w_i(\hat{F}_1)\rho_\tau(g^{-1}(X_i) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}(\tau)) + (1 - w_i(\hat{F}_1))\rho_\tau(g^{-1}(X^{+\infty}) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}(\tau))\right\}. \tag{11}
$$

2.2. *Variable selection procedure.* To select important variables, a penalty function is added to the weighted objective function (11), to obtain the penalized estimator $\tilde{\boldsymbol{\beta}}(\tau)$:

$$
Q_p(\boldsymbol{\beta}(\tau), w_i(\hat{F}_1)) = \sum_{i=1}^n \left\{w_i(\hat{F}_1)\rho_\tau(g^{-1}(X_i) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}(\tau)) + (1 - w_i(\hat{F}_1))\rho_\tau(g^{-1}(X^{+\infty}) - \boldsymbol{Z}_i^\top \boldsymbol{\beta}(\tau))\right\}
$$

$$+\sum_{j=1}^{p} p_\lambda(|\beta_j(\tau)|), \tag{12}$$

where $p_\lambda(\cdot)$ can be LASSO, Adaptive LASSO and so on.

For LASSO and ALASSO penalty, we can easily write $p_\lambda(|\beta_j|) = \lambda_n|\hat{\beta}_j|^{-\gamma}$, where $|\hat{\beta}_j|$ is the j-th element of the initial consistent unpenalized estimator. We choose $\gamma = 0$ for LASSO and $\gamma = 1$ for ALASSO. The minimization of (12) and (11) can be directly solved with R package quantreg without linear programming, leading our proposed methods to conveniently applicable tools.

3. Theoretical Property. To establish the asymptotic results in this paper, we require the following assumptions:

A1 The covariates $\boldsymbol{Z}$ is bounded in probability. There exists a constant $K_{\boldsymbol{z}}$ such that $E\|\boldsymbol{Z}\|^3 \le K_{\boldsymbol{z}}$. and $E(\boldsymbol{Z}\boldsymbol{Z}^\top)$ is a positive definite $p \times p$ matrix.
A2 The functions $F_1(t|\boldsymbol{Z})$ and $G(t|)$ have first derivatives with respect to t, denoted as $f_1(t|\boldsymbol{Z})$ and $g_0(t)$, which are uniformly bounded away from infinity. Additionally, $F_1(t|\boldsymbol{Z})$ and $G(t)$ have bounded (uniformly in t) second-order partial derivatives with respect to $\boldsymbol{Z}$.
A3 For $\boldsymbol{\beta}$ in the neighbourhood of $\boldsymbol{\beta}_0(\tau)$, $E(\boldsymbol{Z}\boldsymbol{Z}^\top g'(\boldsymbol{Z}^\top\boldsymbol{\beta}) f_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z})\{1 - G(g(\boldsymbol{Z}^\top\boldsymbol{\beta}))\})$ and $E(\boldsymbol{Z}\boldsymbol{Z}^\top g'(\boldsymbol{Z}^\top\boldsymbol{\beta}) g_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta})))$ are positive definite.

Assumption A1 states some tail and moment conditions on the covariate $\boldsymbol{Z}$, which are standard for the quantile regression. Assumption A2 is needed for the local Kaplan-Meier estimator. It allows us to obtain the local expansions of $F_1(t|\boldsymbol{z})$ and $G(t)$ in the neighbourhood of $\boldsymbol{Z}^\top\boldsymbol{\beta}_0(\tau)$, to obtain the uniform consistency and the linear representation of $\hat{F}_1(t|\boldsymbol{Z})$. Assumption A3 ensures that the expectation of the estimating function $E\{\mathbf{M}_n(\boldsymbol{\beta}, F_1)\}$ has a unique zero at $\boldsymbol{\beta}_0(\tau)$ and it is needed to establish the asymptotic distribution of $\hat{\boldsymbol{\beta}}(\tau)$.

C1 There exists $\nu > 0$ such that $P(C = \nu) > 0$ and $P(C > \nu)$=0.
C2 $\boldsymbol{\beta}_0(\tau)$ is Lipschitz continuous for $\tau \in [\tau_L, \tau_U]$.
C3 $P(\epsilon = 1|\boldsymbol{Z}) < 1$ a.s.

Assumption C1-C2 are regularity conditions for competing risks quantile regressions. Assumption C3 is easily satisfied for the situation of competing risks; otherwise, it will turns out to be a standard Cox model.

THEOREM 3.1. *Assume that triples $\{\boldsymbol{Z}_i, X_i, \delta_i\epsilon_i\}, i = 1, \cdots, n$ constitute an iid multivariate random sample, and that the censoring variable C_i is independent of T_i conditionally on the covariate $\boldsymbol{Z}_i$. Under model (1) and assumptions A1-A3, C1-C3,*

$$\hat{\boldsymbol{\beta}}(\tau) \to \boldsymbol{\beta}_0(\tau) \tag{13}$$

in probability as $n \to \infty$.

THEOREM 3.2. *Under the assumptions of Theorem 1 and $r < 1/4$, we have*

$$n^{1/2}(\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)) \xrightarrow{D} \mathrm{N}(0, \boldsymbol{\Gamma^{-1}}\mathbf{V}\boldsymbol{\Gamma^{-1}}), \tag{14}$$

where $\boldsymbol{\Gamma^{-1}} = E[\boldsymbol{Z}\boldsymbol{Z}^\top g'(\boldsymbol{Z}^\top\boldsymbol{\beta}_0(\tau))\{1 - G(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0(\tau)))\} f_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0(\tau))|\boldsymbol{Z})]$, and

$$\mathbf{V} = \mathrm{Cov}(\boldsymbol{m}_i(\boldsymbol{\beta}_0, F_1) + (1-\tau)\boldsymbol{\phi}_i), \tag{15}$$

with $\boldsymbol{m}_i(\boldsymbol{\beta}_0, F_1) = \boldsymbol{Z}_i\{\tau - w_i(F_1) I(X_i < g(\boldsymbol{Z}_i^\top\boldsymbol{\beta}_0(\tau)))\}$, $\boldsymbol{\phi}_i$ defined in equation (24) in the Appendix 5.

Theorem 3.1-3.2 established the consistency and asymptotic normality of the unpenalized estimator $\hat{\boldsymbol{\beta}}(\tau)$. We then establish the property of consistency in variable selection of the proposed penalized estimator $\tilde{\boldsymbol{\beta}}(\tau)$. Let $\mathcal{A}(\tau) : \{j : \boldsymbol{\beta}_{0j} \neq 0\}$ and $\mathcal{A}^c(\tau) = \{j : \boldsymbol{\beta}_{0j}(\tau) = 0\}$

THEOREM 3.3. *If A1-A3,C1-C3 hold, and if $n^{-1/2}\lambda_n \to 0$ and $n^{\gamma/2-1}\lambda_n \to \infty$, then*

$$P\left(\{j : \tilde{\beta}_j(\tau) \neq 0\} = \mathcal{A}(\tau)\right) \to 1 \quad \text{as} \quad n \to \infty.$$

Theorem 3.3 states that the proposed procedure is able to select the correct model with probability approaching one. By the remark of Theorem 2 of [18], the oracle properties are satisfied by the proposed estimators. The technical proof is also provided in Appendix.

4. Numerical Studies.

4.1. *Monte Carlo Simulation.* We conduct Monte Carlo simulations to evaluate the performance of the proposed methods. Consider the data-generating ways as [11] with larger dimension of convariates. Specifically, we set the number of predictors to be 13, the sample size to be 200 and 400. For the structure of covariance matrix for covariates, we consider two kinds of setup: $\Sigma_{1,ij} = \rho$ and $\Sigma_{2,ij} = \rho^{|i-j|}$, where $\rho = 0, 0.25, 0.5, 0.75$.

We generate (T, ϵ) satisfying $P(\epsilon = 1|\boldsymbol{Z}) = p_0 I(Z_2 = 0) + p_1 I(Z_2 = 1)$, $P(T \le t|\epsilon = 1, \boldsymbol{Z}) = \Phi(\log t - \boldsymbol{\gamma}_0^\top \boldsymbol{Z})$ and $P(T \le t|\epsilon = 2, \boldsymbol{Z}) = \Phi(\log t - \boldsymbol{\alpha}_0^\top \boldsymbol{Z})$, where $\Phi(\cdot)$ denotes the standard normal distribution function, $p_0 = 0.8, p_1 = 0.6$. Set $\gamma_0 = (-2, -2.5, 0.5, 0, \cdots, 0)$, while $\alpha_0 = -\gamma_0$. Then

$$\log Q_1(\tau|\boldsymbol{Z}) = \Phi^{-1}\left(\frac{\tau}{p_0}\right) + \gamma_0^{(1)} Z^{(1)} + \left\{\gamma_0^{(2)} + \Phi^{-1}\left(\frac{\tau}{p_1}\right) - \Phi^{-1}\left(\frac{\tau}{p_0}\right)\right\} Z^{(2)} + \gamma^{(3)} Z^{(3)},$$

thus the true number of non-zero coefficient is 4 for $\tau \neq 0.4$, and 3 for $\tau = 0.4$ due to $\Phi^{-1}\left(\frac{0.4}{p_0}\right) = 0$.

Besides, we also simulate the heavy-tailed distributions $t(3)$ instead of Gaussian distribution for $P(T \le t|\epsilon = 1, \boldsymbol{Z})$. We generate the covariate vector $\boldsymbol{Z} = (Z_1, Z_2, \cdots, Z_{13})^\top$ as follows: $Z_1 \sim$ Unif(0,1) and $Z_2 \sim$ Bernoulli(0.5), $Z_i \sim N(0, \Sigma), i = 3, \cdots, 13$. Also we consider a different choice for p_0 and p_1 as 0.6 and 0.45 respectively, in order to test the performance under a different probability of $P(\epsilon_i = 1)$. For each scenario, the simulation is repeated 500 times.

We use the following criteria to evaluate the performances: the average number of relevant variables correctly selected (TP), the number of irrelevant variables incorrectly selected (FP), the absolute error $P_1 = \sum_{j=1}^{p} |\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)|$ and the squared error $P_2 = \sum_{j=1}^{p} \|\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau))\|^2$.

We compare our proposed weighted estimators with the estimated estimator of competing risks quantile regression model proposed in Li et al.(2019), denoted as WCQR and CQR respectively, implying weighting method or not. And in simulation tables, we use cqr.l and cqr.a to represent CQR estimators with LASSO and ALASSO penalty respectively. Similarly, our estimators, denoted as wcqr$i$.l and wcqr$i$.a, $i = 0, 1$ stands for administrative censoring where C are known and randomly right censoring cases where X in place of C respectively. wcqr2 and wcqr3 use a different weight:

$$w_i(F_1) = \begin{cases} 1 & \delta_i \epsilon_i = 1 \\ \frac{\tau - \hat{F}_1(C_i)}{1 - \hat{F}_1(C_i) - \hat{F}_2(C_i)} & \delta_i = 0, \hat{F}_1(C_i) < \tau \\ 0 & \text{otherwise.} \end{cases}$$

TABLE 1
Bias and Empirical Coverage, $n = 300, \rho = 0, p_0 = 0.8, p_1 = 0.6,$

τ	Method	Bias				EmpCoverage			
		β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
0.1	cqr	-0.012	-0.003	0.003	0.003	0.952	0.953	0.953	0.958
	wcqr1	-0.027	-0.026	0.031	0.014	0.950	0.952	0.950	0.949
	wcqr2	-0.027	-0.032	0.036	0.018	0.948	0.952	0.946	0.950
	wcqr3	-0.033	-0.038	0.041	0.020	0.948	0.949	0.944	0.950
0.2	cqr	0.000	-0.015	-0.007	0.008	0.946	0.950	0.943	0.952
	wcqr1	-0.024	-0.061	0.042	0.032	0.948	0.944	0.949	0.943
	wcqr2	-0.019	-0.081	0.055	0.040	0.949	0.940	0.952	0.940
	wcqr3	-0.044	-0.097	0.074	0.048	0.946	0.941	0.943	0.939
0.3	cqr	-0.010	-0.023	0.008	0.014	0.947	0.949	0.956	0.953
	wcqr1	-0.042	-0.108	0.095	0.055	0.938	0.952	0.944	0.943
	wcqr2	-0.027	-0.131	0.104	0.065	0.939	0.949	0.934	0.941
	wcqr3	-0.094	-0.207	0.170	0.105	0.933	0.941	0.910	0.929
0.4	cqr	0.219	-1.244	-0.069	0.316	0.999	0.999	0.999	0.999
	wcqr1	-0.065	-0.203	0.168	0.121	0.947	0.943	0.931	0.937
	wcqr2	0.007	-0.153	0.111	0.095	0.949	0.939	0.941	0.922
	wcqr3	-0.326	-0.678	0.493	0.347	0.959	0.930	0.950	0.920
0.5	cqr	-5.878	-25.526	-12.526	10.139	0.975	0.957	0.967	0.953
	wcqr1	-0.313	-0.632	0.594	0.280	0.965	0.943	0.963	0.914
	wcqr2	0.144	0.034	0.041	-0.007	0.913	0.955	0.952	0.953
	wcqr3	-2.007	-2.857	2.081	1.140	0.879	0.830	0.881	0.693

Since the weight above involves the estimation of $F_2(t) = P(T \leq t, \epsilon = 2|\boldsymbol{Z})$, which probably is complicated in practical circumstances, we only use it for comparison in simulations. Here in wcqr2, we apply similar estimating method of F_1 to F_2, whether in wcqr3, we use the Kaplan Meier estimator of $F(t) = P(T \leq t|\boldsymbol{Z})$ and $F_2 = F - F_1$, displaying almost the same with wcqr2. Though our theoretical results don't base on these two estimators of w_i, most simulation results shows that wcqr0 and wcqr1 are considerably close, since the weight is only different at $\delta = 1$, suggesting good estimates in large censoring rate. Research about massive competing risks data with enormous censored observations will appear in our future work.

