Published in final edited form as:

J R Stat Soc Series B Stat Methodol. 2015 January ; 77(1): 107–130. doi:10.1111/rssb.12063.

Quantile Regression Adjusting for Dependent Censoring from Semi-Competing Risks

Ruosha Li and

Ruosha Li, University of Pittsburgh, Pittsburgh, USA.

Limin Peng*

Emory University, Atlanta, USA.

Summary

In this work, we study quantile regression when the response is an event time subject to potentially dependent censoring. We consider the semi-competing risks setting, where time to censoring remains observable after the occurrence of the event of interest. While such a scenario frequently arises in biomedical studies, most of current quantile regression methods for censored data are not applicable because they generally require the censoring time and the event time be independent. By imposing rather mild assumptions on the association structure between the time-to-event response and the censoring time variable, we propose quantile regression procedures, which allow us to garner a comprehensive view of the covariate effects on the event time outcome as well as to examine the informativeness of censoring. An efficient and stable algorithm is provided for implementing the new method. We establish the asymptotic properties of the resulting estimators including uniform consistency and weak convergence. The theoretical development may serve as a useful template for addressing estimating settings that involve stochastic integrals. Extensive simulation studies suggest that the proposed method performs well with moderate sample sizes. We illustrate the practical utility of our proposals through an application to a bone marrow transplant trial.

Keywords

Copula; Dependent censoring; Quantile regression; Semi-competing risks; Stochastic integral equation

1. Introduction

Quantile regression (Koenker and Bassett, 1978) is gaining increasing attention in survival analysis as a valuable alternative to traditional regression models. It targets a spectrum of conditional quantiles to achieve a comprehensive and robust analysis. Specifically, define the τ th conditional quantile of an event time T_1 as $Q_{T_1}(\tau|\mathbf{Z}) = \inf\{t : \Pr(T_1 \leq t|\mathbf{Z}) \geq \tau\}$, where $\tau \in (0, 1)$, $\mathbf{Z} = (1, \mathbf{Z})^T$, and $\mathbf{Z}$ is a $p \times 1$ covariate vector. A quantile regression model may assume

*Address for correspondence: Limin Peng, Department of Biostatistics and Bioinformatics, Emory University, Atlanta, Georgia 30322, USA. lpeng@sph.emory.edu.

$$Q_{T_1}(\tau|\mathbf{Z}) = \exp\left\{\mathbf{Z}^T \boldsymbol{\beta}_0(\tau)\right\}, \quad 0 < \tau < 1, \quad (1)$$

where $\boldsymbol{\beta}_0(\tau)$ is a vector of unknown regression coefficients representing covariate effects at the τ th quantile of the event time. By allowing $\boldsymbol{\beta}_0(\tau)$ to vary with τ , model (1) may accommodate inhomogeneous covariate effects that, for example, differ between subjects of high susceptibility to the event versus those with low risk of the event. Also note that model (1) is a strict extension of the accelerated failure time (AFT) model, which corresponds to model (1) with each component of $\boldsymbol{\beta}_0(\tau)$ being constant except for the intercept. Like with AFT model, quantile regression postulates a direct relationship between the event time and covariates, thereby rendering straightforward physical interpretations. Many research efforts have been devoted to developing methods for quantile regression with survival data subject to independent censoring (Powell 1986; Ying et al. 1995; Portnoy 2003; Peng and Huang 2008; Wang and Wang 2009; Huang 2010; among others). However, the assumption of independent censoring is problematic in many practical situations.

For example, in a multi-center trial of allogeneic bone marrow transplants (BMT) in patients with acute leukemia (Copelan et al., 1991; Klein and Moeschberger, 2005), one intermediate endpoint of interest is the development of chronic graft-versus-host disease (GVHD). However, some patients died without experiencing GVHD. The time to GVHD for these subjects was then censored by death, which may correlate with GVHD. On the other hand, the occurrence of GVHD did not prevent the subsequent patient follow-up in this study. As a result, time to death remained observable after the GVHD endpoint. Such a dependent censoring scenario falls into the paradigm of semi-competing risks (Fine et al., 2001), which permits the observation of censoring after the occurrence of the endpoint of interest, and is the focus of this work. Naively conducting competing risks analysis in the BMT example would ignore the observed data on time to death for patients who experienced GVHD during the study and hence incur unnecessary information loss.

In the presence of dependent censoring to T_1 , it would be problematic to estimate model (1) by using existing quantile regression methods that assume the time to censoring is independent of T_1 . For example, we show via simulation studies that applying Peng and Huang (2008)'s method by treating dependent censoring as independent can incur large biases in the estimation of $\boldsymbol{\beta}_0(\tau)$ (Figure 1). To date, very limited work has been done for quantile regression with dependently censored data. The competing risks quantile regression method by Peng and Fine (2009), like its precursor in one sample setting (Peng and Fine, 2007a), targets the *crude* conditional quantiles, which are conditional quantiles defined based on the cumulative incidence function, not the marginal distribution function. In practice, quantities based on the marginal distribution of T_1 can also be of interest (Jiang et al., 2003). This consideration motivates the study of quantile regression based on model (1), which formulates covariate effects on $Q_{T_1}(\tau|\mathbf{Z})$, and cannot be addressed by the method of Peng and Fine (2009).

In this paper, we develop a quantile regression method, which accommodates dependent censoring and renders inference on *net* conditional quantiles that refer to conditional quantiles defined upon the marginal distribution of the event time of interest. We focus on

the semi-competing risks setting, where the censoring event remains observable after the occurrence of the endpoint of interest, as in the BMT example. The new method provides a useful alternative to existing regression approaches for dependently censored data of semi-competing risks structure. For example, it allows for nonconstant covariate effects, which are not permitted by Lin et al. (1996), Peng and Fine (2006), Hsieh et al. (2008), Ding et al. (2009), and Chen (2011). While other varying coefficient models, such as multiplicative hazards model and additive risks model, have been studied for survival data (we refer to Martinussen and Scheike (2006) for a comprehensive coverage), approaches tailored to the semi-competing risks setting are quite limited. One available method is the functional regression model studied by Peng and Fine (2007b), which generalizes the Cox proportional hazards model with varying coefficient incorporated. In contrast, model (1) formulates covariate effects on the quantiles of T_1 , offering direct physical interpretations, and thus may be preferred in many practical situations.

In our proposals, we first tackle the issue related to the lack of nonparametric identifiability of the marginal distribution of T_1 (Tsiatis, 1975) by assuming the association between T_1 and its censoring time follows a copula model with unspecified copula parameters. Note that, this is a weaker assumption compared to that usually adopted in the competing risks setting, where the copula model is fully specified (Zheng and Klein 1995; Chen 2010; among others). We construct unbiased estimating equations based on the assumed models, effectively utilizing the semi-competing risks feature of the censoring scheme. Estimation and inference procedures are developed to offer not only analyses of *net* conditional quantiles but also insights about how censoring is associated with the event of interest. A stable and efficient algorithm is developed for the implementation of the proposed procedures.

Moreover, we establish the asymptotic properties of the proposed estimators despite considerable technical challenges. Our theoretical development provides a useful template for addressing estimating equations that involve the use of stochastic integrals. Via extensive simulations, we show that the proposed method performs well with moderate sample sizes, and are robust to misspecification of the assumed association model. An application to the BMT example illustrates the utility of our proposals, uncovering findings unattainable through traditional survival regression models.

2. Quantile Regression Procedure

2.1. Data and Model

We begin with a formal introduction of data and notation. Let T_1 be time to the endpoint of interest and T_2 be time to dependent censoring. Define $\tilde{\mathbf{Z}} = (1, \mathbf{Z}^T)^T$ as a $(p + 1) \times 1$ vector extended from covariates recorded in $\mathbf{Z}$. There often involves an administrative censoring time C which is conditionally independent of (T_1, T_2) given $\mathbf{Z}$. Define $X = T_1 \wedge T_2 \wedge C$, $Y = T_2 \wedge C$, $\delta = I(T_1 \leq T_2 \wedge C)$, $\eta = I(T_2 \leq C)$, where $\wedge$ is the minimum operator and $I(\cdot)$ is the indicator function. Observed data include n identically and independently distributed (i.i.d.) replicates of $\{X, Y, \delta, \eta, \mathbf{Z}\}$, denoted by $\{X_i, Y_i, \delta_i, \eta_i, \mathbf{Z}_i\}_{i=1}^n$.

To address the lack of nonparametric identifiability of the *net* conditional quantiles $Q_{T_1}(\tau|\mathbf{Z})$ in the presence of dependent censoring by T_2 (Tsiatis, 1975), we impose additional assumptions/ on the association structure between T_1 and T_2 . Such a strategy has been widely adopted in previous work on dependent censoring. Specifically, we adopt a copula model that assumes

$$Pr(T_1 > s, T_2 > t | \mathbf{Z}) = \Psi \left\{ 1 - F_1(s | \mathbf{Z}), 1 - F_2(t | \mathbf{Z}), g \left(\bar{\mathbf{Z}}^T \mathbf{r}_0 \right) \right\}, \quad (2)$$

where $F_i(t | \mathbf{Z}) = \Pr(T_i \leq t | \mathbf{Z})$ ($i = 1, 2$), $\bar{\mathbf{Z}} = (1, \tilde{\mathbf{Z}}_0^T)^T$ with $\mathbf{Z}_0$ being a sub-vector of $\mathbf{Z}$, or $\mathbf{Z}$ itself, and Ψ is a known function. For a given parameter θ , $\Psi(\cdot, \cdot, \theta)$ is a copula function, for example, Clayton's copula (Clayton, 1978) or Frank's copula (Genest, 1987). In model (2), $\mathbf{r}_0$ is an unknown parameter, characterizing the dependent censoring mechanism that may vary according to $\mathbf{Z}$. When $\bar{\mathbf{Z}} = \mathbf{1}$, the association between T_1 and T_2 is assumed to be homogeneous for all subjects. The copula parameter specified as $g(\bar{\mathbf{Z}}^T \mathbf{r}_0)$ is often closely connected to a measure that characterizes the strength of the association between T_1 and T_2 , where $g(\cdot)$ is a known function. For example, under the Clayton's copula model, where $\Psi(u, v, a) = (u^{-a} + v^{-a} - 1)^{-1/a}$ and $a \geq 0$, $a/(a+2)$ equals the Kendall's tau coefficient (Kendall and Gibbons, 1962). In this case, one may select $g(x) = \exp(x)$, which accounts for the fact that a needs to be non-negative.