Before the variable selection, we also conduct the simulation for unpenalized estimators. In this case, we use $\gamma_0 = (1, -1.5, -0.5)$, $p_0 = 0.8, p_1 = 0.6$ and $\rho = 0$. We repeat 1000 times, and compare the empirical bias (EmpBias) and average coverage probabilities based on 95% confidence intervals computed with empirical variance. The results are summarized in Table 1, where in lower quantiles, the cqr method shows extremely excellence, while in high quantiles, it displays sometime instability. For weighted methods, though inferior to cqr in lower quantiles, these methods behave still well in most simulations, especially wcqr1 and wcqr2. The average coverage probabilities display similar patterns, cqr behaves well until $\tau < 0.4$. In relatively high quantile such as $\tau = 0.5$, wcqr2 behaves the best for most coefficients.

Table 2 represents the TP and FP evaluated for $n = 200$ and all τ and ρ values for Σ_1 . We can find that the TPs are quite close to true numbers of relevant variables, as well as the case of $\tau = 0.4$ and our weighted estimators behave inferior to CQR estimators in TP but slightly superior to it in FP, though both behave well. In high quantile $\tau = 0.5$, the behaviors of both estimators are inferior than lower quantile such as $\tau = 0.1$. It may because 0.5 is close to the probability $P(\epsilon = 1|\boldsymbol{Z})$ for some $\boldsymbol{Z}$, which induces larger biases. It's easy to understand that TP decreases when ρ increases, since when the correlation of covariates increases, it's more difficult for identification. But even when $\rho = 0.75$, the simulation behaviors are still quite good.

TABLE 2
TP, $n = 200, p_0 = 0.8, p_1 = 0.6$,

ρ	τ		cqr.l	cqr.a	wcqr0.l	wcqr0.a	wcqr1.l	wcqr1.a	wcqr2.l	wcqr2.a	wcqr3.l	wcqr3.a
0	0.10	TP	3.82	3.90	3.83	3.91	3.83	3.91	3.83	3.90	3.83	3.90
		FP	4.82	0.60	4.82	0.65	4.82	0.65	4.79	0.63	4.79	0.63
	0.20	TP	3.84	3.93	3.83	3.91	3.83	3.91	3.82	3.92	3.82	3.92
		FP	5.18	0.58	4.99	0.52	4.99	0.52	5.07	0.53	5.07	0.53
	0.30	TP	3.90	3.95	3.86	3.95	3.86	3.95	3.88	3.94	3.88	3.94
		FP	5.21	0.54	4.64	0.50	4.64	0.50	4.71	0.52	4.71	0.52
	0.40	TP	2.90	2.95	2.88	2.95	2.87	2.94	2.88	2.95	2.88	2.95
		FP	6.43	1.88	5.90	1.76	5.86	1.73	6.19	1.70	6.19	1.70
	0.50	TP	3.31	3.60	3.20	3.49	3.18	3.48	3.33	3.58	3.33	3.58
		FP	4.91	1.37	4.18	1.10	4.14	1.03	4.46	1.01	4.46	1.01
0.25	0.10	TP	3.83	3.88	3.77	3.87	3.77	3.87	3.77	3.87	3.77	3.87
		FP	5.01	0.85	4.73	0.79	4.73	0.79	4.79	0.81	4.79	0.81
	0.20	TP	3.95	3.90	3.93	3.89	3.93	3.89	3.93	3.90	3.93	3.90
		FP	4.94	0.76	4.52	0.63	4.52	0.63	4.47	0.61	4.47	0.61
	0.30	TP	3.93	3.96	3.86	3.97	3.86	3.97	3.84	3.96	3.84	3.96
		FP	4.46	0.65	4.12	0.52	4.11	0.52	4.16	0.48	4.16	0.48
	0.40	TP	2.87	2.91	2.82	2.90	2.82	2.90	2.82	2.89	2.82	2.89
		FP	6.29	1.91	5.78	1.86	5.76	1.86	6.17	1.67	6.17	1.67
	0.50	TP	3.52	3.72	3.44	3.61	3.48	3.60	3.57	3.70	3.57	3.70
		FP	4.92	1.63	4.34	1.26	4.52	1.17	4.69	0.92	4.69	0.92
0.5	0.10	TP	3.83	3.89	3.83	3.90	3.83	3.90	3.83	3.90	3.83	3.90
		FP	4.54	0.83	4.49	0.76	4.49	0.76	4.40	0.76	4.40	0.76
	0.20	TP	3.91	3.90	3.88	3.90	3.88	3.90	3.89	3.90	3.89	3.90
		FP	3.81	0.60	3.43	0.58	3.43	0.58	3.47	0.54	3.47	0.54
	0.30	TP	3.95	3.92	3.92	3.92	3.92	3.92	3.93	3.93	3.93	3.93
		FP	3.84	0.43	3.47	0.39	3.45	0.40	3.33	0.38	3.33	0.38
	0.40	TP	2.96	2.94	2.90	2.92	2.89	2.92	2.94	2.90	2.94	2.90
		FP	5.41	1.84	4.92	1.64	4.93	1.65	5.08	1.58	5.08	1.58
	0.50	TP	3.45	3.58	3.46	3.43	3.43	3.44	3.49	3.63	3.49	3.63
		FP	4.36	1.74	3.93	1.24	3.84	1.29	4.19	1.36	4.19	1.36
0.75	0.10	TP	3.76	3.71	3.74	3.67	3.74	3.67	3.77	3.68	3.77	3.68
		FP	3.21	1.05	3.10	0.97	3.10	0.97	3.15	1.00	3.15	1.00
	0.20	TP	3.86	3.77	3.87	3.78	3.87	3.78	3.88	3.77	3.88	3.77
		FP	3.11	1.04	2.89	0.88	2.89	0.88	2.79	0.86	2.79	0.86
	0.30	TP	3.85	3.79	3.83	3.78	3.83	3.78	3.82	3.78	3.82	3.78
		FP	3.23	0.97	2.88	0.76	2.89	0.76	2.86	0.71	2.86	0.71
	0.40	TP	2.79	2.72	2.80	2.76	2.80	2.76	2.80	2.72	2.80	2.72
		FP	4.21	2.11	3.81	1.94	3.84	1.94	4.00	1.92	4.00	1.92
	0.50	TP	3.38	3.41	3.32	3.31	3.32	3.31	3.32	3.44	3.32	3.44
		FP	3.82	1.77	3.29	1.47	3.24	1.38	3.00	1.52	3.00	1.52

The simulation results of P1,P2 are placed in supplementary material, which shows that the ALASSO penalty significantly decreases the FP for both estimators, which suggests the strengthes of ALASSO. Our estimators behave fairly close to CQR estimator in most cases. Though TP of our estimators behave slightly inferior, the FP shows a relatively better performance. Not only WCQR shows smaller deviation about estimated coefficients, but also shows greatly stability, especially for the ALASSO penalty, in the case of higher quantiles $\tau = 0.5$. This shows the meaningful application of our estimators in high quantiles.

Table 3 represents the case of $n = 400$, where the performances of all criteria are greatly improved. Other different scenarios including Σ_2, which presents even a little better results than the case of Σ_1, a different pair of $(p_0, p_1) = (0.6, 0.45)$ and t(3) distribution in place of standard normal distribution, displaying that our estimators still behave very well for heavy-tail distributions are all provided in supplementary material.

TABLE 3
$n = 400, \rho = 0.5, p_0 = 0.8, p_1 = 0.6,$

Criteria	τ	cqr.l	cqr.a	wcqr0.l	wcqr0.a	wcqr1.l	wcqr1.a	wcqr2.l	wcqr2.a	wcqr3.l	wcqr3.a
TP	0.10	3.99	3.99	3.99	3.99	3.99	3.99	3.98	3.99	3.98	3.99
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.20	4.00	3.99	3.99	3.99	3.99	3.99	4.00	3.99	4.00	3.99
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.30	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.40	3.00	2.98	3.00	2.98	3.00	2.98	3.00	2.98	3.00	2.98
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.50	3.96	3.94	3.98	3.94	3.98	3.95	3.99	3.96	3.99	3.96
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FP	0.10	4.33	0.49	4.34	0.48	4.34	0.48	4.34	0.45	4.34	0.45
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.20	4.16	0.41	3.60	0.32	3.60	0.32	3.70	0.28	3.70	0.28
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.30	4.70	0.53	4.05	0.36	4.08	0.36	4.18	0.37	4.18	0.37
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.40	5.56	1.54	4.85	1.45	4.85	1.45	4.85	1.42	4.85	1.42
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.50	5.91	0.93	5.12	0.74	5.01	0.86	5.25	0.88	5.25	0.88
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
P_1	0.10	1.82	1.00	1.81	1.00	1.81	1.00	1.81	1.01	1.81	1.01
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.20	1.67	0.78	1.73	0.79	1.73	0.79	1.71	0.80	1.71	0.80
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.30	1.63	0.80	1.68	0.77	1.67	0.77	1.69	0.78	1.69	0.78
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.40	1.73	0.83	1.65	0.80	1.65	0.82	1.75	0.82	1.75	0.82
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.01
	0.50	2.15	1.21	2.04	1.17	2.06	1.18	2.18	1.27	2.18	1.27
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
P_2	0.10	1.03	0.41	1.02	0.42	1.02	0.42	1.03	0.44	1.03	0.44
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.01
	0.20	0.79	0.25	0.93	0.27	0.93	0.27	0.87	0.28	0.87	0.28
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.30	0.70	0.26	0.81	0.25	0.80	0.25	0.83	0.25	0.83	0.25
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.40	0.80	0.29	0.80	0.28	0.80	0.30	0.87	0.31	0.87	0.31
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.50	1.10	0.59	1.05	0.59	1.10	0.57	1.18	0.59	1.18	0.59
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

To conclude, WCQR estimators behaves comparably with CQR estimator, with slightly worse performance for TP but better for FP. Interestingly, the higher correlations and higher quantiles, the superior performance the WCQR estimators display, which extends our application to more complex data and higher quantiles.

4.2. *Real data analysis.* In this subsection, we use the BMT data set in in [14] for practical application. As the simulation illustrates, WCQR estimators display more stability to complexity of data and high quantiles than existing CQR estimator, which motivates us to conduct the data analysis with our methods.

In this dataset, a total of 177 patients received a stem cell transplant for acute leukemia. The failure event is relapse of the original disease (REL, 56 patients) and death from causes related to the transplant (transplant related mortality, TRM, 75 patients) is the competing

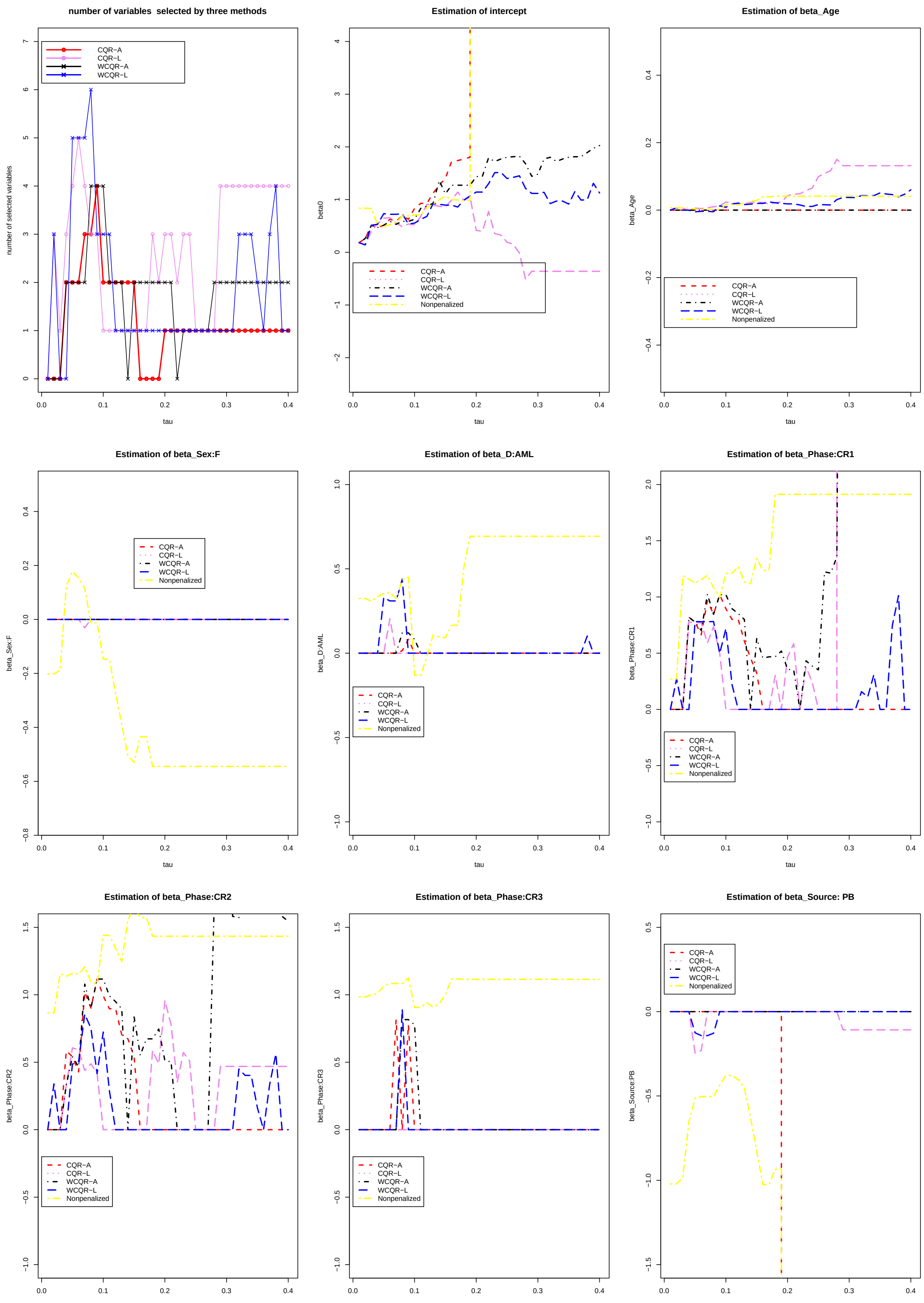

FIG 1. *Variable Selection and Estimation Results*

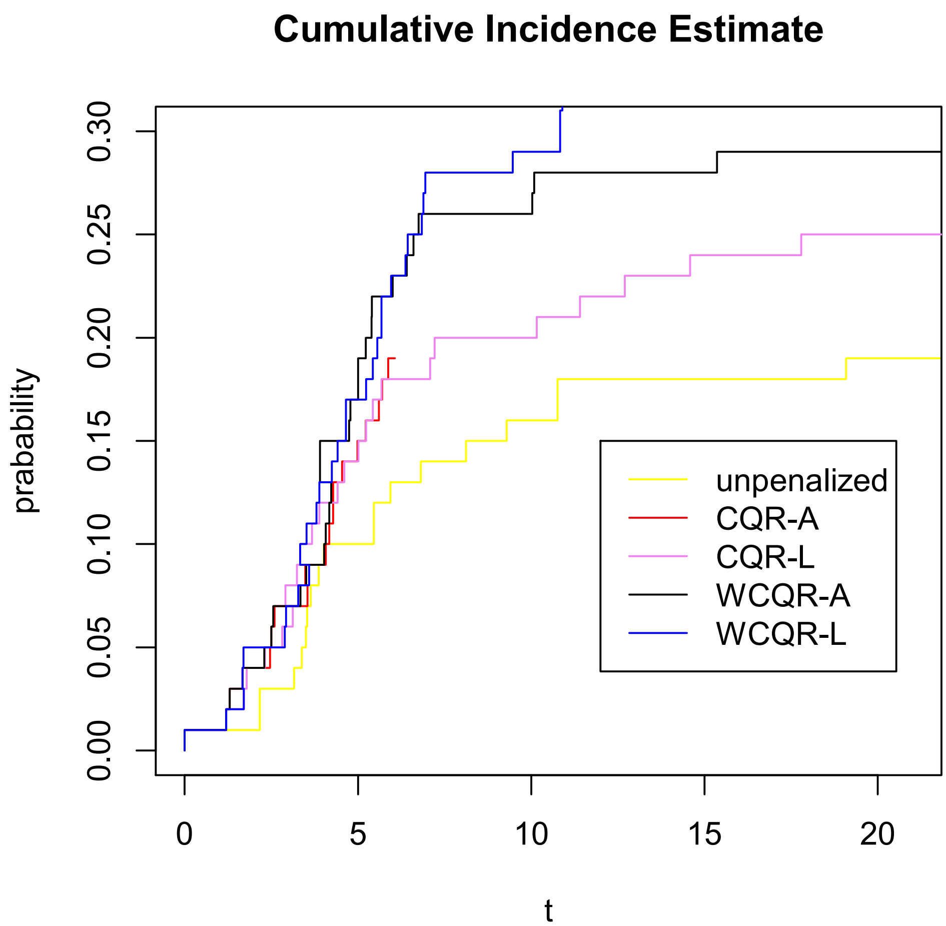

FIG 2. *Estimation of* F_1.

risk. 46 patients are censored, thus censoring rate is 26% . Covariates that affect REL and TRM includes Sex, Disease (lymphoblastic or myeloblastic leukemia), Phase at transplant (Relapse, CR1, CR2, CR3), Source of stem cells (bone marrow and peripheral blood, coded as BM+PB, or peripheral blood, coded as PB), and Age. The link function is assumed to be exponential.

Figure 1 reports the numbers of selected variables as well as coefficient estimates by our weighted estimators compared with penalized quantile estimating equations proposed by [10] and the penalty-free methods with τ ranging from 0 to 0.4. From the figure we can see mainly our estimators select similar numbers of variables to CQR estimators at lower quantiles, but in higher quantiles, the WQR estimators lie between the CQR-LASSO and CQR-ALASSO. For the intercept, in lower quantiles, five estimators appears conincidence with each other, though CQR-ALASSO estimator tends to be unstable while WCQR estimators shows stability here. For Age, all estimators behave to regard this variable as unimportant, except that the two LASSO estimators probably overestimate the importance. For Sex: F, almost all estimators shrink the corresponding coefficients to zero. The ALASSO estimators tend to treat D:AML as unimportant variable, except for quantile around 0.1. For Phase:CR1 and Phase:CR2, all estimators tend to select them in lower quantiles , but WCQR tends to select Phase:CR1 at higher quantiles bigger than 0.21 but neglect Phase:CR2 from 0.22 to 0.27. For Phase:CR3, all the estimators show analogue performances but for slightly shifts. For Source:PB, the WCQR estimators perform more stably than CQR for all quantiles.

To conclude, our WCQR estimators present stability and keep similar performances to the results of CQR estimators. More importantly, our weighted estimates provide a relatively general objective function for researchers to directly use R packages for application.

5. Conclusion. In this paper, we propose a locally weighted method for competing risks quantile regression model to transform the estimating equation to a common weighted objective function, and apply the LASSO and ALASSO penalization for variable selection. We establish the consistency and asymptotic normality for penalty-free estimators as well as the consistency of variable selection. Monte Carlo simulations are conducted for several scenarios, presenting the good variable selection performance and stability. Last, a real data set is utilized to illustrate the application of our methods, which is comparable with the other method.

APPENDIX

To simplify the presentation, we omit τ in such expressions as $\boldsymbol{\beta}(\tau)$.

Since the weights w_i depend on F_1^* , we take w_i as $w_i(F_1^*)$. Additionally, we define $\mathbf{M}_n(\boldsymbol{\beta}, F_1^*) = n^{-1}\sum_{i=1}^n \mathbf{m}_i(\boldsymbol{\beta}, F_1^*)$ as the subgradient of the weighted quantile objective function (11), where

$$\begin{aligned}
\mathbf{m}_i(\boldsymbol{\beta}, F_1^*) &= \boldsymbol{Z}_i\{\tau - w_i(F_1^*)I(g^{-1}(X_i) \le \boldsymbol{Z}_i^\top \boldsymbol{\beta})\} \\
&= \boldsymbol{Z}_i\Big(\tau - I\{\epsilon_i = 1, T_i \le C_i, C_i \le g(\boldsymbol{Z}_i^\top\boldsymbol{\beta})\} \\
&\quad - I\{\epsilon_i = 1, T_i \le C_i, g^{-1}(T_i) \le \boldsymbol{Z}_i^\top\boldsymbol{\beta}, C_i > g(\boldsymbol{Z}_i^\top\boldsymbol{\beta})\} \\
&\quad - \frac{\tau - F_1^*(C_i)}{1 - F_1^*(C_i)}\Big[I\{F_1^*(C_i) \le \tau, C_i \le g(\boldsymbol{Z}_i^\top\boldsymbol{\beta})\}(1 - I\{T_i \le C_i, \epsilon_i = 1\})\Big] \\
&\quad - \frac{\tau - F_1^*(C_i)}{1 - F_1^*(C_i)} I\{\epsilon_i = 2, F_1^*(C_i|\boldsymbol{Z}_i) \le \tau, T_i \le g(\boldsymbol{Z}_i^\top\boldsymbol{\beta}), C_i > g(\boldsymbol{Z}_i^\top\boldsymbol{\beta})\}\Big)
\end{aligned}$$

Let $\mathbf{M}(\boldsymbol{\beta}, F_1^*) = E\{\mathbf{m}_n(\boldsymbol{\beta}, F_1^*)\} = E\{\boldsymbol{Z}\left(\tau - H(g(\boldsymbol{Z}^\top\boldsymbol{\beta})) - R(\boldsymbol{\beta}, F_1^*) - J(\boldsymbol{\beta}, F_1^*)\right)\}$, where

$$H(t|\boldsymbol{Z}) = \int_{-\infty}^{t} F_1(u)g_0(u)\mathrm{d}u + (1 - G(t))F_1(t|\boldsymbol{Z}),$$

$$\begin{aligned}
R(\boldsymbol{\beta}, F_1^*) &= E_{C|\boldsymbol{Z}}\frac{\tau - F_1^*(C)}{1 - F_1^*(C)} I\{F_1^*(C) \le \tau, C \le g(\boldsymbol{Z}^\top\boldsymbol{\beta})\}(1 - I\{T \le C, \epsilon = 1\}) \\
&= \int_0^{g(\boldsymbol{Z}^\top\boldsymbol{\beta})} g_0(u)I\{F_1^*(u) \le \tau\}(1 - F_1(u|\boldsymbol{Z}))\frac{\tau - F_1^*(u)}{1 - F_1^*(u)}\mathrm{d}u,
\end{aligned}$$

$$\begin{aligned}
J(\boldsymbol{\beta}, F_1^*) &= E_{C|\boldsymbol{Z}} I\{\epsilon = 2, F_1^*(C) \le \tau, T \le g(\boldsymbol{Z}^\top\boldsymbol{\beta}), C > g(\boldsymbol{Z}^\top\boldsymbol{\beta})\}\frac{\tau - F_1^*(C)}{1 - F_1^*(C)} \\
&= (F_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z}) - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z}))\int_{g(\boldsymbol{Z}^\top\boldsymbol{\beta})}^{\infty} I\{F_1^*(u) \le \tau\}\frac{\tau - F_1^*(u)}{1 - F_1^*(u)} g_0(u)\mathrm{d}u
\end{aligned}$$

where $g_0(u)$ is the density of censoring variable C conditionally on $\boldsymbol{Z}$, and $F_0(t|\boldsymbol{Z}) = P(T \le t|\boldsymbol{Z})$. It's noteworthy that $J(\boldsymbol{\beta}_0, F_1) \equiv 0$, and it's easy to derive that $\mathrm{M}(\boldsymbol{\beta}_0, F_1) \equiv 0$.

LEMMA A.1. *Assume assumptions A1-A3,C1-C3 hold. Then*

$$\|\hat{F}_1 - F_1\|_{\mathcal{H}} \doteq \sup_t \sup_z |\hat{F}_1(t|\boldsymbol{z}) - F_1(t|\boldsymbol{z})| = o_p(n^{-1/2+r}) \tag{16}$$

for every $r > 0$.