While model (2) allows identifying the *net* conditional quantiles of T_1 , the estimation of model (1) is further facilitated by the fact that T_2 is only subject to independent censoring by C , and therefore standard censored regression techniques can be applied to estimate $F_2(t|\mathbf{Z})$. Here, we assume that the regression model for T_2 takes the same form as that for T_1 :

$$Q_{T_2}(\tau|\mathbf{Z}) = \exp \left\{ \mathbf{Z}^T \boldsymbol{\alpha}_0(\tau) \right\}, \quad \tau \in (0, 1), \quad (3)$$

where $\boldsymbol{\alpha}_0(\tau)$ is a $(p+1) \times 1$ vector of regression coefficients, which may be estimated by using Peng and Huang (2008)'s approach. Denote the resulting estimator as $\hat{\boldsymbol{\alpha}}(\cdot)$. Note that the modeling in (1)–(3) accommodates situations where T_1 and T_2 are influenced by different sets of covariates. This can be achieved by letting $\mathbf{Z}$ include all covariates that impact T_1 or T_2 and then appropriately setting zero coefficients in $\boldsymbol{\beta}_0(\tau)$ and $\boldsymbol{\alpha}_0(\tau)$. In practice, one may consider a different model for T_2 when there exists strong evidence for the inadequacy of model (3). In principle, model (3) may be replaced by any model that can generate “good” estimates for $Q_{T_2}(\tau|\mathbf{Z})$ with randomly censored data. The resulting change in the estimating equations would be substituting $\exp \left\{ \mathbf{Z}_i^T \boldsymbol{\alpha}(\tau) \right\}$ with the estimate for $Q_{T_2}(\tau|\mathbf{Z})$. In the sequel, models (1)–(3) are assumed without further mentioning.

It is interesting to note that the model studied here reduces to a bivariate location-shift model, when $\bar{\mathbf{Z}} = \mathbf{1}$ and both $\boldsymbol{\alpha}_0(\tau)$ and $\boldsymbol{\beta}_0(\tau)$ are constant except in the intercept terms. The method proposed below will provide a valuable complement with enhanced flexibility and interpretability to the regression methods by Lin et al. (1996) and Peng and Fine (2006).

2.2. Estimating Equations

Our proposal for estimating $\beta_0(\tau)$ and r_0 is based on the following two equalities:

$$Pr(X > t | Y > t, \mathbf{Z}) = \frac{Pr(T_1 > t, T_2 > t | \mathbf{Z})}{Pr(T_2 > t | \mathbf{Z})} = K_A \left\{ F_1(t | \mathbf{Z}), F_2(t | \mathbf{Z}), g \left(\bar{\mathbf{Z}}^T \mathbf{r}_0 \right) \right\}, \quad (4)$$

and for $s \leq t$,

$$Pr(X \leq s | Y > t, \mathbf{Z}) = \frac{Pr(T_1 \leq s, T_2 > t | \mathbf{Z})}{Pr(T_2 > t | \mathbf{Z})} = K_B \left\{ F_1(s | \mathbf{Z}), F_2(t | \mathbf{Z}), g \left(\bar{\mathbf{Z}}^T \mathbf{r}_0 \right) \right\}, \quad (5)$$

where $K_A(u, v, \theta) = \Psi(1 - u, 1 - v, \theta) / (1 - v)$ and $K_B(u, v, \theta) = \{1 - v, -\Psi(1 - u, 1 - v, \theta) / (1 - v)\}$.

It is suggested by (4) and (5) that the independent censoring by C may be handled through conditioning $X > t$ and $X \leq s$ on $Y > t$ with $s \leq t$. Note that (4) is derived based on the assumed joint distribution of (T_1, T_2) on the diagonal line, while (5) utilizes the information on the upper wedge of (T_1, T_2) . Therefore, these two equalities well capture the semi-competing risks feature of the observed data.

To construct estimating equations for $\beta_0(\tau)$ and r_0 based on (4) and (5), we need to further bridge $\alpha_0(\cdot)$ and $\beta_0(\cdot)$ with the distribution functions of T_1 and T_2 . This may be done based on the facts that $F_1 \left[\exp \left\{ \mathbf{Z}^T \beta_0(\tau) \right\} | \mathbf{Z} \right] = \tau$ for any $\tau \in (0, 1)$, and

$$F_2(t | \mathbf{Z}) = \int_0^1 I \{ F_2(t | \mathbf{Z}) \geq u \} du = \int_0^1 I \left[t \geq \exp \left\{ \mathbf{Z}^T \alpha_0(u) \right\} \right] du. \quad (6)$$

Motivated by (4), (5) and (6), we consider the following two quantities:

$$P_i^0(\beta, \alpha, \mathbf{r}, \tau) = I \left\{ \log X_i > \mathbf{Z}_i^T \beta(\tau) \right\} - I \left\{ \log Y_i > \mathbf{Z}_i^T \beta(\tau) \right\} \times K_A \left\{ \tau, \int_0^1 I \left\{ \mathbf{Z}_i^T \beta(\tau) \geq \mathbf{Z}_i^T \alpha(u) \right\} du, g \left(\bar{\mathbf{Z}}_i^T \mathbf{r} \right) \right\}, \quad (7)$$

$$\begin{aligned} Q_i^0(\beta, \alpha, \mathbf{r}, \tau) &= \int_0^\infty I \left\{ \mathbf{Z}_i^T \beta(\tau) \leq \log t < \log Y_i \right\} \\ &\times \left(I \left\{ \log X_i \leq \mathbf{Z}_i^T \beta(\tau) \right\} - K_B \left\{ \tau, \int_0^1 I \left\{ \log t \geq \mathbf{Z}_i^T \alpha(u) \right\} du, g \left(\bar{\mathbf{Z}}_i^T \mathbf{r} \right) \right\} \right) dt. \end{aligned} \quad (8)$$

In the above definitions, the β and α in $P_i^0(\beta, \alpha, \mathbf{r}, \tau)$ and $Q_i^0(\beta, \alpha, \mathbf{r}, \tau)$ stand for functions defined from $[0, 1]$ to R^{p+1} . The same rule applies to other similar notation without further mentioning. Under the assumed models (1)–(3), we have $E \left\{ P_i^0(\beta_0, \alpha_0, \mathbf{r}_0, \tau) | \mathbf{Z} \right\} = 0$. This can be seen from the equality $E \{ I(X > t) - I(Y > t) \Pr(X > t | Y > t, \mathbf{Z}) | \mathbf{Z} \} = 0$ with (4) plugged in.

Moreover, we can show $E \left\{ \int_{\tau} Q_i^0(\beta_0, \alpha_0, \mathbf{r}_0, \tau) d\tau | \mathbf{Z}_i \right\} = 0$ by noting that

$$E \left[\int_s \int_t I(s \leq t, t < Y) \times \{I(X \leq s) - Pr(X \leq s | Y > t, \mathbf{Z})\} dt ds | \mathbf{Z} \right] = 0,$$

and applying the monotone variable transformation $s = \exp \mathbf{Z}^T \beta_0(t)$. The double integration with respect to s and t reflects a good use of the data in the whole upper wedge of (T_1, T_2) .

Based on the above results, one may consider the following estimating equations:

$$n^{-1/2} \sum_{i=1}^n \mathbf{Z}_i P_i^0(\beta, \hat{\alpha}, \mathbf{r}, \tau) = 0, \quad n^{-1/2} \sum_{i=1}^n \int_{\tau_a}^{\tau_b} \bar{\mathbf{Z}}_i Q_i^0(\beta, \hat{\alpha}, \mathbf{r}, \tau) d\tau = 0, \quad (9)$$

where τ_a and τ_b are some prespecified constants in $(0, 1)$. Please note, in the absence of censoring (i.e. $Y_i = \infty$ for all i), under an independent copula model (i.e. $\Psi(x, y, \theta) = xy$), the first estimating equation in (9) would reduce to

$n^{-1/2} \sum_{i=1}^n \mathbf{Z}_i \left[I \left\{ \log X_i > \mathbf{Z}_i^T \beta(\tau) \right\} - \tau \right] = 0$, which is the same as that presented in Ying et al. (1995) for the case without censoring. The estimating function in this special case equals $(-1/\sqrt{n})$ times the “pseudo” derivative of the classical objective function for defining sample regression quantiles (Koenker, 2005), $\sum_{i=1}^n \rho_{\tau}(\log X_i - \mathbf{Z}_i^T \mathbf{b})$, where $\rho_{\tau}(x) = x\{\tau - I(x < 0)\}$. Such a connection may help understand the inclusion of $\mathbf{Z}_i$ in (9).

There are some potential issues with the equations in (9). More specifically, with T_2 subject to the independent censoring from C , $\alpha_0(\tau)$ may not be identifiable for some τ close to 1 (Peng and Huang, 2008). In this case, problems would arise because $P_i^0(\beta, \hat{\alpha}, \mathbf{r}, \tau)$ and $Q_i^0(\beta, \hat{\alpha}, \mathbf{r}, \tau)$ requires estimating $\alpha_0(\tau)$ for all $\tau \in (0, 1)$, which may not be feasible due to the lack of identifiability of $\alpha_0(\tau)$ in the upper tail of τ . To circumvent this issue, we modify the equations in (9) based on the idea of “truncating” $P_i^0(\beta_0, \alpha_0, \mathbf{r}_0, \tau)$ and $Q_i^0(\beta_0, \alpha_0, \mathbf{r}_0, \tau)$ conditioning on $\mathbf{Z}_i^T \beta_0(\tau) \leq \mathbf{Z}_i^T \alpha_0(\tau_{U,2})$ and $t \leq \exp \left\{ \mathbf{Z}_i^T \alpha_0(\tau_{U,2}) \right\}$ respectively, where $\tau_{U,2}$ is an upper bound of a τ -range, in which $\alpha_0(\tau)$ is identifiable. Some empirical rule for selecting $\tau_{U,2}$ is provided in Peng and Huang (2008). Applying this truncation idea leads to the consideration of

$$P_i(\beta, \alpha, \mathbf{r}, \tau) = I \left\{ \log X_i > \mathbf{Z}_i^T \beta(\tau) \right\} - I \left\{ \log Y_i > \mathbf{Z}_i^T \beta(\tau) \right\} \times K_A \left\{ \tau, \int_0^{\tau_{U,2}} I \left\{ \mathbf{Z}_i^T \beta(\tau) \geq \mathbf{Z}_i^T \alpha(\mathbf{u}) \right\} d\mathbf{u}, g \left(\frac{\tau}{\mathbf{Z}_i^T \mathbf{r}} \right) \right\},$$

which involves $\alpha(u)$ for u up to $\tau_{U,2}$ instead of 1. We have

$$\begin{aligned} Q_i(\beta, \alpha, \mathbf{r}, \tau) &= \int_{t \in (0, \infty)} I \left\{ \mathbf{Z}_i^T \beta(\tau) \leq \log t \leq \mathbf{Z}_i^T \alpha(\tau_{U,2}) \wedge \log Y_i \right\} \\ &\times \left(I \left\{ \log X_i \leq \mathbf{Z}_i^T \beta(\tau) \right\} - K_B \left[\tau, \int_0^{\tau_{U,2}} I \left\{ \log t \geq \mathbf{Z}_i^T \alpha(\mathbf{u}) \right\} d\mathbf{u}, g \left(\frac{\tau}{\mathbf{Z}_i^T \mathbf{r}} \right) \right] \right) dt, \end{aligned}$$

for $\tau \in (0, 1)$. The first equality holds because

$$\int_0^{\tau_{U,2}} I \left\{ \mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \geq \mathbf{Z}_i^T \boldsymbol{\alpha}_0(u) \right\} du = \int_0^1 I \left\{ \mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \geq \mathbf{Z}_i^T \boldsymbol{\alpha}_0(u) \right\} du \text{ when}$$