PROOF OF LEMMA A.1. By condition C1 and A1, [11] has developed that for every $r > 0$, $\sup_{t<\nu} |\hat{G}(t) - G(t)| = o(n^{-1/2+r})$, a.s. This coupled with C2, implies that

$$\sup_x \left\| n^{-1} \sum_{i=1}^n \left(\frac{I\{X_i \geq x\} I\{\delta_i \epsilon_i = 1\}}{1 - \hat{G}(X_i)} \right) \right.$$

$$\left. -n^{-1} \sum_{i=1}^n \left(\frac{I\{X_i \leq g(\boldsymbol{Z}_i^\top \boldsymbol{b})\} I(\delta_i \epsilon_i = 1)}{1 - G(X_i)} \right) \right\| = o(n^{-1/2+r}), a.s., \tag{17}$$

Simultaneously, for $t < \nu$, $1 - G(t)$ is uniformly bounded away from 0 ,thus by Chebyshev's inequality,

$$P \left\{ \left| n^{-1} \sum_{i=1}^n \left(\frac{I\{X_i \leq x\} I(\delta_i \epsilon_i = 1)}{1 - G(X_i)} \right) \right. \right.$$

$$\left. \left. -n^{-1} \sum_{i=1}^n E\left(\frac{I\{X_i \leq x\} I(\delta_i \epsilon_i = 1)}{1 - G(X_i)} \right) \right| \geq \varepsilon | \boldsymbol{Z}_i \right\}$$

$$\leq \frac{n^{-1} \mathrm{Var}(I\{X_i \leq x\} I(\delta_i \epsilon_i = 1))}{\varepsilon^2}$$

and this inequality holds for any x and $\boldsymbol{Z}$, that is, for every $r > 0$,

$$\sup_{x,z} \left\| n^{-1} \sum_{i=1}^n \left(\frac{I\{X_i \leq x\} I(\delta_i \epsilon_i = 1)}{1 - G(X_i)} \right) - F_1(x|\boldsymbol{Z}_i) \right\| = o_p(n^{-1/2+r}). \tag{18}$$

Combining Eq.(17) and (18), we have

$$\sup_{x,z} \|\hat{F}_1(x|\boldsymbol{Z}) - F_1(x|\boldsymbol{Z})\| = o_p(n^{-1/2+r})$$

holds uniformly for $\boldsymbol{Z}$, that is,

$$\|\hat{F}_1 - F_1\|_{\mathcal{H}} \doteq \sup_t \sup_z |\hat{F}_1(t|\boldsymbol{z}) - F_1(t|\boldsymbol{z})| = o_p(n^{-1/2+r}).$$

□

Remark: Lemma A.1 directly guarantees the consistency of our weight estimation $w_i(\hat{F}_1)$ to $w_i(F_1)$, which is the w_{0i} in equation (6).

LEMMA A.2. *For all positive values* $\varepsilon_n = o(1)$, *we have*

$$\sup_{\|\boldsymbol{\beta} - \boldsymbol{\beta}_0\| \leq \varepsilon_n, \|F_1^* - F_1\| \leq \varepsilon_n} \|\boldsymbol{M}_n(\boldsymbol{\beta}, F_1^*) - \boldsymbol{M}(\boldsymbol{\beta}, F_1^*) - \boldsymbol{M}_n(\boldsymbol{\beta}_0, F_1)\| = o_p(n^{-1/2}) \tag{19}$$

PROOF OF LEMMA A.2. Let Z_{ij} and m_{ij} denote the $j-$th coordinates of $\boldsymbol{Z}_i$ and $\mathbf{m}_i$, respectively. For notational simplicity, in the following we omit the subscript i in various expressions such as $\boldsymbol{Z}_i, Z_{ij}, T_i, C_i$. Let $K_j, j = 1, \cdots, 5$ be some positive constants. Note that for $j = 1, \cdots, p$,

$$|m_j(\boldsymbol{\beta}, F_1^*) - m_j(\boldsymbol{\beta}', F_1^{*'})|^2 \leq B_1 + B_2 + B_3 + B_4,$$

where

$$
\begin{aligned}
B_1 &= Z_j^2 |I\{\epsilon = 1, T \le C, C \le g(\boldsymbol{Z}^\top \boldsymbol{\beta})\} - I\{\epsilon = 1, T \le C, C \le g(\boldsymbol{Z}^\top \boldsymbol{\beta}')\}| \\
B_2 &= Z_j^2 |I\{\epsilon = 1, T \le C, T \le g(\boldsymbol{Z}^\top \boldsymbol{\beta}), C > g(\boldsymbol{Z}^\top \boldsymbol{\beta})\} \\
&\quad - I\{\epsilon = 1, T \le C, T \le g(\boldsymbol{Z}^\top \boldsymbol{\beta}'), C > g(\boldsymbol{Z}^\top \boldsymbol{\beta}')\}| \\
B_3 &= Z_j^2 \left| \frac{\tau - F_1^*(C)}{1 - F_1^*(C)} \left[I\{F_1^*(C) \le \tau, C \le g(\boldsymbol{Z}^\top \boldsymbol{\beta})\}(1 - I\{T \le C, \epsilon = 1\}) \right] \right. \\
&\quad \left. - \frac{\tau - F_1^{*'}(C)}{1 - F_1^{*'}(C)} \left[I\{F_1^{*'}(C) \le \tau, C \le g(\boldsymbol{Z}^\top \boldsymbol{\beta}')\}(1 - I\{T \le C, \epsilon = 1\}) \right] \right| \\
B_4 &= Z_j^2 \left| \frac{\tau - F_1^*(C)}{1 - F_1^*(C)} I\{\epsilon = 2, F_1^*(C|\boldsymbol{Z}) \le \tau, T \le g(\boldsymbol{Z}^\top \boldsymbol{\beta}), C > g(\boldsymbol{Z}^\top \boldsymbol{\beta})\} \right. \\
&\quad \left. - \frac{\tau - F_1^{*'}(C)}{1 - F_1^{*'}(C)} I\{\epsilon = 2, F_1^{*'}(C) \le \tau, T \le g(\boldsymbol{Z}^\top \boldsymbol{\beta}'), C > g(\boldsymbol{Z}^\top \boldsymbol{\beta}')\} \right|
\end{aligned}
$$

It's easy to verify that

$$
\sup_{\beta' : \|\beta - \beta'\| \le \varepsilon_n} |I(g(\boldsymbol{Z}^\top \boldsymbol{\beta}) < C) - I(g(\boldsymbol{Z}^\top \boldsymbol{\beta}') < C)| \le \|\boldsymbol{Z}\| \{ I(g(\boldsymbol{Z}^\top \boldsymbol{\beta}) - \varepsilon_n < C) - I(g(\boldsymbol{Z}^\top \boldsymbol{\beta}) + \varepsilon_n < C) \}
$$

or multiplied by a constant, by assumption C3. Therefore, by assumptions A1 and A2,

$$
\begin{aligned}
&E \left(\sup_{\beta' : \|\beta - \beta'\| \le \varepsilon_n} B_1 \right) \\
&\quad = E \left[\sup_{\beta' : \|\beta - \beta'\| \le \varepsilon_n} Z_j^2 |I\{C \le g(\boldsymbol{Z}^\top \boldsymbol{\beta})\} - I\{C \le g(\boldsymbol{Z}^\top \boldsymbol{\beta}')\}| \right] \\
&\quad \le E \|\boldsymbol{Z}\|^3 \{ G(g(\boldsymbol{Z}^\top \boldsymbol{\beta}) + \varepsilon_n) - G(g(\boldsymbol{Z}^\top \boldsymbol{\beta}) - \varepsilon_n) \} \\
&\quad \le K_1 \varepsilon_n
\end{aligned}
$$

Following the similar arguments, we can show that

$$
\begin{aligned}
E \left(\sup_{\beta' : \|\beta - \beta'\| \le \varepsilon_n} B_2 \right) &= E \left[\sup_{\beta' : \|\beta - \beta'\| \le \varepsilon_n} Z_j^2 I(\epsilon_i = 1) \right. \\
&\quad \left. \times |I\{T \le g(\boldsymbol{Z}^\top \boldsymbol{\beta}), C > g(\boldsymbol{Z}^\top \boldsymbol{\beta})\} - I\{T \le g(\boldsymbol{Z}^\top \boldsymbol{\beta}'), C > g(\boldsymbol{Z}^\top \boldsymbol{\beta}')\}| \right] \\
&\le E[\|\boldsymbol{Z}\|^3 \{ G(g(\boldsymbol{Z}^\top \boldsymbol{\beta}) + \varepsilon_n) - G(g(\boldsymbol{Z}^\top \boldsymbol{\beta}) - \varepsilon_n) \} \\
&\quad + \|\boldsymbol{Z}\|^3 \{ F_1(g(\boldsymbol{Z}^\top \boldsymbol{\beta}) + \varepsilon_n) - F_1(g(\boldsymbol{Z}^\top \boldsymbol{\beta}) - \varepsilon_n) \}] \\
&\le K_2 \varepsilon_n
\end{aligned}
$$

Note that

$$
\begin{aligned}
B_3 &\le Z_j^2 \left| \left[1 - \frac{1 - \tau}{1 - F_1^*(C)} \right] I\{F_1^*(C) \le \tau\} - \left[1 - \frac{1 - \tau}{1 - F_1^{*'}(C)} \right] I\{F_1^{*'}(C) \le \tau\} \right| \\
&\quad + Z_j^2 \left| I\{C \le g(\boldsymbol{Z}^\top \boldsymbol{\beta})\} - I\{C \le g(\boldsymbol{Z}^\top \boldsymbol{\beta}')\} \right| \\
&\doteq B_{31} + B_{32}
\end{aligned}
$$

Similarly to B_1, it's easy to verify that $E\left(\sup_{\beta':\|\beta-\beta'\|\le\varepsilon_n} B_{32}\right) \le K_1\varepsilon_n$. Then

$$\begin{aligned}
B_{31} &= Z_j^2 I\{F_1^*(C) < \tau, F_1^{*'}(C) < \tau\} \frac{(1-\tau)[F_1^*(C) - F_1^{*'}(C)]}{(1-F_1^*(C))(1-F_1^{*'}(C))} \\
&\quad + Z_j^2 I\{F_1^*(C) < \tau < F_1^{*'}(C)\} \frac{1-\tau}{1-F_1^*(C)} \\
&\quad + Z_j^2 I\{F_1^{*'}(C) < \tau < F_1^*(C)\} \frac{1-\tau}{1-F_1^{*'}(C)} \\
&\le Z_j^2 \frac{F_1^*(C) - F_1^{*'}(C)}{(1-\tau)} \\
&\quad + Z_j^2 I\{F_1^*(C) < \tau < F_1^{*'}(C)\} + Z_j^2 I\{F_1^{*'}(C) < \tau < F_1^*(C)\}
\end{aligned}$$

Since

$$\begin{aligned}
& E\left[\sup_{F_1^{*'}:\|F_1^* - F_1^{*'}\|_{\mathcal{H}}} I\{F_1^*(C) < \tau < F_1^{*'}(C)\}\right] \\
&\le P\{F_1^*(C) < \tau < F_1^*(C) + \varepsilon_n\} \\
&\le G\{F_1^{*-1}(\tau)\} - G\{F_1^{*-1}(\tau - \tau)\} \le K_3\varepsilon_n.
\end{aligned}$$

Then by assumption A1, we have $E\left(\sup_{\beta':\|\beta-\beta'\|\le\varepsilon_n} B_{31}\right) \le K_4\varepsilon_n$. Consequently,

$$E\left(\sup_{\beta':\|\beta-\beta'\|\le\varepsilon_n} B_3\right) \le K_5\varepsilon_n.$$

Similar arguments to proving B_3, by adding and subtracting $\frac{\tau - F_1^{*'}(C)}{1-F_1^{*'}(C)} I\{\epsilon = 2, F_1^{*'}(C) \le \tau, T \le g(\boldsymbol{Z}^\top\boldsymbol{\beta}), C > g(\boldsymbol{Z}^\top\boldsymbol{\beta})\}$, yields

$$\begin{aligned}
B_4 &\le Z_j^2 \left| \frac{\tau - F_1^*(C)}{1-F_1^*(C)} I\{F_1^*(C) \le \tau\} - \frac{\tau - F_1^{*'}(C)}{1-F_1^{*'}(C)} I\{F_1^{*'}(C) \le \tau\} \right| \\
&\quad + Z_j^2 \left| I\{T \le g(\boldsymbol{Z}^\top\boldsymbol{\beta}'), C > g(\boldsymbol{Z}^\top\boldsymbol{\beta}')\} - I\{T \le g(\boldsymbol{Z}^\top\boldsymbol{\beta}), C > g(\boldsymbol{Z}^\top\boldsymbol{\beta})\} \right| \\
&\doteq B_{41} + B_{42}.
\end{aligned}$$

By the proof of B_{31} and B_2, we can easily get that $E\left(\sup_{\beta':\|\beta-\beta'\|\le\varepsilon_n} B_{41}\right) \le K_4\varepsilon_n$ and $E\left(\sup_{\beta':\|\beta-\beta'\|\le\varepsilon_n} B_{42}\right) \le K_2\varepsilon_n$. Thus $E\left(\sup_{\beta':\|\beta-\beta'\|\le\varepsilon_n} B_4\right) \le K_5\varepsilon_n$.

Therefore, condition (3.2) of [2] holds with $r = 2$ and $s_j = 1/2$, and condition (3.3) is satisfied by remark 3(ii) of their paper. Thus, Lemma A.2 holds by applying theorem 3 of [2]. □

PROOF OF THEOREM 3.1. Note that $F_1(t|\boldsymbol{Z}) < \tau$ is equivalent to $t < g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)$ and $F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)) = \tau$. Therefore, when plugging in the true $\boldsymbol{\beta}_0$ and F_1 into $\mathbf{M}$, we get

$$\mathbf{M}(\boldsymbol{\beta}_0, F_1) = E\{\boldsymbol{Z}\left(\tau - H(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)) - R(\boldsymbol{\beta}_0, F_1) - J(\boldsymbol{\beta}_0, F_1)\right)\} = 0.$$

Because $\boldsymbol{\beta}_0$ is the solution of $\mathbf{M}(\boldsymbol{\beta}, F_1)$ with $\mathbf{M}(\boldsymbol{\beta}, F_1)$ being a continuous function of $\boldsymbol{\beta}$ in a compact parameter neighbourhood $\mathcal{B}$.

Therefore, the consistency of $\hat{\boldsymbol{\beta}}$ is the direct conclusion of theorem 1 of [2], and we only need verify the conditions (1.1)-(1.2) and (1.5') in their paper, as (1.3) is trivially satisfied and (1.4) follows from Lemma A.1.

(1.1) By the subgradient condition of quantile regression ([9]), there exists a vector $\boldsymbol{v}$ with coordinates $|v_i| \le 1$ such that

$$\|\mathbf{M}_n(\hat{\boldsymbol{\beta}}, \hat{w})\| = n^{-1}\|(\boldsymbol{Z}_i v_i) : i \in \Xi\| = o_p(n^{-1/2}) \tag{20}$$

by assumption A.1, where Ξ denotes a $p-$element subset of $\{1, 2, \cdots, n\}$.

(1.2) For any $\varepsilon > 0$ and $\boldsymbol{\beta} \in \mathcal{B}$,

$$\begin{aligned}
&\inf_{\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\| \ge \varepsilon} \|\mathrm{M}(\boldsymbol{\beta}, F_1)\| \\
&= \inf_{\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\| \ge \varepsilon} \|\mathrm{M}(\boldsymbol{\beta}, F_1) - \mathrm{M}(\boldsymbol{\beta}_0, F_1)\| \\
&\ge \inf_{\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\| \ge \varepsilon} \|E[\boldsymbol{Z}\boldsymbol{Z}^\top(\boldsymbol{\beta} - \boldsymbol{\beta}_0)]g'(\xi^*)\{1 - G(g(\xi^*)|\boldsymbol{Z})\}f_1(g(\xi^*)|\boldsymbol{Z})\|,
\end{aligned}$$

which is strictly positive under assumptions A1 and A3. Here ξ^* is some value between $\boldsymbol{Z}^\top\boldsymbol{\beta}$ and $\boldsymbol{Z}^\top\boldsymbol{\beta}_0$.

(1.5') Let $\{a_n\}$ be a sequence of positive numbers approaching zero as $n \to \infty$. Note that $E\{\|\boldsymbol{Z}_i w_i I(X_i \le g(\boldsymbol{Z}_i^\top\boldsymbol{\beta}))\|^2\} \le E(\|\boldsymbol{Z}_i\|^2) \le K_{\boldsymbol{z}}$, under assumption A1. It then follows from Chebyshev's inequality that

$$\sup_{\boldsymbol{\beta}\in\mathcal{B}, \|F_1^* - F_1\|_{\mathcal{H} \le a_n}} \|\mathbf{M}_n(\boldsymbol{\beta}, F_1^*) - \mathbf{M}(\boldsymbol{\beta}, F_1^*)\| = o_p(1).$$

Then the proof of theorem 3.1 is complete with the conclusion theorem 1 of [2]. □

PROOF OF THEOREM 3.2. The asymptotic normality of $\hat{\boldsymbol{\beta}}$ relies on the results of theorem 2 in [2]. We need to prove the conditions (2.1)-(2.4), (2.5'), and (2.6') in their paper. The conditions (2.1), (2.4), and (2.5') hold directly by (20), Lemma A.1 and A.2, respectively.

Note that for any C_i lying above the τ-th conditional quantile $\boldsymbol{Z}_i^\top\beta_0$, the quantile fit will not be affected if we assign the entire weight to either $(\boldsymbol{Z}_i, C_i)$ or $(\boldsymbol{Z}_i, X^{+\infty})$. Then we obtain

$$\begin{aligned}
\boldsymbol{\Gamma}_1(\boldsymbol{\beta}_0, F_1) &= \frac{\partial \mathbf{M}(\boldsymbol{\beta}, F_1)}{\partial \boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} \\
&= -E[\boldsymbol{Z}\boldsymbol{Z}^\top g'(\boldsymbol{Z}^\top\beta_0)\{1 - G(g(\boldsymbol{Z}^\top\beta_0)|\boldsymbol{Z})\}f_1(g(\boldsymbol{Z}^\top\beta_0)|\boldsymbol{Z})]
\end{aligned}$$

which is continuous at β_0 and of full rank under assumption A3. For all $\boldsymbol{\beta} \in \mathcal{B}$, we define the functional derivative of $\mathbf{M}(\boldsymbol{\beta}, F_1^*)$ at F_1 in the direction $[F_1^* - F_1]$ as

$$\begin{aligned}
&\boldsymbol{\Gamma}_2(\boldsymbol{\beta}, F_1)[F_1^* - F_1] \\
&= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}[\mathbf{M}\{\boldsymbol{\beta}, F_1 + \varepsilon(F_1^* - F_1)\} - \mathbf{M}\{\boldsymbol{\beta}, F_1\}] \\
&= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}E\left[R(\boldsymbol{\beta}, F_1) - R(\boldsymbol{\beta}, F_{1\varepsilon}) + J(\boldsymbol{\beta}, F_1) - J(\boldsymbol{\beta}, F_{1\varepsilon})\right]
\end{aligned}$$

where $F_{1\varepsilon} = F_1 + \varepsilon(F_1^* - F_1)$. Since

$$\begin{aligned}
&\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}E\boldsymbol{Z}\left[R(\boldsymbol{\beta}, F_1) - R(\boldsymbol{\beta}, F_{1\varepsilon})\right] \\
&= E\boldsymbol{Z}[A_1(\boldsymbol{\beta}) + A_2(\boldsymbol{\beta})] + (1-\tau)E\boldsymbol{Z}\int_0^{g(\boldsymbol{Z}^\top\boldsymbol{\beta})} g_0(u)I\{F_1(u|\boldsymbol{Z}) \le \tau\}\frac{F_1^*(u|\boldsymbol{Z}) - F_1(u|\boldsymbol{Z})}{1 - F_1(u|\boldsymbol{Z})}du
\end{aligned}$$

where

$$A_1(\boldsymbol{\beta}) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^{g(\boldsymbol{Z}^\top\boldsymbol{\beta})} g_0(u)(1-F_1(u))\left[I\{F_1(u|\boldsymbol{Z})\leq\tau\} - I\{F_{1\varepsilon}(u|\boldsymbol{Z})\leq\tau\}du\right]$$

$$A_2(\boldsymbol{\beta}) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^{g(\boldsymbol{Z}^\top\boldsymbol{\beta})} g_0(u)(1-F_1(u))(1-\tau)\frac{I\{F_{1\varepsilon}(u|\boldsymbol{Z})\leq\tau\} - I\{F_1(u|\boldsymbol{Z})\leq\tau\}}{1-F_{1\varepsilon}(u|\boldsymbol{Z})}du$$

Similarly, we can derive

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}E\boldsymbol{Z}\left[J(\boldsymbol{\beta},F_1) - J(\boldsymbol{\beta},F_{1\varepsilon})\right]\\
&= E\boldsymbol{Z}[A_3(\boldsymbol{\beta}) + A_4(\boldsymbol{\beta})] + (1-\tau)E\boldsymbol{Z}[F_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z}) - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z})]\\
&\quad\times\int_{g(\boldsymbol{Z}^\top\boldsymbol{\beta})}^{\infty} g_0(u)I\{F_1(u|\boldsymbol{Z})\leq\tau\}\frac{F_1^*(u|\boldsymbol{Z}) - F_1(u|\boldsymbol{Z})}{(1-F_1(u|\boldsymbol{Z}))^2}du
\end{aligned}$$

where

$$\begin{aligned}
A_3(\boldsymbol{\beta}) &= [F_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z}) - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z})]\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{g(\boldsymbol{Z}^\top\boldsymbol{\beta})}^{\infty} g_0(u)\\
&\quad I\{F_1(u|\boldsymbol{Z})\leq\tau\} - I\{F_{1\varepsilon}(u|\boldsymbol{Z})\leq\tau\}du\\
A_4(\boldsymbol{\beta}) &= [F_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z}) - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z})](1-\tau)\\
&\quad\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{g(\boldsymbol{Z}^\top\boldsymbol{\beta})}^{\infty} g_0(u)\frac{I\{F_{1\varepsilon}(u|\boldsymbol{Z})\leq\tau\} - I\{F_1(u|\boldsymbol{Z})\leq\tau\}}{1-F_{1\varepsilon}(u|\boldsymbol{Z})}du
\end{aligned}$$

For $\boldsymbol{\beta}$ such that $g(\boldsymbol{Z}^\top\boldsymbol{\beta}) < g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0),\ A_1(\boldsymbol{\beta}) = 0, A_2(\boldsymbol{\beta}) = 0$. For sufficiently small ε, $F_{1\varepsilon}^{-1}(\tau) > g(\boldsymbol{Z}^\top\boldsymbol{\beta})$, then

$$A_3(\boldsymbol{\beta}) = [F_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z}) - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z})]\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left\{G(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)) - G(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z}))\right\}$$

$$A_4(\boldsymbol{\beta}) = (1-\tau)[F_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z}) - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z})]\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left\{\tilde{G}(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z})) - \tilde{G}(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0|\boldsymbol{Z}))\right\}$$

where $\frac{d\tilde{G}(u|\boldsymbol{Z})}{du} = \frac{g_0(u)}{1-F_1(u|\boldsymbol{Z})}$.

For $\boldsymbol{\beta}$ such that $g(\boldsymbol{Z}^\top\boldsymbol{\beta}) > g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0),\ A_3(\boldsymbol{\beta}) = 0, A_4(\boldsymbol{\beta}) = 0$. For sufficiently small ε, $F_{1\varepsilon}^{-1}(\tau) < g(\boldsymbol{Z}^\top\boldsymbol{\beta})$, then

$$A_1(\boldsymbol{\beta}) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left\{\breve{G}(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0|\boldsymbol{Z})) - \breve{G}(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z}))\right\}$$

where $\frac{d\breve{G}(u|\boldsymbol{Z})}{du} = g_0(u)(1-F_1(u|\boldsymbol{Z}))$ and

$$A_2(\boldsymbol{\beta}) = (1-\tau)\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left\{G(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z})) - G(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))\right\}$$

For $\boldsymbol{\beta} = \boldsymbol{\beta}_0$, note that $I\{F_0(t|\boldsymbol{Z}) < \tau\} = 1$ for $t \in (0, g(\boldsymbol{Z}^\top\boldsymbol{\beta}))$, then

$$A_1(\boldsymbol{\beta}) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left\{\breve{G}(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0|\boldsymbol{Z})) - \breve{G}(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z}))\right\}$$

$$A_2(\boldsymbol{\beta}) = (1-\tau)\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left\{G(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z})) - G(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))\right\}$$

And $I\{F_0(t|\boldsymbol{Z}) < \tau\} = 0$ for $t \in (g(\boldsymbol{Z}^\top\boldsymbol{\beta}), \infty)$, then

$$A_3(\boldsymbol{\beta}) = [F_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z}) - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z})] \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left\{G(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)) - G(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z}))\right\}$$

$$A_4(\boldsymbol{\beta}) = (1-\tau)[F_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z}) - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z})] \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left\{\tilde{G}(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z})) - \tilde{G}(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0|\boldsymbol{Z}))\right\}$$

By expanding $\tilde{G}(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z}))$ (treated as a function of ε) around $\varepsilon = 0$, and using the fact that $\frac{d}{d\varepsilon}F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z})|_{\varepsilon=0} = \frac{\tau - F_1^*(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}{f_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}$ (example 20.5 in [19]), we obtain

$$G(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z})) = G(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)) + g_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))\frac{\tau - F_1^*(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}{f_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}\varepsilon + O(\varepsilon^2).$$

Similarly, we have

$$\breve{G}(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z})) = \breve{G}(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)) + g_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))(1 - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)))\frac{\tau - F_1^*(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}{f_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}\varepsilon + O(\varepsilon^2).$$

$$\tilde{G}(F_{1\varepsilon}^{-1}(\tau|\boldsymbol{Z})) = \tilde{G}(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)) + \frac{g_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}{1 - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}\frac{\tau - F_1^*(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}{f_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}\varepsilon + O(\varepsilon^2).$$