$\mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \leq \mathbf{Z}_i^T \boldsymbol{\alpha}_0(\tau_{U,2})$, and the second and the third equalities follow directly from (4) and (6). With similar arguments, we can also show that the integrand in $Q_i(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau)$ has expectation 0, and hence

$$\begin{aligned} & E \left(I \left\{ \mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \leq \mathbf{Z}_i^T \boldsymbol{\alpha}_0(\tau_{U,2}) \right\} P_i(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau) \right) \\ &= E \left(I \left\{ \mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \leq \mathbf{Z}_i^T \boldsymbol{\alpha}_0(\tau_{U,2}) \right\} \left[I \left\{ \log X_i > \mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \right\} - I \left\{ \log Y_i > \mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \right\} \times K_A \left\{ \tau, \int_0^1 I \left\{ \mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \geq \mathbf{Z}_i^T \boldsymbol{\alpha}_0(u) \right\} du \right\} \right] \right) \\ &= E \left(I \left\{ \mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \leq \mathbf{Z}_i^T \boldsymbol{\alpha}_0(\tau_{U,2}) \right\} \left[I \left\{ \log X_i > \mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \right\} - I \left\{ \log Y_i > \mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \right\} \times K_A \left(F_1 \left[\exp \left\{ \mathbf{Z}_i^T \boldsymbol{\beta}_0(\tau) \right\} | \mathbf{Z} \right], I \right) \right] \right) \end{aligned} \quad (10)$$

By (10) and (11), we propose the following estimating equations:

$$EQ_i(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau) = 0. \quad (11)$$

where

$$n^{1/2} S_n(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}, \mathbf{r}, \tau) = 0, \quad n^{1/2} \int_{\tau_a}^{\tau_b} \mathbf{W}_n(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}, \mathbf{r}, \tau) d\tau = 0, \quad (12)$$

Taking into consideration the identifiability of $\boldsymbol{\beta}_0(\tau)$ due to censoring, we restrict our attention to $\boldsymbol{\beta}_0(\tau)$ with $\tau \in (0, \tau_{U,1}]$, where $\tau_{U,1}$ is a constant less than 1 subject to certain regularity conditions given in Appendix A. In practice, $\tau_{U,1}$ and $\tau_{U,2}$ may be selected in an adaptive manner as suggested by Peng and Huang (2008) for randomly censored data. We suggest choosing $[\tau_a, \tau_b]$ such that the interval covers most of $(0, \tau_{U,1}]$ but stays away from the boundaries. The proposed estimators of $\boldsymbol{\beta}_0(\tau)$ and $\mathbf{r}_0$ can be obtained as the solutions to the equations in (12), denoted by $\hat{\boldsymbol{\beta}}(\tau)$ and $\hat{\mathbf{r}}$ respectively.

Remark: To compute $Q_i(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}, \mathbf{r}, \tau)$, one only needs to evaluate the integration over

$t \in (0, \max_{i=1}^n Y_i \wedge \exp \left\{ \mathbf{Z}_i^T \hat{\boldsymbol{\alpha}}(\tau_{U,2}) \right\})$. When $\boldsymbol{\beta}(\cdot)$ is a cadlag function, the integrand in $Q_i(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}, \mathbf{r}, \tau)$ is a piecewise constant function t , and hence $Q_i(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}, \mathbf{r}, \tau)$ can be calculated exactly.

2.3 Computational Algorithm

We developed an efficient iterative algorithm to jointly estimate $\boldsymbol{\beta}_0(\cdot)$ and $\mathbf{r}_0$ based on (12). In the sequel, we may use $\boldsymbol{\beta}$ as a shorthand for $\boldsymbol{\beta}(\tau)$ and $\boldsymbol{\alpha}$ for $\boldsymbol{\alpha}(\tau)$. Similarly, we may write $\int_{\tau_a}^{\tau_b} \mathbf{W}_n(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}, \mathbf{r}, \tau) d\tau$ as $\mathbf{W}_n(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}, \mathbf{r})$ for brevity.

First, let $\mathcal{G}_{L(n)} = \{0 < \tau_0 = \tau_1 = \dots < \tau_{L(n)} = \tau_{U,1} < 1\}$ be a prespecified grid on $[0, 1]$, and $\|\mathcal{G}_{L(n)}\| = \max \{\tau_j - \tau_{j-1}; j=1, 2, \dots, L(n)\}$ be the size of the grid. For any given $\mathbf{r}$, let $\hat{\boldsymbol{\beta}}(\tau; \mathbf{r})$, with a shorthand $\hat{\boldsymbol{\beta}}(\mathbf{r})$, denote a right-continuous step function of τ which jumps

only on $\mathcal{G}_{L(n)}$ and satisfies $n^{1/2} S_n \left\{ \hat{\beta}(\tau), \hat{\alpha}, \tau, \tau \right\} = o(1)$ for $\tau \in \mathcal{G}_{L(n)}$. The detailed algorithm follows.

Step A Estimate $\left\{ \hat{\alpha}(\tau), \tau \in (0, \tau_{U,2}] \right\}$ using Peng and Huang (2008)'s method.

Step B Set $k = 0$ and choose an initial value for $\hat{\beta}$ and $\hat{r}$, denoted by $\hat{\beta}^{[k]}$ and $\hat{r}^{[k]}$.

Step C Update $\hat{\beta}^{[k]}$ with $\hat{\beta}^{[k+1]} = \hat{\beta}(\hat{r}^{[k]})$.

Step D Solve $\hat{r}^{[k]}$ from $n^{1/2} W_n \left(\hat{\beta}^{[k]}, \hat{\alpha}, \hat{r} \right) = 0$. Then increase k by 1 and go to Step C until certain convergence criteria are satisfied.

In practice, $\hat{\beta}^{[0]}$ may be chosen as the naive estimate obtained by treating (X, δ) as independently censored survival data and employing Peng and Huang (2008)'s method, and $\hat{r}^{(0)}$ can be chosen as the $\hat{r}$ that correspond to the Kendall's tau coefficient between X_i and Y_i ,

$i = 1, 2, \dots, n$. Since $n^{1/2} W_n \left(\hat{\beta}^{[k]}, \hat{\alpha}, \hat{r} \right)$ is a smooth function of $\hat{r}$, the root-finding in Step D can be easily implemented with existing statistical functionalities, like the *optim* function

in R. To solve $n^{1/2} S_n \left(\hat{\beta}, \hat{\alpha}, \hat{r}^{[k]}, \tau \right) = 0$ in Step C, we propose to obtain

$\hat{\beta}^{[k+1]}(\tau) \equiv \hat{\beta}(\tau; \hat{r}^{[k]})$ through the following iterative procedure:

C.0 Set $m = 0$ and let $\hat{\beta}^{[k+1,m]}(\tau) = \hat{\beta}^{[k]}(\tau)$.

C.1 Find $\hat{\beta}^{[k+1,m+1]}(\tau)$ by solving

$$\begin{aligned} S_n(\beta, \alpha, \tau) &= n^{-1} \sum_{i=1}^n Z_i I \left\{ Z_i^T \beta(\tau) \leq Z_i^T \alpha(\tau_{U,2}) \right\} P_i(\beta, \alpha, \tau), \\ W_n(\beta, \alpha, \tau) &= n^{-1} \sum_{i=1}^n \bar{Z}_i Q_i(\beta, \alpha, \tau). \end{aligned}$$

for $\mathbf{b}$, where $\hat{\beta}_i^{[m]}(\tau) = I \left\{ Z_i^T \hat{\beta}^{[k+1,m]}(\tau) \leq Z_i^T \hat{\alpha}(\tau_{U,2}) \right\}$ and

$$n^{-1/2} \sum_{i=1}^n Z_i \hat{B}_i^{[m]}(\tau) \left\{ I \left(\log X_i > Z_i^T \mathbf{b} \right) - \hat{A}_i^{[m]}(\tau) \right\} = 0 \quad (13)$$

C.2 Increase m by 1 and go to C.1 until certain convergence criteria are satisfied.

It can be shown that the estimating function in (13) is a monotone random field in $\mathbf{b}$ (Fyngenson and Ritov, 1994), and equals the derivative of the following L_1 -type function

$$\hat{A}_i^{[m]}(\tau) = I \left\{ \log Y_i > \mathbf{Z}_i^T \hat{\beta}^{[k+1,m]}(\tau) \right\} K_A \left[\tau, \int_0^{\tau_{U,2}} I \left\{ \mathbf{Z}_i^T \hat{\beta}^{[k+1,m]}(\tau) \geq \mathbf{Z}_i^T \hat{\alpha}(u) \right\} du, g \left(\frac{\tau}{\mathbf{Z}_i^T \hat{\mathbf{r}}^{[k]}} \right) \right].$$

where M is an extremely large number that can bound $|\sum_{i=1}^n \hat{\beta}_i^{[m]}(\tau) \mathbf{Z}_i^T \mathbf{b}|$ and $|2 \sum_{i=1}^n \hat{\beta}_i^{[m]}(\tau) \hat{A}_i^{[m]}(\tau) \mathbf{Z}_i^T \mathbf{b}|$ from the above. Such a L_1 -type minimization problem can be readily implemented in the *rq* function in R or the *lfit* function in Spls. In Section 3, we show that the proposed algorithm can produce fast and stable implementation of the proposed estimation method. More details of the algorithm, including convergence criteria adopted for iterations, are provided in Appendix D.

2.4. Asymptotic Results

In this subsection we outline the asymptotic properties of the proposed estimators. In the below, Theorem 1 states the consistency of $\hat{\mathbf{r}}$ and the uniform consistency of $\hat{\beta}(\tau)$, and Theorem 2 gives the results on the limiting distribution of $n^{1/2}(\hat{\mathbf{r}} - \mathbf{r}_0)$ and $n^{1/2} \left\{ \hat{\beta}(\tau) - \beta_0(\tau) \right\}$. In order to establish two theorems, we need regularity conditions C1-C5, the details of which are relegated to Appendix A.

THEOREM 1. *Under conditions C1-C5, if $\lim_{n \rightarrow \infty} \|\mathcal{G}_{L(n)}\| = 0$, then there exists $\left\{ \hat{\beta}(\tau), \hat{\mathbf{r}} \right\}$ in a neighborhood of $\{\beta_0(\tau), \mathbf{r}_0\}$, such that $\sup_{\tau \in [\nu_1, \tau_{U,1}]} \|\hat{\beta}(\tau) - \beta_0(\tau)\| \xrightarrow{p} 0$, for any $0 < \nu_1 < \tau_{U,1}$ and $\hat{\mathbf{r}} \xrightarrow{p} \mathbf{r}_0$.*

THEOREM 2. *Under conditions C1-C5, if $\lim_{n \rightarrow \infty} n^{1/2} \|\mathcal{G}_{L(n)}\| = 0$, then $n^{1/2}(\hat{\mathbf{r}} - \mathbf{r}_0)$ converges in distribution to a Normal distribution with mean 0, and $n^{1/2} \left\{ \hat{\beta}(\tau) - \beta_0(\tau) \right\}$ converges weakly to a mean 0 Gaussian process for $\tau \in [\nu_1, \tau_{U,1}]$, where $0 < \nu_1 < \tau_{U,1}$.*

The proofs for Theorems 1–2 involve extensive use of empirical process theory and stochastic integrals. In the following we sketch the major steps taken to prove the theorems, which can provide a useful template for asymptotic studies in similar settings. To this end, we need to introduce additional notation. Let $s(\beta, \alpha, \mathbf{r}, \tau) = E\{S_n(\beta, \alpha, \mathbf{r}, \tau)\}$, $w(\beta, \alpha, \mathbf{r}, \tau) = E\{W_n(\beta, \alpha, \mathbf{r}, \tau)\}$, and $w(\beta, \alpha, \mathbf{r}) = \int_{\tau_a}^{\tau_b} w(\beta, \alpha, \mathbf{r}, \tau) d\tau$.

To prove the consistency in Theorem 1, we first show that $\int_0^{\tau_{U,2}} I \left[t \geq \exp \left\{ \mathbf{z}^T \hat{\alpha}(u) \right\} \right] du$ provides a uniformly consistent estimate for $F_2(t|z) \wedge \tau_{U,2}$. By this result, applications of empirical process theory lead to $\sup_{\tau \in (0, \tau_{U,1}]} s \left\{ \hat{\beta}(\tau), \alpha_0, \mathbf{r}, \tau \right\} = o_p(1)$ for any fixed $\mathbf{r} \in \mathcal{R}(d_R)$, which implies the uniform consistency of $\hat{\beta}(\tau; \mathbf{r})$ to $\tilde{\beta}(\tau; \mathbf{r})$, the solution of $s(\beta, \alpha_0, \mathbf{r}, \tau) = 0$. Similarly, we can show $\|\tilde{w}(\hat{\alpha}, \mathbf{r}) - \tilde{w}(\alpha_0, \mathbf{r})\| \rightarrow_p 0$, where $\tilde{w}(\alpha, \mathbf{r}) = w(\tilde{\beta}(\mathbf{r}), \alpha, \mathbf{r})$, and then $w(\tilde{\alpha}_0, \mathbf{r}) = o_p(1)$, which implies the consistency of $\hat{\mathbf{r}}$.

Next, we show that $\tilde{\beta}(\tau; \mathbf{r})$ has a bounded derivative with respect to $\mathbf{r}$ at $\mathbf{r} = \mathbf{r}_0$. This, coupled with the consistency of $\mathbf{r}$, gives $\sup_{\tau \in (\tau_U, 1]} \|\tilde{\beta}(\tau; \hat{\mathbf{r}}) - \tilde{\beta}(\tau; \mathbf{r}_0)\| = o_p(1)$, and hence the uniform consistency of $\hat{\beta}(\tau) \equiv \hat{\beta}(\tau; \hat{\mathbf{r}})$.

Establishing the asymptotic normality of $\mathbf{r}$ and the weak convergence of $\beta(\tau)$ requires considerable efforts to address difficulties due to the non-smoothness of the estimating functions and also the complicated entanglement between $\hat{\beta}(\tau) \equiv \hat{\beta}(\tau; \hat{\mathbf{r}})$ and $\mathbf{r}$. A brief outline of our proof follows. First, we establish the weak convergence of $n^{1/2} [s\{\hat{\beta}(\mathbf{r}_0), \alpha_0, \mathbf{r}_0, \tau\} - s\{\beta_0, \alpha_0, \mathbf{r}_0, \tau\}]$ and that of $n^{1/2} \mathbf{W}_n\{\hat{\beta}(\mathbf{r}_0), \hat{\alpha}, \mathbf{r}_0\}$. The former result entails a uniform i.i.d. representation of $n^{1/2}\{\hat{\beta}(\tau; \mathbf{r}_0) - \beta_0(\tau)\}$. Secondly, we show the asymptotic linearity of $n^{1/2}\{\hat{\beta}(\tau; \hat{\mathbf{r}}) - \hat{\beta}(\tau; \mathbf{r}_0)\}$ with respect to $n^{1/2}(\mathbf{r} - \mathbf{r}_0)$. This result facilitates deriving an i.i.d. representation of $n^{1/2}(\mathbf{r} - \mathbf{r}_0)$ from the arguments for the weak convergence of $n^{1/2} \mathbf{W}_n\{\hat{\beta}(\mathbf{r}_0), \hat{\alpha}, \mathbf{r}_0\}$. Lastly, noting that $n^{1/2}\{\hat{\beta}(\tau) - \beta_0(\tau)\} = n^{1/2}\{\hat{\beta}(\tau; \hat{\mathbf{r}}) - \hat{\beta}(\tau; \mathbf{r}_0) + \hat{\beta}(\tau; \mathbf{r}_0) - \beta_0(\tau)\}$, the asymptotic linearization of $n^{1/2}\{\hat{\beta}(\tau; \hat{\mathbf{r}}) - \hat{\beta}(\tau; \mathbf{r}_0)\}$, coupled with the i.i.d. representations of $n^{1/2}\{\hat{\beta}(\tau; \mathbf{r}_0) - \beta_0(\tau)\}$ and $n^{1/2}(\mathbf{r} - \mathbf{r}_0)$, renders the weak convergence of $n^{1/2}\{\hat{\beta}(\tau) - \beta_0(\tau)\}$. The detailed proofs of Theorems 1–2 are provided in Appendices B–C.

2.5. Other Inferences

Due to the complexity involved in the limit distribution of $\hat{\beta}(\tau)$ and $\mathbf{r}$, we propose to use a bootstrap method for the inference on $\beta_0(\tau)$ and $\mathbf{r}_0$. Specifically, we employ the paired-bootstrap scheme (Efron, 1979) and apply the algorithm presented in Section 2.3 to each of the B bootstrapped datasets to acquire $\{\hat{\alpha}_b^*, \hat{\beta}_b^*, \hat{\mathbf{r}}_b^*\}_{b=1}^B$. For a fixed $\tau \in (0, \tau_{U,1}]$, one may use the sample variance of $\{\hat{\beta}_b^*(\tau)\}_{b=1}^B$ to estimate the asymptotic variance of $\hat{\beta}(\tau)$. Similarly, the variance of $\mathbf{r}$ can be approximated by the sample variance of $\{\hat{\mathbf{r}}_b^*\}_{b=1}^B$. The confidence intervals for $\beta_0(\tau)$ and $\mathbf{r}_0$ can be constructed using normal approximation, or with the empirical percentiles of the bootstrap estimates.

We can also perform second stage inferences for exploring the varying pattern of covariate effects over $\tau \in [l, u]$, where $[l, u] \subset (0, \tau_{U,1}]$ denotes a prespecified interval of interest. For example, one may let $[l, u] = [\tau_a, \tau_b]$. First, we define a trimmed mean effect,

$\Phi_{1,q} = \left\{ \int_l^u \beta_0^{(q)}(\tau) d\tau \right\} / (u - l)$, to summarize the effect of $Z^{(q)}$. Here and hereafter, we use $u^{(q)}$ to denote the q th component of a vector $\mathbf{u}$. As a natural estimate for $\Phi_{1,q}$,

$\hat{\Phi}_{1,q} = \left\{ \int_l^u \hat{\beta}^{(q)}(\tau) d\tau \right\} / (u - l)$ can be shown to be consistent and asymptotically normal under mild assumption on the functional form of $\beta_0(\cdot)$. The limit distribution of $\hat{\Phi}_{1,q}$ can be

estimated by the empirical distribution of the bootstrap realizations,

$\left\{ \left\{ \int_l^u \hat{\beta}_b^{*(q)}(\tau) d\tau \right\} / (u-l) \right\}_{b=1}^B$. This result can be easily extended to a Wald-type or percentile-based test based on $\hat{\Phi}_{1,q}$ for evaluating whether a covariate $Z^{(q)}$ ($2 \leq q \leq p+1$) has a significant effect on a range of quantiles with $\tau \in [l, u]$, namely, $H_{01} : \beta^{(q)}(\tau) = 0, \tau \in [l, u]$. In practice, one may be interested in accessing whether the effect of a covariate $Z^{(q)}$ is constant for $\tau \in [l, u]$. The null hypothesis may be formulated as $H_{02} : \beta_0^{(q)}(\tau) = c_0, \tau \in [l, u]$, where c_0 is an unspecified constant. With a predetermined nonconstant weight function $\Xi(\tau)$ satisfying $\int_l^u \Xi(\tau) d\tau = 1$, the test statistic can be constructed as

$\mathcal{T}_{2,q} = \int_l^u \Xi(\tau) \hat{\beta}^{(q)}(\tau) d\tau - \hat{\Phi}_{1,q}$, the distribution of which can be approximated by the

empirical distribution of $\left\{ \int_l^u \Xi(\tau) \hat{\beta}_b^{*(q)}(\tau) d\tau - \left\{ \int_l^u \hat{\beta}_b^{*(q)}(\tau) d\tau \right\} / (u-l) \right\}_{b=1}^B$. Wald-type or percentile-based hypothesis testing can be performed accordingly. Rejection of H_{02} indicates that the effect of $Z^{(q)}$ may not be constant across the quantiles. Justifications of the presented second-stage inferences would follow similar lines of Peng and Huang (2008) and thus are omitted in the paper