Therefore, for $\boldsymbol{\beta}$ such that $g(\boldsymbol{Z}^\top\boldsymbol{\beta}) < g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)$,

$$\begin{aligned} A_3(\boldsymbol{\beta}) + A_4(\boldsymbol{\beta}) = &-[F_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z}) - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z})]g_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)) \\ &\frac{\tau - F_1^*(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}{f_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}\frac{F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)) - \tau}{1 - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))} \equiv 0 \end{aligned}$$

for $\boldsymbol{\beta}$ such that $g(\boldsymbol{Z}^\top\boldsymbol{\beta}) \geq g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)$,

$$A_1(\boldsymbol{\beta}) + A_2(\boldsymbol{\beta}) = g_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))(F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)) - \tau)\frac{\tau - F_1^*(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))}{f_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0))} \equiv 0$$

That is

$$\begin{aligned} &\boldsymbol{\Gamma}_2(\boldsymbol{\beta}_0, F_1)[F_1^* - F_1] \\ &= (1-\tau)E\boldsymbol{Z}\int_0^{g(\boldsymbol{Z}^\top\boldsymbol{\beta})} g_0(u)I\{F_1(u|\boldsymbol{Z}) \leq \tau\}\frac{F_1^*(u|\boldsymbol{Z}) - F_1(u|\boldsymbol{Z})}{1 - F_1(u|\boldsymbol{Z})}du \\ &\quad + (1-\tau)E\boldsymbol{Z}[F_0(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z}) - F_1(g(\boldsymbol{Z}^\top\boldsymbol{\beta})|\boldsymbol{Z})] \\ &\qquad \times \int_{g(\boldsymbol{Z}^\top\boldsymbol{\beta})}^{\infty} g_0(u)I\{F_1(u|\boldsymbol{Z}) \leq \tau\}\frac{F_1^*(u|\boldsymbol{Z}) - F_1(u|\boldsymbol{Z})}{(1 - F_1(u|\boldsymbol{Z}))^2}du \end{aligned} \tag{21}$$

With the process of Taylor expansion, we can verify condition (2.3) of [2] under assumption A1 and A2.

Then we verify condition (2.6). Combining (21) and the analysis above, we have

$$\begin{aligned} &\boldsymbol{\Gamma}_2(\boldsymbol{\beta}_0, F_1)[\hat{F}_1 - F_1] \\ &= (1-\tau)E\boldsymbol{Z}\int_0^{g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)} g_0(u)\frac{\hat{F}_1(u|\boldsymbol{Z}) - F_1(u|\boldsymbol{Z})}{1 - F_1(u|\boldsymbol{Z})}du \end{aligned} \tag{22}$$

Denote $F_1^G(t|\boldsymbol{Z}) = \frac{1}{n}\sum_{i=1}^n \frac{I\{X_i \leq t, \delta_i\epsilon_i = 1\}}{1 - G(X_i)}$, $N_i^G(t) = I(X_i \leq t, \delta_i\epsilon_i = 0)$, $Y_i(t) = I(X_i \geq t)$, $y(t) = P(X \geq t)$, $\lambda^G(t) = \lim_{\Delta\to 0} P(X \in (t, t+\Delta)|X \geq t)$, $\Lambda^G(t) = \int_0^t \lambda^G(s)ds$ and $M_i^G(t) = N_i^G - \int_0^\infty Y_i(s)d\Lambda^G(s)$. Follow the proof in [11], $\sup_{t\in[0,\nu)}\|n^{1/2}\{\hat{G}(t) - G(t) -$

$n^{-1/2}\sum_{i=1}^n G(t)\int_0^t y(s)^{-1}dM_i^G\}\| \to 0$, from [12], and $n^{-1}\sum_{i=1}^n Y_i(t)I\{X_i \le x\}I(\delta_i\epsilon_i = 1)(1-G(X_i))^{-1}$ converges to $\pi(x,t)$ uniformly in both $x \in R$ and $t \in [0,\nu)$, where $\pi(x,t) = EY_i(t)I\{X_i \le x\}I(\delta_i\epsilon_i = 1)(1-G(X_i))^{-1}$. Then

$$\begin{aligned}
\hat{F}_1(x|\boldsymbol{Z}) - F_1(x|\boldsymbol{Z}) &= F_1^G(x|\boldsymbol{Z}) - F_1(x|\boldsymbol{Z}) + \hat{F}_1(x|\boldsymbol{Z}) - F_1^G(x|\boldsymbol{Z}) \\
&= \frac{1}{n}\sum_{i=1}^n \xi_{1,i}(x) - \frac{1}{n}\sum_{i=1}^n \frac{\hat{G}(X_i) - G(X_i)}{\hat{G}(X_i)G(X_i)} I(X_i \le x)I(\delta_i\epsilon_i = 1) \\
&\approx \frac{1}{n}\sum_{i=1}^n \xi_{1,i}(x) - \frac{1}{n}\sum_{i=1}^n \frac{n^{-1}\sum_{j=1}^n Y_i(s)y(s)^{-1}dM_j^G}{G(X_i)} I(X_i \le x)I(\delta_i\epsilon_i = 1) \\
&= \frac{1}{n}\sum_{i=1}^n \xi_{1,i}(x) - \frac{1}{n}\sum_{i=1}^n \int_0^\infty \left(\sum_{j=1}^n \frac{Y_j(s)I(X_j \le x)I(\delta_j\epsilon_j = 1)}{nG(X_j)}\right)\frac{dM_i^G(s)}{y(s)} \\
&\approx \frac{1}{n}\sum_{i=1}^n \xi_{1,i}(x) - \frac{1}{n}\sum_{i=1}^n \int_0^\infty \pi(x,s)\frac{dM_i^G(s)}{y(s)} \\
&= \frac{1}{n}\sum_{i=1}^n \{\xi_{1,i}(x) - \xi_{2,i}(x)\},
\end{aligned}$$

where $\approx$ denotes asymptotic equivalence uniformly in $\tau \in [\tau_L, \tau_U]$, $\xi_{1,i}(x) = I(X_i \le x)I(\delta_i\epsilon_i = 1)G(X_i)^{-1} - F_1(x|\boldsymbol{Z})$ and $\xi_{2,i} = \int_0^\infty \pi(x,s)y(s)^{-1}dM_i^G(s), i = 1, ..., n$. Similarly derived as [11], $\int_0^\infty \pi(x,s)y(s)^{-1}dM_i^G$ is Lipshitz in x, $\hat{F}_1(x|\boldsymbol{Z}) - F_1(x|\boldsymbol{Z})$ converges weakly to a mean zero Guassian process with covariance matrix $\boldsymbol{\Sigma}(x) = E\{\boldsymbol{\xi}_1(x)'\boldsymbol{\xi}_1(x)\}$. Then by (22),

$$\begin{aligned}
&\boldsymbol{\Gamma}_2(\boldsymbol{\beta}_0, F_1)[\hat{F}_1 - F_1] \\
&\approx (1-\tau)n^{-1}\sum_{i=1}^n E_{\boldsymbol{z}}\left[\boldsymbol{Z}\int_0^{g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)} g_0(u)\frac{\xi_{1,i}(u) - \xi_{2,i}(u)}{1 - F_1(u|\boldsymbol{Z})}du\right] \\
&= (1-\tau)n^{-1}\sum_{i=1}^n \boldsymbol{\phi}_i \qquad (23)
\end{aligned}$$

where

$$\boldsymbol{\phi}_i = E_{\boldsymbol{z}}\boldsymbol{Z}\int_0^{g(\boldsymbol{Z}^\top\boldsymbol{\beta}_0)} g_0(u)\frac{\xi_{1,i}(u) - \xi_{2,i}(u)}{1 - F_1(u|\boldsymbol{Z})}du \qquad (24)$$

is a random vector with mean 0 and $E\|\boldsymbol{\phi}_i\|^2 < \infty$ by assumption A1-A3 .

Recall $\mathbf{M}_n(\boldsymbol{\beta}_0, F_1) = n^{-1}\sum_{i=1}^n \mathbf{m}_i(\boldsymbol{\beta}_0, F_1)$ being independent mean 0 random vectors.

$$\begin{aligned}
&\mathbf{m}_i(\boldsymbol{\beta}_0, F_1) \\
&= \boldsymbol{Z}_i\Big(\tau - I\{\epsilon_i = 1, T_i \le C_i, C_i \le g(\boldsymbol{Z}_i^\top\boldsymbol{\beta}_0)\} \\
&\quad - I\{\epsilon_i = 1, T_i \le C_i, g^{-1}(T_i) \le \boldsymbol{Z}_i^\top\boldsymbol{\beta}_0, C_i > g(\boldsymbol{Z}_i^\top\boldsymbol{\beta}_0)\} \\
&\quad - \frac{\tau - F_1(C_i)}{1 - F_1(C_i)}\left[I\{F_1(C_i) \le \tau, C_i \le g(\boldsymbol{Z}_i^\top\boldsymbol{\beta}_0)\}(1 - I\{T_i \le C_i, \epsilon_i = 1\})\right]\Big) \\
&\doteq \boldsymbol{Z}_i(\tau - D_1 - D_2 - D_3)
\end{aligned}$$

Since $E\mathbf{m}_i(\boldsymbol{\beta}_0, F_1) = 0$, and $D_i D_j = 0$ for $i \neq j$, it's easy to verify

$$\begin{aligned}
&\mathrm{Cov}\{\mathbf{m}_i(\boldsymbol{\beta}_0, F_1)\} \\
&= E_{\boldsymbol{Z},C} E\left\{\boldsymbol{Z}_i\boldsymbol{Z}_i^\top\left[\tau(1-\tau)I(C_i > g(\boldsymbol{Z}_i^\top\boldsymbol{\beta}_0)) + I(C_i \le g(\boldsymbol{Z}_i^\top\boldsymbol{\beta}_0))\frac{F_1(C_i)(1-\tau)^2}{1-F_1(C_i)}\right]\right\} \doteq d_1
\end{aligned}$$

Then applying the central limit theorem gives

$$n^{1/2}\{\mathbf{M}_n(\boldsymbol{\beta}_0, F_1) + \boldsymbol{\Gamma}_2(\boldsymbol{\beta}_0, F_1)[\hat{F}_1 - F_1]\} \xrightarrow{D} N(0, \mathbf{V}),$$

where

$$\begin{aligned}
&\mathbf{V} = \mathrm{Cov}\{\mathbf{m}_i(\boldsymbol{\beta}_0, F_1) + (1-\tau)\boldsymbol{\phi}_i\} \doteq d_1 + d_2 + d_2 \\
&d_1 = (1-\tau)E\{\mathbf{m}_i(\boldsymbol{\beta}_0, F_1)\boldsymbol{\phi}^\top\} \\
&d_2 = (1-\tau)^2 E\{\boldsymbol{\phi}^\top\boldsymbol{\phi}\}
\end{aligned}$$

Then the proof for (14) is thus complete by theorem 2 of [2]. □

Proof of theorem 3.3. Let $\hat{\mathcal{A}}_n = \{j : \tilde{\beta}_j \neq 0\}$. We first show that for any $j \neq \mathcal{A}$, $P(j \in \hat{\mathcal{A}}_n) \to 0$ as $n \to \infty$. Suppose there exists a $k \in \hat{\mathcal{A}}^c$ such that $|\tilde{\beta}_k| \neq 0$. Let $\boldsymbol{\beta}^*$ be a vector constructed by replacing $\tilde{\beta}_k$ with 0 in $\tilde{\boldsymbol{\beta}}$. For simplicity, we write $\hat{w}_i = w_i(\hat{F}_1)$. Note that $|\rho_\tau(a) - \rho_\tau(b)| \le |a-b|\max\{\tau, 1-\tau\} < |a-b|$. Therefore, for large enough n,

$$\begin{aligned}
&Q_p(\tilde{\boldsymbol{\beta}}, \hat{w}_i) - Q_p(\boldsymbol{\beta}^*, \hat{w}_i) \\
&= \sum_{i=1}^n \hat{w}_i\left\{\rho_\tau(g^{-1}(X_i) - \boldsymbol{Z}_i^\top\tilde{\boldsymbol{\beta}}) - \rho_\tau(g^{-1}(X_i) - \boldsymbol{Z}_i^\top\boldsymbol{\beta}^*)\right\} \\
&+ \sum_{i=1}^n (1-\hat{w}_i)\left\{\rho_\tau(g^{-1}(X^{+\infty}) - \boldsymbol{Z}_i^\top\tilde{\boldsymbol{\beta}}) - \rho_\tau(g^{-1}(X^{+\infty}) - \boldsymbol{Z}_i^\top\boldsymbol{\beta}^*)\right\} + p_{\lambda_n}(|\tilde{\beta}_k|) \\
&\ge -2\sum_{i=1}^n \|\boldsymbol{Z}_i\| \cdot |\hat{\beta}_k| + \lambda_n|\hat{\beta}_k|^{-\gamma}|\tilde{\beta}_k|
\end{aligned}$$

By theorem 3.1, $\hat{\beta}_k - \beta_k = O_p(n^{-1/2})$ and β_k, thus $\hat{\beta}_k = O_p(n^{-1/2})$. As $\sum_{i=1}^n \|\boldsymbol{Z}_i\| = O_p$ and $n^{-1}\lambda_n|\hat{\beta}_k|^{-\gamma} \ge n^{r/\gamma-1}\lambda_n \to \infty$, which yields

$$\begin{aligned}
&Q_p(\tilde{\boldsymbol{\beta}}, \hat{w}_i) - Q_p(\boldsymbol{\beta}^*, \hat{w}_i) \ge -2\sum_{i=1}^n \|\boldsymbol{Z}_i\| \cdot |\hat{\beta}_k| + \lambda_n|\hat{\beta}_k|^{-r}|\tilde{\beta}_k| \\
&\ge |\tilde{\beta}_k| n\left[-O_p(1) + n^{-1}\lambda_n|\hat{\beta}_k|^{-\gamma} \ge n^{\gamma/2-1}\lambda_n\right] > 0, \text{as} \quad n \to \infty. \qquad (25)
\end{aligned}$$

This contradicts the fact that $Q_p(\tilde{\boldsymbol{\beta}}, \hat{w}_i) \le Q_p(\boldsymbol{\beta}^*, \hat{w}_i)$.