3. Simulation Studies

We evaluated the finite-sample performance of the proposed estimators via extensive Monte Carlo simulations. Let $\mathbf{Z} = (1, Z_1, Z_2)^T$, where $Z_1 \sim \text{Unif}(0, 1)$ and $Z_2 \sim \text{Bernoulli}(0.5)$. The model used for generating T_1 is $\log T_1 = b_1 Z_1 + b_2 Z_2 + \varepsilon_1$, where the error term ε_1 is $\text{Normal}(0, 0.25^2)$ when $Z_2 = 0$, and $\text{Normal}(0, 0.5^2)$ when $Z_2 = 1$. With this heteroscedastic error structure, the underlying regression quantile $\beta_0(\tau) = \{\beta_0^{(1)}(\tau), \beta_0^{(2)}(\tau), \beta_0^{(3)}(\tau)\}$, where $\beta_0^{(1)}(\tau) = Q_{\text{norm}}(\tau, 0, 0.25^2)$, $\beta_0^{(2)}(\tau) = b_1$ and $\beta_0^{(3)}(\tau) = Q_{\text{norm}}(\tau, 0, 0.5^2) - Q_{\text{norm}}(\tau, 0, 0.25^2) + b_2$. Here $Q_{\text{norm}}(\tau, \mu, \sigma^2)$ denotes the τ -th quantile of a $N(\mu, \sigma^2)$ variable. Notice that both $\beta_0^{(1)}(\tau)$ and $\beta_0^{(2)}(\tau)$ vary with τ while $\beta_0^{(3)}(\tau)$ is constant. We generated the dependent censoring time T_2 from a log-linear model with i.i.d. errors, $\log T_2 = a_1 Z_1 + a_2 Z_2 + \varepsilon_2$, where $\varepsilon_2 \sim \text{Normal}(\mu_2, 0.5^2)$. For the association structure between T_1 and T_2 , we considered two types of copulas: the Clayton's copula and the Frank's copula. The bivariate survival function $\Pr(T_1 > s, T_2 > t | \mathbf{Z})$ is given by

$$-\frac{n^{-1/2}}{2} \left[\sum_{i=1}^n \hat{B}_i^{[m]}(\tau) |\log(X_i) - \mathbf{Z}_i^T \mathbf{b}| + |M - \sum_{i=1}^n \hat{B}_i^{[m]}(\tau) \mathbf{Z}_i^T \mathbf{b}| + |M + 2 \sum_{i=1}^n \hat{B}_i^{[m]}(\tau) \hat{A}_i^{[m]}(\tau) \mathbf{Z}_i^T \mathbf{b}| \right], \quad (14)$$

when the Clayton's copula is adopted, and by

$$\left\{ \Pr(T_1 > s | \mathbf{Z})^{-\exp(r_c)} + \Pr(T_2 > t | \mathbf{Z})^{-\exp(r_c)} - 1 \right\}^{-1/\exp(r_c)}$$

when the Frank's copula is adopted. For a fixed $\mathbf{Z}$, the Kendall's tau coefficient between T_1 and T_2 equals $\exp(r_c)/\{\exp(r_c) + 2\}$ under the Clayton's model, and equals $1 + 4 \{D_1(r_f) -$

$1\}/r_f$ under the Frank's model, where $D_1(r) = [\int_0^r t / \{exp(t) - 1\} dt] r$. We set $r_c = \log(2)$ and $r_f = 5.75$ so that the corresponding Kendall's tau coefficients equal 0.5 under both copula models. The independent censoring time C was set to follow $Unif(0, U_C)$. We considered different combinations of $\mu_2, (a_1, a_2), (b_1, b_2)$ and U_C , which led to the 4 set-ups presented in Table 1.

Under each set-up, we generated 1000 simulated datasets and implemented the proposed numerical algorithm to obtain $\hat{\beta}(\tau)$ and $\mathbf{r}$. We chose sample size $n = 200$, grid size $\|\mathcal{G}_{L(n)}\| = 0.01$, and bootstrap sample size $B = 150$. We set $(\tau_a, \tau_b) = (0.1, 0.75)$ under S1.C and S2.F, and let $(\tau_a, \tau_b) = (0.1, 0.65)$ under S2.C and S2.F, which involve heavier censoring. The simulation results on $\hat{\beta}(\tau)$ at different τ 's are summarized in Table 2. We report the empirical biases (EBias), the empirical standard errors (ESD), and the average resampling-based standard errors (ASD) for $\hat{\beta}(\tau)$, as well as the empirical coverage probabilities of the 95% Wald-type (ECP_W) and percentile-based (ECP_P) confidence intervals of $\beta_0(\tau)$. It is shown that $\hat{\beta}(\tau)$ has small biases under all set-ups, and the bootstrap standard errors closely match the empirical standard errors. Both the Wald-type and the percentile-based confidence intervals achieve empirical coverage probabilities that are close to the nominal level. It is observed that the percentile-based confidence intervals may perform better than the Wald-type intervals when τ is small.

Table 3 summarizes the simulation results on $\mathbf{r}$. We present the same set of summary statistics including EBias, ESD, ASD, ECP_W and ECP_P that are presented in Table 2, along with the true values of $\mathbf{r}_0$ and Kendall's tau coefficient, which is expressed as a known function of $\mathbf{r}_0, \mathcal{K}(\mathbf{r}_0)$; see the column labeled as TRUE. We can see that $\mathbf{r}$ is virtually unbiased, and the estimated standard errors are close to their empirical counterparts. For the estimation of $\mathbf{r}_0$ and $\mathcal{K}(\mathbf{r}_0)$, we note that the coverage rates of the Wald-type confidence intervals tend to be higher than those of the percentile-based intervals. In Table 3, we also present the empirical biases and empirical standard errors of the corresponding estimates for Kendall's tau, denoted by $\mathcal{K}(\hat{\mathbf{r}})$. It is shown that, as an established measure of association, Kendall's tau coefficient can be accurately estimated with small standard errors.

The results of second stage inferences are summarized in Table 4. We present the EBias,

ESD and ASD of the trimmed mean effect estimates $\hat{\Phi}_1^{(q)} = \left\{ \int_l^u \hat{\beta}^{(q)}(\tau) d\tau \right\} / (u - l)$, where $q = 2, 3$, and $(l, u) = (\tau_a, \tau_b)$. We also summarize the performance of the proposed hypothesis tests for $H_{01}: \beta_0^{(q)}(\tau) = 0$, and $H_{02}: \beta_0^{(q)}(\tau) \equiv c_0$, following the procedures in Section 2.5. To construct the test statistics for H_{02} , we chose the weight function $\Xi(\tau)$ as

$2I\left(\tau \leq \frac{\tau_L + \tau_U}{2}\right) / (\tau_U - \tau_L)$. For each test, we report the empirical rejection rates based on normal approximation (ERR_W) and those based on percentiles (ERR_P). We find that the

trimmed mean effect estimates $\hat{\Phi}_1^{(q)}, q=2, 3$, are well-performed in terms of biases and standard error estimates. For H_{01} , we observe that both Wald-type test and percentile-based test achieve empirical sizes close to the nominal level 0.05 and have satisfactory power.

Regarding the constancy test, the percentile-based test, as compared to the Wald-based test,

may be more conservative in terms of size but have higher power in the presence of non-constant effects. In Supplementary Material, we present additional simulation results with $n = 400$, which indicate better performance of the proposed method with a larger sample size while delivering similar observations to those from Tables 2–4.

We also perform simulation studies to evaluate the robustness of the proposed estimators. Specifically, we examine the proposed estimators of $\beta_0(\tau)$ and r_0 when the association model (2) is misspecified. We generated data from the set-ups S2.C and S2.F listed in Table 1, where the Clayton's copula and the Frank's copula hold respectively. For each set-up, we implemented the proposed method separately by assuming the Clayton's copula and the Frank's copula. Figure 1 presents the empirical averages of $\hat{\beta}(\tau)$ under a correctly specified copula model and an incorrectly specified copula model. For comparison, we also plot the empirical averages of Peng and Huang (2008)'s estimates obtained by naively treating T_2 as independent censoring. Not surprisingly, the naive estimator yields large biases for $\tau \in [0.1, 0.65]$ in all scenarios. When the association model (2) is correctly specified, the empirical averages of $\hat{\beta}(\tau)$ are close to the true regression quantiles. This agrees with the results in Table 2. With a misspecified association structure, the proposed estimator of $\beta_0(\tau)$ exhibits quite small deviations from the true coefficients. For example, the biases of $\hat{\beta}(\tau)$, when we incorrectly adopt Frank's copula in set-up S2.C, are nearly negligible. We have very similar observations for the estimation of the association parameter. The empirical average of the estimated Kendall's tau coefficient, when Frank's copula is incorrectly assumed in set-up S2.C, equals 0.475, and that when we assume Clayton's copula in set-up S2.F equals 0.472. Both empirical averages are not far from the true Kendall's coefficient 0.5. Overall, our simulations suggest that the proposed estimation of $\beta_0(\tau)$ and r_0 is quite robust to the misspecification of the parametric form of the copula model. The empirical results from our robustness study agree with previous reports in competing risks literature. For example, the marginal regression analysis proposed by Chen (2010) was found to be robust to the specification of the functional form of the copula when the level of association is correctly assumed. Compared to competing risk methods such as those by Zheng and Klein (1995) and Chen (2010), the proposed method only requires the assumption on the parametric form of the copula while leaving the copula parameter unspecified. Consequently, our robustness result entails a better utility of the proposed method than the sensitive analysis permitted by competing risk data.

4. A Real Example

We illustrate the proposed method via an application to the bone marrow transplant (BMT) example, where data were collected on patients with acute leukemia following allogeneic bone marrow transplantation (Copelan et al., 1991). One intermediate endpoint of interest is the development of chronic graft-versus-host disease (GVHD), a common complication following transplantation. The GVHD endpoint was subject to potentially dependent censoring due to death. With the origin at transplantation, we let T_1 denote time to the GVHD endpoint and T_2 denote time to death, while C represents time to end of follow-up, which is not fixed but random. For the 137 subjects in this dataset, 61 (44.5%) were observed to experience chronic GVHD, 81 (59%) died during the follow-up period, among

which 52 patients (38%) died before the onset of chronic GVHD. The values of P_{00} , P_{10} , P_{01} and P_{11} defined in Table 1 are 17.5%, 38.0%, 23.4%, 21.2% respectively. Based on disease type, patients were classified into three groups, referred to as acute lymphoblastic leukemia (ALL), acute myelocytic leukemia (AML) low risk and AML high risk, respectively. The recorded data also include patients' age at the time of study enrollment. In this analysis, we are interested in assessing the effects of disease type and age on the development of GVHD. The covariate vector $\mathbf{Z} = (1, Z_1, Z_2, Z_3)$, where Z_1 and Z_2 are indicator variables of the AML low risk group and AML high risk group, respectively, and Z_3 represents patients' age.

We fitted models (1)–(3) to the BMT data, adopting Frank's copula and setting $(\tau_a, \tau_b) = (0.05, 0.55)$. Standard error estimates and confidence intervals were obtained based on 400 bootstrap resamples. The estimated association parameter r equals 4.65. The corresponding Kendall's tau equals 0.43, with a percentile-based 95% interval of $(-0.04, 0.60)$ and a 95% Wald-type interval of $(0.05, 0.63)$. This result provides some evidence for a moderate positive association between GVHD and death, conditionally upon disease type and age. In Figure 2, we present the fitted regression coefficients, coupled with percentile-based 95% confidence intervals. Figure 2 suggests that patients in AML low risk group may have overall significantly slower progression to GVHD compared to patients in ALL group. The benefit in the timing of GVHD associated with being in AML high risk group versus ALL group is much less evident, with the 95% confidence intervals excluding 0 only for $\tau < 0.15$. In addition, the estimation of the coefficients for Z_3 suggests little age effect on the timing of GVHD development. For comparison purpose, we also present in Figure 2 Peng and Huang (2008)'s estimator $\hat{\beta}_N(\tau)$ which treats death as independent censoring. We observe that the intercept term of $\hat{\beta}_N(\tau)$ tends to be higher than that of $\hat{\beta}(\tau)$, which agrees with our observation in Figure 1 from the simulation studies that $\hat{\beta}_N(\tau)$ may tend to overestimate the intercept term when T_1 and T_2 are positively associated. There are rather large discrepancies in the estimated coefficients for Z_2 and Z_3 . For example, naively using $\hat{\beta}_N(\tau)$ may lead to an impression that younger patients have more rapid progression to GVHD compared to older patients, which seems counterintuitive. Compared to the naive analysis based on $\hat{\beta}_N(\tau)$, the proposed method may provide us higher confidence on the results by appropriately accounting for to the dependence between time GVHD and time to death.

We next conducted second-stage inferences on covariate effects, setting $(l, u) = (0.05, 0.55)$. The estimated trimmed mean effect of Z_2 (AML low risk versus ALL) equals 0.65, with an estimated standard error of 0.32. The percentile-based 95% confidence interval is $(0.26, 1.44)$. These results confirm the beneficial effect of being in AML low risk group (versus ALL group) observed in Figure 2. For Z_2 (AML high risk versus ALL), the trimmed mean effect estimate is 0.17, with an estimated standard error of 0.14 and a percentile-based 95% confidence interval of $(-0.03, 0.50)$. This indicates little evidence to support the difference in time to GVHD between AML high risk group and ALL group. The trimmed mean estimate for Z_3 (Age) equals -0.003 , with an estimated standard error of 0.009 and a percentile-based 95% confidence interval of $(-0.020, 0.016)$. This suggests a non-significant age effect on GVHD endpoint. Regarding the constancy of covariate effects, we obtained

$\mathcal{T}_{2,2} = -0.22$ for Z_1 , coupled with a percentile-based 95% confidence interval of $(-0.75, -0.04)$. This confirms our visual impression from Figure 2 and implies that the difference between AML low risk group and ALL group may be more pronounced in patients who are less prone to GVHD. For Z_2 and Z_3 , the test statistics $\mathcal{T}_{2,3}$ and $\mathcal{T}_{2,4}$ equal 0.03 and 0.002, respectively, both with percentile-based confidence intervals covering 0. Therefore, constant effects may be adequate for AML high risk (versus ALL) and Age.

Finally, we present the estimated quantiles of time to GVHD endpoint for patients with different disease types in Figure 3, with age set at its mean of 28.4 years. These curves provide useful alternative survival summaries by using quantiles of time to the endpoint of interest, which have direct physical interpretations. We observe that the differences in estimated quantiles based on $\hat{\beta}(\tau)$ and $\hat{\beta}_N(\tau)$ seem to increase with τ . For example, for ALL, AML low and AML high risk groups, the differences between the estimated 40th quantiles based on these two estimators are 0.9, 2.3 and 1.7 months respectively. Further unreported analyses suggest that the discrepancies in estimated quantiles increase with patients' age. It is also observed from Figure 3 that estimated quantiles of time to GVHD for the AML low risk group are much larger than those for the ALL group and AML high risk group, and this beneficial effect of being in AML low risk group appears more apparent when τ increases. By comparison, the estimated quantiles for AML high risks patients are quite similar to those for ALL patients. These observations are consistent with the findings from our second-stage inferences

5. Remarks

In this paper, we propose a quantile regression method that can properly accommodate dependent censoring situations that fall into the paradigm of semi-competing risks. Our method offers a useful tool for investigating nonterminating endpoints that often arise in clinical follow-up studies and their relationship with subsequent competing endpoints. The *net* quantile inference pursued in this paper is useful when covariate effects with the removal of dependent censoring are of substantive relevance.

It is worth noting that the inequalities (4) and (5) adopted for estimating model (1) entail a rather comprehensive use of semi-competing risks information, which is not limited to the diagonal line of the joint distribution of (T_1, T_2) as in Peng and Fine (2007b). While formal efficiency comparisons between these two approaches are tricky due to their distinct modeling strategies, some simulation studies (unreported) show 25-35% more efficient coefficient estimation from the new method compared to that from Peng and Fine (2007b).

We impose assumptions on the dependent censoring scheme through a general class of copula models. As noted in the paper, this is necessary for addressing the identifiability issue inherited with the dependent censoring problem. Simulations have shown that the proposed estimators are quite robust to misspecification of the parametric form of the copula model. This robustness feature is expected to enhance the practical utility of the proposed method.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors are grateful to the editor, the associate Editor, and the two referees for many helpful comments. This research has been supported by National Science Foundation under Grant Number DMS-1007660 and National Institute of Health under Award Number R01HL 113548.

APPENDIX A: REGULARITY CONDITIONS

For a vector $\mathbf{x}$, let $\mathbf{x}^{\otimes 2}$ denote $\mathbf{x}\mathbf{x}^T$, and $\|\mathbf{x}\|$ denote the Euclidean norm of $\mathbf{x}$. For a random variable T , let $f_T(\cdot|z)$ denote its conditional density function given $\mathbf{Z} = z$. We define $s(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}, v) = ES_n(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}, v)$, $w(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}, v) = EW_n(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{r}, v)$, and $w(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}) = \int_{\tau_a}^{\tau_b} w(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}, \tau) d\tau$. Let $B_b(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}, v) = s(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}, v) / \boldsymbol{\beta}(v)$, $B_r(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}, v) = s(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}, v) / \mathbf{r}$, $L_b(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}, v) = w(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}, v) / \boldsymbol{\beta}(v)$, and $L_r(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}) = w(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{r}) / \mathbf{r}$. Next, define $\boldsymbol{\mu}(a) = E[\mathbf{Z}I\{Y = \exp(\mathbf{Z}^T a), \eta = 1\}]$ as in Peng and Huang (2008). For $d > 0$, let

$$\mathcal{A}(d) = \left\{ a \in R^{p+1} : \inf_{\tau \in (0, \tau_{U,2}]} \|\boldsymbol{\mu}(a) - \boldsymbol{\mu}(\boldsymbol{\alpha}_0(\tau))\| \leq d \right\} \text{ and}$$

$$\mathcal{B}(d) = \left\{ \mathbf{r} \in R^{q+1} : \|\mathbf{r} - \mathbf{r}_0\| \leq d \right\}. \text{ Define}$$

$$\bar{\mathcal{B}}(d) = \left\{ \boldsymbol{\beta} \in R^{p+1} : \max_{\tau \in (\nu_1, \tau_{U,1}]} \|\boldsymbol{\beta}(\tau) - \boldsymbol{\beta}_0(\tau)\| \leq d \right\}. \text{ Let } d_A, d_B \text{ and } d_R \text{ be positive constants that determine the spans of the neighborhoods of } \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, \mathbf{r}_0 \text{ respectively. For}$$

presentation simplicity, we exclude the indicator functions $I\{\mathbf{Z}_i^T \boldsymbol{\beta}(\tau) \leq \mathbf{Z}_i^T \boldsymbol{\alpha}(\tau_{U,2})\}$ in S_n and $I\{\log t < \mathbf{Z}_i^T \boldsymbol{\alpha}(\tau_{U,2})\}$ in W_n respectively, corresponding to the situation with $\tau_{U,2}$ close to 1. This incurs little essential difference in asymptotic arguments.

We require the following regularity conditions:

C1 The covariate space $\mathcal{Z}$ is bounded, i.e., $\sup_i \|\mathbf{Z}_i\| < \infty$.

C2 (i) The regularity conditions in Peng and Huang (2008) hold for $(Y, \eta, \mathbf{Z})$ and $\boldsymbol{\alpha}_0(\tau)$, $\tau \in (0, \tau_{U,2}]$. (ii) For $\mathbf{B}_a(a) = \boldsymbol{\mu}(a) / a$, there exists a constant C_F , such that each component of $\|f_{T_2}\{\exp(\mathbf{z}^T a) | \mathbf{z}\} \exp(\mathbf{z}^T a) \times \mathbf{B}_a(a)^{-1}\|$ is bounded by C_F uniformly in $\mathbf{z} \in \mathcal{Z}$ and $a \in \mathcal{A}(d_A)$, where d_A positive constant. (iii) Define

$$-\frac{1}{r_f} \log \left(1 + \frac{[\exp\{-r_f \Pr(T_1 > s | \mathbf{Z})\} - 1] \times [\exp\{-r_f \Pr(T_2 > t | \mathbf{Z})\} - 1]}{\exp(-r_f) - 1} \right)$$

where $d_A(u, v, \theta) = K_A(u, v, \theta) / v$ and $d_B(u, v, \theta) = K_B(u, v, \theta) / v$. For $\mathbf{Z}$ containing continuous covariates, $\mathbf{V}_A(a, v)$ and $\mathbf{V}_B(a)$ are differentiable with respect to a , and every component of $[E\{\mathbf{V}_A(a, v) / a\} \mathbf{B}_a(a)^{-1}]$ and $[E\{\mathbf{V}_B(a) / a\} \mathbf{B}_a(a)^{-1}]$ are bounded uniformly for $a \in \mathcal{A}(d_A)$ and $v \in (0, \tau_{U,1}]$.