We next show that for any $j \in \mathcal{A}$, $P(j \notin \hat{\mathcal{A}}_n) \to 0$. We write $b_{\mathcal{A}} = (b_j, j \in \mathcal{A})$ for any vector $b \in R^p$, and $B_{\mathcal{A}\mathcal{A}}$ as the sub-matrix of a $p \times p$ matrix B with both row and column indices in $\mathcal{A}$. By Taylor expansion

$$\begin{aligned}
\mathbf{M}_n(\boldsymbol{\beta}_{\mathcal{A}}, F_1^*) &= \mathbf{M}_n(\boldsymbol{\beta}_{0\mathcal{A}}, F_1) + \boldsymbol{\Gamma}_{1\mathcal{A}\mathcal{A}}(\boldsymbol{\beta}_{\mathcal{A}} - \boldsymbol{\beta}_{0\mathcal{A}}) \\
&+ \boldsymbol{\Gamma}_{2\mathcal{A}\mathcal{A}}(\boldsymbol{\beta}_{0\mathcal{A}}, F_1)[F_1^* - F_1] + o_p(n^{-1/2}) \qquad (26)
\end{aligned}$$

uniformly over $\boldsymbol{\beta}_{\mathcal{A}}, F_1$ such that $\|\boldsymbol{\beta}_{\mathcal{A}} - \boldsymbol{\beta}_{0\mathcal{A}}\| = O(n^{-1/2})$ and $\|F_1^* - F_1\|_{\mathcal{H}} = o(n^{-1/2+r})$. Let $\boldsymbol{\beta}_{\mathcal{A}} - \boldsymbol{\beta}_{0\mathcal{A}} = n^{-1/2}\mathbf{u}$, we have

$$
\begin{aligned}
n\mathbf{u}^\top \mathbf{M}_n(\boldsymbol{\beta}_{\mathcal{A}}, \hat{F}_1) = n\mathbf{u}^\top\{&\mathbf{M}_n(\boldsymbol{\beta}_{0\mathcal{A}}, F_1) + \boldsymbol{\Gamma}_{2\mathcal{A}\mathcal{A}}\} \\
&+ n^{1/2}\mathbf{u}^\top \boldsymbol{\Gamma}_{1\mathcal{A}\mathcal{A}}\mathbf{u} + o_p(n^{1/2})
\end{aligned} \tag{27}
$$

where $\boldsymbol{\Gamma}_{2\mathcal{A}\mathcal{A}} = \boldsymbol{\Gamma}_{2\mathcal{A}\mathcal{A}}(\boldsymbol{\beta}_{0\mathcal{A}}, F_1)[\hat{F}_1 - F_1]$. Therefore, with probability tending to 1,

$$
\begin{aligned}
-n\mathbf{u}^\top \mathbf{M}_n(\boldsymbol{\beta}_{\mathcal{A}}, \hat{F}_1) &\geq -n\mathbf{u}^\top\{\mathbf{M}_n(\boldsymbol{\beta}_{0\mathcal{A}}, F_1) + \boldsymbol{\Gamma}_2\} - n^{1/2}\mathbf{u}^\top \boldsymbol{\Gamma}_{1\mathcal{A}\mathcal{A}}\mathbf{u} + o(n^{1/2}) \\
&\geq k_0 n^{1/2+r}
\end{aligned} \tag{28}
$$

for some positive k_0 and $r > 0$. However, the subgradient condition (20) requires that

$$
\|n\mathbf{u}^\top \mathbf{M}_n(\boldsymbol{\beta}_{\mathcal{A}}, \hat{F}_1)\| + \lambda_n \sum_{j\in\mathcal{A}} |\hat{\beta}_j|^{-r}|\tau - I(\tilde{\beta}_j < 0)| \leq O_p(\max_i \|\boldsymbol{Z}_i\|). \tag{29}
$$

When $\lambda_n = o(n^{1/2})$ and assumption A1 holds, (28) and (29) suggest that the subgradient condition cannot hold if $\|\tilde{\boldsymbol{\beta}}_{\mathcal{A}} - \boldsymbol{\beta}_{0\mathcal{A}}\| = Kn^{-1/2}$ for some positive K. Using the monotonicity argument in [6], we can show that the subgradient condition also cannot hold if $\|\tilde{\boldsymbol{\beta}}_{\mathcal{A}} - \boldsymbol{\beta}_{0\mathcal{A}}\| > Kn^{-1/2}$. Therefore, $\|\tilde{\boldsymbol{\beta}}_{\mathcal{A}} - \boldsymbol{\beta}_{0\mathcal{A}}\| \leq Kn^{-1/2}$ with probability tending to 1. Equivalently speaking, for all $j \in \mathcal{A}$, $P(j \in \hat{\mathcal{A}}_n) \to 1$ or $P(j \notin \hat{\mathcal{A}}_n) \to 0$. The proof of theorem 3.3 is thus complete. □

Acknowledgements. The first author was supported by Scientific Research Foundation of North China University of Technology (No.110051360002). The second author was supported in part by the National Natural Science Foundation of China (No.11861042), and the China Statistical Research Project (No.2020LZ25)

SUPPLEMENTARY MATERIAL

Supplementary simulation results
Table 6 is the performance for Σ_2, which presents even a little better results than the case of Σ_1. Table 7 is for a different pair of $(p_0, p_1) = (0.6, 0.45)$ and our τs ranges from 0 to 0.4, and $\tau = 0.3$ turns out to be the quantile of 3 nonzero coefficients , which fits our simulation results. Table 8 simulates t(3) distribution in place of standard normal distribution, displaying that our estimators still behave very well for heavy-tail distributions.

REFERENCES

[1] AHN, K. W. and KIM, S.(2018).Variable selection with group structure in competing risks quantile regression. *Statistics in Medicine* **37(9)** 1577–1586.

[2] CHEN, X. and LINTON, O. and VAN KEILEGOM, I. (2003). Estimation of semiparametric models when the criterion function is not smooth. *Econometrica* **71** 1591–1608.

[3] 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(446)** 496–509.

[4] FU, Z. and PARIKH, C. R. and ZHOU, B. (2017). Penalized variable selection in competing risks regression. *Lifetime data analysis* **23(3)** 353–376.

[5] 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.

[6] JUREČKOVÁ, J. (1977). Asymptotic relations of m-estimates and r-estimates in linear regression model. *The Annals of Statistics* **5** 464-472.

[7] KAPLAN, E. L. and MEIER, P (1958). Nonparametric estimation from incomplete observations nonparametric estimation from incomplete observations. *Journal of the American statistical association* **53(282)** 457–481.

[8] KOENKER, R. W. and BASSETT, G. (1978). Regression quantile. *Econometrica* **46(1)** 33–50.

TABLE 4
P_1, $n = 200, p_0 = 0.8, p_1 = 0.6$,

ρ	τ	cqr.l	cqr.a	wcqr0.l	wcqr0.a	wcqr1.l	wcqr1.a	wcqr2.l	wcqr2.a	wcqr3.l	wcqr3.a
0	0.10	2.65	1.37	2.67	1.32	2.67	1.32	2.67	1.32	2.67	1.32
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.20	2.53	1.25	2.55	1.22	2.55	1.22	2.55	1.24	2.55	1.24
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.30	2.33	1.26	2.44	1.21	2.44	1.21	2.46	1.18	2.46	1.18
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.40	2.40	1.29	2.53	1.29	2.53	1.31	2.50	1.30	2.50	1.30
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.50	3.40	64.81	3.57	2.76	3.59	2.73	3.52	2.43	3.52	2.43
		0.01	2.82	0.01	0.03	0.01	0.03	0.01	0.02	0.01	0.02
0.25	0.10	2.63	1.26	2.74	1.30	2.74	1.30	2.75	1.31	2.75	1.31
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.20	2.27	1.24	2.36	1.23	2.36	1.23	2.38	1.24	2.38	1.24
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.30	2.32	1.20	2.40	1.18	2.40	1.18	2.43	1.22	2.43	1.22
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.40	2.47	1.46	2.54	1.46	2.53	1.47	2.57	1.47	2.57	1.47
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.50	3.27	48.32	3.35	1.98	3.32	1.99	3.11	1.93	3.11	1.93
		0.01	3.27	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.5	0.10	2.76	1.46	2.77	1.52	2.77	1.52	2.77	1.53	2.77	1.53
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.20	2.35	1.22	2.37	1.21	2.37	1.21	2.39	1.21	2.39	1.21
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.30	2.22	1.17	2.30	1.16	2.30	1.17	2.34	1.19	2.34	1.19
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.40	2.48	1.36	2.47	1.35	2.48	1.36	2.47	1.39	2.47	1.39
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.50	3.33	58.28	3.33	2.27	3.30	2.31	3.32	2.20	3.32	2.20
		0.01	3.40	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
0.75	0.10	2.56	1.59	2.60	1.58	2.60	1.58	2.58	1.58	2.58	1.58
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.20	2.26	1.37	2.24	1.33	2.24	1.33	2.25	1.35	2.25	1.35
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.30	2.22	1.41	2.18	1.31	2.18	1.31	2.18	1.35	2.18	1.35
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.40	2.40	1.46	2.33	1.47	2.33	1.46	2.34	1.45	2.34	1.45
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.50	3.40	42.46	3.39	2.77	3.40	2.74	3.40	2.71	3.40	2.71
		0.01	3.06	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02

[9] KOENKER, R. (2005). *Quantile Regression*. Cambridge University Press, New York.

[10] LI, E. and TIAN, M. and TANG, M.(2019). Variable selection in competing risks models based on quantile regression. *Statistics in Medicine* **38(23)** 4670–4685.

[11] PENG, L and FINE, J. P. (2009). Competing risks quantile regression. *Journal of the American Statistical Association*, **104(488)** 1440–1453.

[12] PEPE, M. S. (1991). Inference for Events With Dependent Risks in Multiple Endpoint Studies. *Journal of the American Statistical Association*, **86** 770ÍC778.

[13] ROBINS, J. M. and ROTNITZKY,A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. *AIDS Epidemiology Theorethodological Issues* JEWELL, N. , DIETZ, K. and FAREWELL,V., Birkhäuser,Boston, 24–33.

[14] SCRUCCA, L., SANTUCCI, A. and AVERSA,F. (2010). Regression modeling of competing risk using R: an in depth guide for clinicians. *Bone Marrow Transplant* **45(9)** 1388–1395.

[15] SUN, Y. Q., WANG, H. and GILBERT, P. (2012). Quantile regression for competing risks data with missing cause of failure. *Statistics Sinica* **22** 703–728.