C3 (i) Each component of $s\{\beta_0, \alpha_0, r_0, \tau\}$ and $w\{\beta_0, \alpha_0, r_0, \tau\}$ is a Lipschitz function of τ when $\tau \in (0, \tau_{U,1}]$, (ii) let $d(u, v, \theta) = \Psi(u, v, \theta) / \theta, \|d(u, v, g(z^T r_0))\| g'(z^T r_0)$ is uniformly bounded for $z \in \mathcal{Z}, u \leq \tau_{U,1}$ and $v \leq \tau_{U,2}$, (iii) the conditional density functions $f_X(t|z), f_Y(t|z), f_{T_1}(t|z)$ and $f_{T_2}(t|z)$ are bounded above uniformly in t and $z \in \mathcal{Z}$.

C4 (i) $k_b \equiv \inf_{\tau \in (\tau_L, \tau_{U,1}]} \text{eigmin}\{B_b(\beta_0, \alpha_0, r_0, \tau)\} > 0$, where $\text{eigmin}(\cdot)$ denotes the minimum eigenvalue of a matrix. (ii) For any fixed $r \in \mathcal{R}(d_R)$, $\tilde{\beta}(\tau; r)$ is the unique solution to $s\{\beta_0, \alpha_0, r, \tau\} = 0$ for $\beta \in \bar{\mathcal{B}}(d_B)$ and $k_r \equiv \inf_{r \in \mathcal{R}(d_R)} \text{eigmin}\left[\partial w\{\tilde{\beta}(r), \alpha_0, r\} / \partial r\right] > 0$. (iii) Every component of $L_b(\beta, \alpha_0, r_0, \tau) \times B_b(\beta, \alpha_0, r_0, \tau)^{-1}$ is uniformly bounded for $\beta \in \bar{\mathcal{B}}(d_B)$.

C5 (i) For any $z \in \mathcal{Z}$, we have $z^T \beta_0(\tau_{U,2})$. (ii) Both $\inf_{z \in \mathcal{Z}} \Pr\left[C > \exp\left\{Z^T \beta_0(\tau_{U,1})\right\} | Z\right]$ and $\inf_{z \in \mathcal{Z}} \Pr\left[C > \exp\left\{Z^T \alpha_0(\tau_{U,2})\right\} | Z\right]$ are bounded away from 0.

Condition C1 requires bounded covariates and is often met in practice. Condition C2 imposes mild assumptions on the dependent censoring time T_2 . Specifically, it ensures that $\int_0^{\tau_{U,2}} I\left[t \geq \exp\left\{Z^T \hat{\alpha}(u)\right\}\right] du$ serves an appropriate estimator of $F_2(t|Z) \wedge \tau_{U,2}$, and that $s(\beta, \hat{\alpha}, r, \tau)$ preserves the nice asymptotic properties of $\hat{\alpha}(\cdot)$ shown in Peng and Huang (2008). By condition C3, $s\{\beta_0(\tau), \alpha_0, r, \tau\}$ and $w\{\beta_0(\tau), \alpha_0, r, \tau\}$ are smooth in both τ and r , and the conditional density functions of X, Y, T_1 and T_2 are uniformly bounded. Condition C4, coupled with the boundary constraints in C5, entails the identifiability of the proposed estimator in a neighborhood of β_0 and r_0 . Specifically, C4(i) and C4(ii) requires that the derivative of $s(\beta, \alpha, r, \tau)$ over β and that of $w(\tilde{\beta}, \alpha, r, \tau)$ over r are bounded from below at the true parameters, and C4(iii) ensures that the derivative of $\tilde{\beta}(\tau; r)$ over r is bounded at $r = r_0$. Note that C5(i) can be removed when equation (8) is adopted.

Our following proofs utilizes the asymptotic properties of Peng and Huang (2008)'s estimator $\hat{\alpha}(\cdot)$. It is worth pointing out that the asymptotic results on $\tilde{\beta}(\cdot)$ remain valid when other modeling strategies are adopted for T_2 , provided that the strategy entail reasonable estimates of $F_2(t|Z)$. The proof can be carried out following the lines of Appendix B and C in the following, and the detailed arguments are omitted here.

APPENDIX B: PROOF OF THEOREM 1

Lemma 1. For $\mu(a = EZI\{Y > \exp(Z^T a), \eta = 1\})$, we have

$$\begin{aligned} V_A(a, \tau) &= -Z \left[I\left\{\log Y > Z^T \beta_0(\tau)\right\} d_A\left\{\tau, F_2\left[\exp\left\{Z^T \beta_0(\tau)\right\} | Z\right], g\left(Z^T r_0\right)\right\} \times I\left\{Z^T \beta_0(\tau) \geq Z^T a\right\} \right] \\ V_B(a) &= -Z \left\{ \int_{\tau_a}^{\tau_b} \int_{t \in (0, \infty)} I\left\{Z^T \beta_0(\tau) \leq \log t < \log Y\right\} \times d_B\left\{\tau, F_2(t|Z), g\left(Z^T r_0\right)\right\} I\left\{\log t \geq \exp\left(Z^T a\right)\right\} d\tau dt \right\}, \end{aligned}$$

Proof of Lemma 1. By regularity condition C2(ii) and Taylor expansion, we have

$$\sup_{z \in \mathcal{Z}, t} \left| \int_0^{\tau_{U,2}} I \left[t \geq \exp \left\{ z^T \hat{\alpha}(u) \right\} \right] du - \int_0^{\tau_{U,2}} I \left[t \geq \exp \left\{ z^T \alpha_0(u) \right\} \right] du \right| \leq 2C_F \sup_{\tau \in (0, \tau_{U,2})} \left\| \mu \left\{ \hat{\alpha}(\tau) \right\} - \mu \left\{ \alpha_0(\tau) \right\} \right\|.$$

where $F_2[\exp\{z^T \alpha_0(\tau)\}|z] = \tau$. Let $\varepsilon_F = \sup_{z \in \mathcal{Z}, \tau \in (0, \tau_{U,2})} |F_2[\exp\{z^T \hat{\alpha}(\tau)\}|z] - \tau|$, we then have $\varepsilon_F \leq C_F \sup_{\tau \in (0, \tau_{U,2})} \left\| \mu \left\{ \hat{\alpha}(\tau) \right\} - \mu \left\{ \alpha_0(\tau) \right\} \right\|$.

Note that

$$\begin{aligned} \sup_{z \in \mathcal{Z}, t} \left| \int_0^{\tau_{U,2}} I \left[t \geq \exp \left\{ z^T \hat{\alpha}(u) \right\} \right] du - \int_0^{\tau_{U,2}} I \left[t \geq \exp \left\{ z^T \alpha_0(u) \right\} \right] du \right| &\leq \sup_{z \in \mathcal{Z}, t} \int_0^{\tau_{U,2}} I \left\{ F_2(t|z) \in [F_2[\exp\{z^T \hat{\alpha}(u)\}|z] - \varepsilon_F, F_2[\exp\{z^T \hat{\alpha}(u)\}|z] + \varepsilon_F] \right\} du \\ &= \sup_{z \in \mathcal{Z}, t} \int_0^{\tau_{U,2}} I \left\{ u \in [F_2(t|z) - \varepsilon_F, F_2(t|Bz) + \varepsilon_F] \right\} du \leq 2\varepsilon_F. \end{aligned}$$

where $\vee$ is the maximum operator. This immediately completes the proof of Lemma 1.

Proof of Theorem 1. For any fixed θ , the copula function $\Psi(u, v, \theta)$ satisfies the Lipschitz condition in u and v (Nelsen, 2006). Hence $K_A(u, v, \theta)$ and $K_B(u, v, \theta)$ are both Lipschitz continuous in v when $v \in \tau_{U,2}$. By Lemma 1 and the fact that

$\sup_{\tau \in (0, \tau_{U,2})} \left\| \mu \left\{ \hat{\alpha}(\tau) \right\} - \mu \left\{ \alpha_0(\tau) \right\} \right\| \xrightarrow{P} 0$ (Peng and Huang, 2008), we can show with Taylor expansion that

$$\sup_{z \in \mathcal{Z}, \mathbf{r}, \mathbf{b}, \tau} \left| K_A \left\{ \tau, \int_0^{\tau_{U,2}} I \left\{ z^T \mathbf{b} \geq z^T \hat{\alpha}(u) \right\} du, g(z^T \mathbf{r}) \right\} - K_A \left\{ \tau, \int_0^{\tau_{U,2}} I \left\{ z^T \mathbf{b} \geq z^T \alpha_0(u) \right\} du, g(z^T \mathbf{r}) \right\} \right| \xrightarrow{P} 0.$$

It follows that

$$\sup_{\beta, \mathbf{r}, \tau} \left\| s(\beta, \hat{\alpha}, \mathbf{r}, \tau) - s(\beta, \alpha_0, \mathbf{r}, \tau) \right\| \xrightarrow{P} 0. \quad (\text{A1})$$

Next, we claim that $\mathcal{G}_1 = \left\{ Z_i P_i(\beta, \alpha, \mathbf{r}, \tau) : Z_i \in \mathcal{Z}, \beta, \alpha \in \mathbf{R}^{p+1}, \mathbf{r} \in \mathbf{R}^{q+1}, \tau \in (0, 1) \right\}$ is Donsker thus Glivenko Cantelli. This follows by noting that the class of indicator functions is a VC-class, and by using the permanence properties of the Donsker class (Van der Vaart and Wellner, 1996). Therefore, Glivenko-Cantelli theorem gives

$$\sup_{\beta, \alpha, \mathbf{r}, \tau} \left\| S_n(\beta, \alpha, \mathbf{r}, \tau) - s(\beta, \alpha, \mathbf{r}, \tau) \right\| = o_p(1). \quad (\text{A2})$$

Since $\tilde{\beta}(\tau; \mathbf{r})$ is the solution to $s(\tilde{\beta}, \alpha_0, \mathbf{r}, \tau) = 0$, we can use simple manipulations and get

$$\begin{aligned}
& \|s \{ \hat{\beta}(r), \alpha_0, r, \tau \} \\
& \quad - s \{ \tilde{\beta}(r), \alpha_0, r, \tau \} \| \\
& = \|s \{ \hat{\beta}(r), \alpha_0, r, \tau \} \| \leq \|s \{ \hat{\beta}(r), \hat{\alpha}, r, \tau \} \| \quad (A3) \\
& + \|s \{ \hat{\beta}(r), \alpha_0, r, \tau \} \\
& \quad - s \{ \tilde{\beta}(r), \hat{\alpha}, r, \tau \} \|.
\end{aligned}$$

Combining (A2) and the fact that $\sup_{r, \tau \in [\nu_1, \tau_{U,1}]} \|S_n \{ \hat{\beta}(r), \hat{\alpha}, r, \tau \} \| = o_p(1)$, we get

$\sup_{r, \tau \in [\nu_1, \tau_{U,1}]} \|s \{ \hat{\beta}(r), \hat{\alpha}, r, \tau \} \| \xrightarrow{p} 0$. This, coupled with (A1) and (A3), implies

$$\sup_{r, \tau \in [\nu_1, \tau_{U,1}]} \|s \{ \hat{\beta}(r), \alpha_0, r, \tau \} - s \{ \tilde{\beta}(r), \alpha_0, r, \tau \} \| \xrightarrow{p} 0. \quad (A4)$$

By condition C4(i), the inverse function theorem implies that there exists $\hat{\beta}(\tau, r) \in \bar{\mathcal{B}}(d_B)$.