TABLE 5
P_2, $n = 200, p_0 = 0.8, p_1 = 0.6,$

ρ	τ	cqr.l	cqr.a	wcqr0.l	wcqr0.a	wcqr1.l	wcqr1.a	wcqr2.l	wcqr2.a	wcqr3.l	wcqr3.a
0	0.10	2.31	0.77	2.36	0.70	2.36	0.70	2.38	0.70	2.38	0.70
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.20	2.05	0.62	2.10	0.61	2.10	0.61	2.11	0.62	2.11	0.62
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.30	1.68	0.70	1.92	0.67	1.92	0.67	1.94	0.63	1.94	0.63
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.40	1.64	0.68	1.92	0.68	1.94	0.73	1.82	0.69	1.82	0.69
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.50	3.49	56251.49	3.98	8.01	4.05	8.02	3.82	4.28	3.82	4.28
		0.03	2508.07	0.03	0.34	0.03	0.34	0.03	0.21	0.03	0.21
0.25	0.10	2.13	0.60	2.43	0.66	2.43	0.66	2.45	0.66	2.45	0.66
		0.02	0.01	0.03	0.01	0.03	0.01	0.03	0.01	0.03	0.01
	0.20	1.48	0.61	1.73	0.62	1.73	0.62	1.73	0.63	1.73	0.63
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.30	1.62	0.63	1.87	0.61	1.87	0.61	1.89	0.64	1.89	0.64
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.40	1.67	0.92	1.98	0.93	1.96	0.93	1.96	0.93	1.96	0.93
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.50	3.05	17820.66	3.36	1.61	3.23	1.62	2.91	1.55	2.91	1.55
		0.03	1266.88	0.03	0.02	0.03	0.02	0.03	0.02	0.03	0.02
0.5	0.10	2.29	0.91	2.33	0.99	2.33	0.99	2.34	0.99	2.34	0.99
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.20	1.73	0.61	1.90	0.61	1.90	0.61	1.88	0.61	1.88	0.61
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.30	1.48	0.57	1.66	0.55	1.66	0.56	1.74	0.58	1.74	0.58
		0.01	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.40	1.63	0.73	1.86	0.77	1.88	0.79	1.72	0.82	1.72	0.82
		0.01	0.01	0.02	0.01	0.02	0.01	0.01	0.01	0.01	0.01
	0.50	3.08	22616.57	3.22	1.96	3.25	1.98	3.10	1.68	3.10	1.68
		0.03	1339.64	0.03	0.03	0.03	0.03	0.03	0.02	0.03	0.02
0.75	0.10	1.94	0.86	2.03	0.89	2.03	0.89	1.94	0.88	1.94	0.88
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.20	1.50	0.62	1.49	0.63	1.49	0.63	1.48	0.65	1.48	0.65
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.30	1.41	0.67	1.42	0.62	1.41	0.62	1.43	0.69	1.43	0.69
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.40	1.59	0.77	1.59	0.77	1.59	0.76	1.58	0.77	1.58	0.77
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.50	3.11	21245.71	3.27	4.46	3.30	4.43	3.43	4.27	3.43	4.27
		0.03	1532.70	0.03	0.19	0.03	0.19	0.03	0.21	0.03	0.21

[16] TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Series B Stat. Methodol.* **58(1)** 267–288.

[17] WANG, H. J. and WANG, L. (1956). Locally weighted censored quantile regression. *Journal of the American Statistical Association* **104(487)** 1117–1128.

[18] WANG, H. J., ZHOU, J. and LI,Y. (2013). Variable selection for censored quantile regresion. *Statistica Sinica* **23(1)** 145–167.

[19] VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press, Cambridge.

[20] ZOU, H. (2006). The adaptive lasso and its oracle properties. *The Annals of Statistics* **101(476)** 1418–1429.

TABLE 6

$n = 400, \rho = 0.5, p_0 = 0.8, p_1 = 0.6, \Sigma_2,$

Criteria	τ	cqr.l	cqr.a	wcqr0.l	wcqr0.a	wcqr1.l	wcqr1.a	wcqr2.l	wcqr2.a	wcqr3.l	wcqr3.a
TP	0.10	4.00	3.99	3.98	3.99	3.98	3.99	3.98	3.99	3.98	3.99
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.20	4.00	3.99	4.00	3.99	4.00	3.99	4.00	3.98	4.00	3.98
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.30	4.00	4.00	3.99	4.00	3.99	4.00	4.00	4.00	4.00	4.00
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.40	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.50	3.98	3.97	3.95	3.94	3.96	3.95	3.98	3.96	3.98	3.96
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FP	0.10	4.83	0.43	4.54	0.38	4.54	0.38	4.56	0.37	4.56	0.37
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.20	5.00	0.45	4.31	0.36	4.31	0.36	4.23	0.36	4.23	0.36
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.30	4.42	0.37	4.08	0.20	4.11	0.20	3.99	0.18	3.99	0.18
		0.02	0.01	0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.00
	0.40	5.62	1.37	4.83	1.34	4.90	1.34	4.94	1.27	4.94	1.27
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.50	6.09	1.22	5.33	1.03	5.21	1.09	5.45	1.08	5.45	1.08
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
P_1	0.10	1.78	0.91	1.85	0.93	1.85	0.93	1.86	0.93	1.86	0.93
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.20	1.59	0.81	1.68	0.79	1.68	0.79	1.69	0.78	1.69	0.78
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.30	1.57	0.76	1.55	0.72	1.55	0.72	1.63	0.72	1.63	0.72
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.40	1.67	0.84	1.69	0.83	1.67	0.83	1.74	0.84	1.74	0.84
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.50	2.10	1.33	2.07	1.36	2.08	1.35	2.21	1.36	2.21	1.36
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
P_2	0.10	0.93	0.37	1.10	0.39	1.10	0.39	1.11	0.40	1.11	0.40
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.20	0.69	0.27	0.86	0.26	0.86	0.26	0.86	0.25	0.86	0.25
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.30	0.74	0.24	0.75	0.22	0.75	0.22	0.83	0.22	0.83	0.22
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.40	0.78	0.31	0.90	0.31	0.87	0.31	0.94	0.32	0.94	0.32
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.50	1.05	0.72	1.13	0.75	1.14	0.71	1.29	0.73	1.29	0.73
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

TABLE 7
$n = 400, \rho = 0.5, p_0 = 0.6, p_1 = 0.45, \Sigma_1,$

Criteria	τ	cqr.l	cqr.a	wcqr0.l	wcqr0.a	wcqr1.l	wcqr1.a	wcqr2.l	wcqr2.a	wcqr3.l	wcqr3.a
TP	0.10	4.00	3.98	4.00	3.98	4.00	3.98	4.00	3.98	4.00	3.98
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.20	4.00	3.98	4.00	3.98	4.00	3.98	4.00	3.98	4.00	3.98
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.30	2.99	2.97	2.98	2.94	2.98	2.94	2.99	2.94	2.99	2.94
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.40	3.37	3.57	3.36	3.45	3.33	3.44	3.63	3.68	3.63	3.68
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
FP	0.10	5.10	0.52	4.72	0.44	4.72	0.44	4.85	0.39	4.85	0.39
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.20	5.21	0.84	4.57	0.57	4.57	0.57	4.58	0.59	4.58	0.59
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.30	6.33	2.02	5.83	1.83	5.68	1.88	5.94	1.86	5.94	1.86
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.40	5.35	2.10	5.04	1.77	5.08	1.83	5.36	1.48	5.36	1.48
		0.03	0.02	0.03	0.01	0.03	0.01	0.03	0.01	0.03	0.01
P_1	0.10	1.86	0.97	1.90	1.00	1.90	1.00	1.87	0.99	1.87	0.99
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.20	1.87	1.00	1.91	0.98	1.91	0.98	1.93	0.97	1.93	0.97
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.01
	0.30	2.07	1.19	2.07	1.16	2.10	1.18	2.06	1.18	2.06	1.18
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.40	3.53	47.28	3.38	2.59	3.46	2.59	3.12	2.10	3.12	2.10
		0.01	1.91	0.01	0.02	0.01	0.02	0.01	0.01	0.01	0.01
P_2	0.10	0.91	0.39	0.99	0.43	1.00	0.43	0.96	0.42	0.96	0.42
		0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.20	0.87	0.37	1.00	0.39	1.00	0.38	1.01	0.38	1.01	0.38
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.30	1.04	0.56	1.12	0.59	1.17	0.60	1.07	0.55	1.07	0.55
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.40	3.25	21806.45	3.09	3.16	3.17	3.35	2.53	1.53	2.53	1.53
		0.03	876.37	0.03	0.10	0.03	0.11	0.03	0.02	0.03	0.02

TABLE 8

$n = 400, \rho = 0.5, p_0 = 0.8, p_1 = 0.6, \Sigma_2, t(3),$

Criteria	τ	cqr.l	cqr.a	wcqr0.l	wcqr0.a	wcqr1.l	wcqr1.a	wcqr2.l	wcqr2.a	wcqr3.l	wcqr3.a
TP	0.10	3.89	3.89	3.84	3.90	3.84	3.90	3.83	3.91	3.83	3.91
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.20	3.98	4.00	3.98	4.00	3.98	4.00	3.98	4.00	3.98	4.00
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.30	3.95	4.00	3.97	3.99	3.97	3.99	3.96	3.99	3.96	3.99
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.40	2.99	3.00	2.98	3.00	2.98	3.00	2.96	2.99	2.96	2.99
		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.50	3.75	3.80	3.66	3.74	3.64	3.74	3.73	3.82	3.73	3.82
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.00
FP	0.10	5.04	0.81	4.67	0.86	4.61	0.86	4.69	0.88	4.69	0.88
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.20	4.63	0.51	4.25	0.32	4.33	0.30	4.18	0.35	4.18	0.35
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.30	3.77	0.25	3.51	0.18	3.54	0.20	3.56	0.16	3.56	0.16
		0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.00
	0.40	5.61	1.59	5.23	1.45	5.13	1.48	5.10	1.42	5.10	1.42
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.50	5.92	1.52	5.18	1.03	5.11	1.13	5.41	1.03	5.41	1.03
		0.03	0.01	0.02	0.01	0.03	0.01	0.03	0.01	0.03	0.01
P_1	0.10	2.64	1.61	2.75	1.62	2.77	1.64	2.76	1.64	2.76	1.64
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.20	2.03	1.03	2.08	1.05	2.07	1.06	2.07	1.06	2.07	1.06
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.30	1.95	0.90	1.97	0.88	1.97	0.89	2.00	0.91	2.00	0.91
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.40	2.01	1.05	1.99	1.02	2.03	1.04	2.12	1.04	2.12	1.04
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	0.50	2.73	6.96	2.66	1.71	2.78	1.71	2.80	1.75	2.80	1.75
		0.01	0.52	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
P_2	0.10	2.20	1.12	2.51	1.13	2.53	1.15	2.52	1.16	2.52	1.16
		0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
	0.20	1.22	0.42	1.34	0.45	1.29	0.46	1.33	0.45	1.33	0.45
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.30	1.27	0.35	1.29	0.34	1.29	0.35	1.31	0.37	1.31	0.37
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.40	1.16	0.44	1.20	0.44	1.28	0.46	1.40	0.45	1.40	0.45
		0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00
	0.50	2.01	2685.65	2.07	1.29	2.28	1.27	2.19	1.21	2.19	1.21
		0.02	268.44	0.02	0.02	0.03	0.02	0.02	0.01	0.02	0.01

IRTG 1792 Discussion Paper Series 2021

For a complete list of Discussion Papers published, please visit http://irtg1792.hu-berlin.de.

001 "Surrogate Models for Optimization of Dynamical Systems" by Kainat Khowaja, Mykhaylo Shcherbatyy, Wolfgang Karl Härdle, January 2021.

002 "FRM Financial Risk Meter for Emerging Markets" by Souhir Ben Amor, Michael Althof, Wolfgang Karl Härdle, February 2021.

003 "K-expectiles clustering" by Bingling Wang, Yingxing Li, Wolfgang Karl Härdle, March 2021.

004 "Understanding Smart Contracts: Hype or Hope?" by Elizaveta Zinovyev, Raphael C. G. Reule, Wolfgang Karl Härdle, March 2021.

005 "CATE Meets ML: Conditional Average Treatment Effect and Machine Learning" by Daniel Jacob, March 2021.

006 "Coins with benefits: on existence, pricing kernel and risk premium of cryptocurrencies" by Cathy Yi-Hsuan Chen, Dmitri Vinogradov, April 2021.

007 "Rodeo or Ascot: which hat to wear at the crypto race?" by Konstantin Häusler, Wolfgang Karl Härdle, April 2021.

008 "Financial Risk Meter based on Expectiles" by Rui Ren, Meng-Jou Lu, Yingxing Li, Wolfgang Karl Härdle, April 2021.

009 "Von den Mühen der Ebenen und der Berge in den Wissenschaften" by Annette Vogt, April 2021.

010 "A Data-driven Explainable Case-based Reasoning Approach for Financial Risk Detection" by Wei Li, Florentina Paraschiv, Georgios Sermpinis, July 2021.

011 "Valuing cryptocurrencies: Three easy pieces" by Michael C. Burda, July 2021.

012 "Correlation scenarios and correlation stress testing" by Natalie Packham, Fabian Woebbeking, July 2021.

013 "Penalized Weigted Competing Risks Models Based on Quantile Regression" by Erqian Li, Wolfgang Karl Härdle, Xiaowen Dai, Maozai Tian, July 2021.

IRTG 1792, Spandauer Strasse 1, D-10178 Berlin
http://irtg1792.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the IRTG 1792.