More-over, Taylor expansion of $s \{ \tilde{\beta}(r), \alpha_0, r, \tau \}$ around $\tilde{\beta}(\tau; r)$ gives

$$\sup_{r \in \mathcal{R}(d_R), \tau \in [\nu_1, \tau_{U,1}]} \| \hat{\beta}(\tau; r) - \tilde{\beta}(\tau; r) \| \leq \sup_{r \in \mathcal{R}(d_R), \tau \in [\nu_1, \tau_{U,1}]} \frac{1}{k_b} \|s \{ \hat{\beta}(r), \alpha_0, r, \tau \} - s \{ \tilde{\beta}(r), \alpha_0, r, \tau \} \| \xrightarrow{p} 0. \quad (A5)$$

This fact, combined with the Glivenko-Cantelli theorem on $W_n(\beta, \alpha, r)$, implies that

$\sup_{r \in \mathcal{R}(d_R)} \|W_n(\hat{\alpha}, r) - \tilde{w}(\hat{\alpha}, r)\| \xrightarrow{p} 0$, where $W_n(\alpha, r) = W_n \{ \hat{\beta}(r), \alpha, r \}$ and

$\tilde{w}(\alpha, r) = w \{ \tilde{\beta}(r), \alpha, r \}$. Similarly, we can use Lemma 1 and the uniform consistency of

$\mu \{ \hat{\alpha}(\tau) \}, \tau \in (0, \tau_{U,2}]$ to show $\sup_{r \in \mathcal{R}(d_R)} \| \tilde{w}(\hat{\alpha}, r) - \tilde{w}(\alpha_0, r) \| \xrightarrow{p} 0$. It follows that

$$\sup_{r \in \mathcal{R}(d_R)} \|W_n(\hat{\alpha}, r) - \tilde{w}(\alpha_0, r)\| \leq \sup_{r \in \mathcal{R}(d_R)} [\|W_n(\hat{\alpha}, r) - \tilde{w}(\hat{\alpha}, r)\| + \|\tilde{w}(\hat{\alpha}, r) - \tilde{w}(\alpha_0, r)\|] \xrightarrow{p} 0. \quad (A6)$$

Therefore, we can see $w(\tilde{\alpha}_0, r) = o_p(1)$ from $W_n(\hat{\alpha}, \hat{r}) = o_p(1)$. By regularity conditions C4(ii), $w(\tilde{\alpha}_0, r_0) = 0$ and $\|\hat{r} - r_0\| \leq k_r^{-1} \|\tilde{w}(\alpha_0, \hat{r}) - \tilde{w}(\alpha_0, r_0)\|$. The consistency of $\hat{r}$ follows.

To show the uniform consistency of $\hat{\beta}(\cdot)$, we first need to show the partial derivative of $\tilde{\beta}(\tau; r)$ with respect to r is bounded at $r = r_0$. Taking partial derivative $\partial / \partial r$ on both sides of $s \{ \tilde{\beta}(r), \alpha_0, r, \tau \} = 0$, we get

$$\frac{\partial \tilde{\beta}(\tau; r)}{\partial r} \Big|_{r=r_0} = -B_b(\beta_0, \alpha_0, r_0, \tau)^{-1} B_r(\beta_0, \alpha_0, r_0, \tau). \quad (A7)$$

By condition C3(ii) and C4(i), $\frac{\partial \tilde{\beta}(\tau; r)}{\partial r}$ is uniformly bounded for $\tau \in [\nu_1, \tau_{U,1}]$, and thus

$$\sup_{\tau \in [\nu_1, \tau_{U,1}]} \|\tilde{\beta}(\tau; \hat{r}) - \tilde{\beta}(\tau; r_0)\| = o_p(1) \quad (\text{A8})$$

follows from the consistency of $\hat{r}$. Based on (A5) and (A8), the following inequality,

$$\|\hat{\beta}(\tau) - \beta_0(\tau)\| = \|\hat{\beta}(\tau; \hat{r}) - \tilde{\beta}(\tau; r_0)\| \leq \|\hat{\beta}(\tau; \hat{r}) - \tilde{\beta}(\tau; \hat{r})\| + \|\tilde{\beta}(\tau; \hat{r}) - \tilde{\beta}(\tau; r_0)\|,$$

implies $\sup_{\tau \in [\nu_1, \tau_{U,1}]} \|\hat{\beta}(\tau) - \beta_0(\tau)\| \xrightarrow{p} 0$. This completes the proof of Theorem 1.

APPENDIX C: PROOF OF THEOREM 2

LEMMA 2. For any sequence $\{\bar{\beta}_n(\tau), \tau \in [\nu_1, \tau_{U,1}]\}_{n=1}^{\infty} \in \bar{\mathcal{B}}(d_B)$ satisfying

$$\sup_{\tau \in [\nu_1, \tau_{U,1}]} \|s(\bar{\beta}_n, \alpha_0, r_0, \tau) - s(\beta_0, \alpha_0, r_0, \tau)\| \xrightarrow{p} 0, \text{ we have}$$

$$\sup_{\tau \in [\nu_1, \tau_{U,1}]} n^{1/2} \|\{S_n(\bar{\beta}_n, \alpha_0, r_0, \tau) - S_n(\beta_0, \alpha_0, r_0, \tau)\} - \{s(\bar{\beta}_n, \alpha_0, r_0, \tau) - s(\beta_0, \alpha_0, r_0, \tau)\}\| \xrightarrow{p} 0.$$

Similarly, for any sequence $\{\bar{\alpha}_n(\tau), \tau \in (0, \tau_{U,2}]\}_{n=1}^{\infty}$ satisfying

$$\sup_{\tau \in (0, \tau_{U,2}]} \|\mu\{\bar{\alpha}_n(\tau)\} - \mu\{\alpha_0(\tau)\}\| \xrightarrow{p} 0,$$

$$\sup_{\tau \in [\nu_1, \tau_{U,1}]} n^{1/2} \|\{S_n(\beta_0, \bar{\alpha}_n, r_0, \tau) - S_n(\beta_0, \alpha_0, r_0, \tau)\} - \{s(\beta_0, \bar{\alpha}_n, r_0, \tau) - s(\beta_0, \alpha_0, r_0, \tau)\}\| \xrightarrow{p} 0.$$

Proof of Lemma 2. Define $\sigma_B^2(\beta) = \sup_{\tau \in [\nu_1, \tau_{U,1}]} \text{Var}\{P_i(\beta, \alpha_0, r_0, \tau) - P_i(\beta_0, \alpha_0, r_0, \tau)\}$,

Following the arguments of Lai and Ying (1988) and Peng and Huang (2008), it is sufficient

for the first part to hold if $\sigma_B^2(\bar{\beta}_n) \xrightarrow{p} 0$. By condition C4(i)–(iii), one can use Taylor

expansion on $s(\bar{\beta}_n, \alpha_0, r_0, \tau)$ around $\beta_0(\tau)$ to show $\sup_{\tau \in [\nu_1, \tau_{U,1}]} \|\bar{\beta}_n(\tau) - \beta_0(\tau)\| \xrightarrow{p} 0$.

Furthermore, the conditional density functions $f_X(t|z)$, $f_Y(t|z)$ and $f_{T_2}(t|z)$ are bounded above uniformly in t and z , and $K_A(u, v, \theta)$ is Lipschitz continuous in v . These facts allow us to

mimic lemma B.1. in Peng and Huang (2008) and get $\sigma_B^2(\bar{\beta}_n) \xrightarrow{p} 0$.

Similarly, define $\sigma_A^2(\alpha) = \sup_{\tau \in [\nu_1, \tau_{U,1}]} \text{Var}\{P_i(\beta_0, \alpha, r_0, \tau) - P_i(\beta_0, \alpha_0, r_0, \tau)\}$, a

sufficient condition for the second part is $\sigma_A^2(\bar{\alpha}_n) \xrightarrow{p} 0$, which follows directly from Lemma 1 and the Lipschitz continuity of $K_A(u, v, \theta)$ in v . This completes the proof of Lemma 2.

$$\text{Weak Convergence } n^{1/2} [s\{\hat{\beta}(r_0), \alpha_0, r_0, \tau\} - s(\beta_0, \alpha_0, r_0, \tau)]$$

By simple manipulations, we can show $n^{1/2} [S_n \{\hat{\beta}(r_0), \hat{\alpha}, r_0, \tau\} - S_n(\beta_0, \alpha_0, r_0, \tau)]$ equals

$$n^{1/2} [S_n(\beta_0, \hat{\alpha}, r_0, \tau) - S_n(\beta_0, \alpha_0, r_0, \tau)] + n^{1/2} [S_n \{\hat{\beta}(r_0), \alpha_0, r_0, \tau\} - S_n(\beta, \alpha_0, r_0, \tau)] + \varepsilon_S \quad (\text{A9})$$

where

$$\varepsilon_S = n^{1/2} [S_n \{\hat{\beta}(r_0), \hat{\alpha}, r_0, \tau\} - S_n(\beta_0, \hat{\alpha}, r_0, \tau)] - n^{1/2} [S_n \{\hat{\beta}(r_0), \alpha_0, r_0, \tau\} - S_n(\beta_0, \alpha_0, r_0, \tau)]$$

can be shown to be $o_p^{\tau \in [\nu_1, \tau_{U,1}]}(1)$ and asymptotically negligible. Based on the uniform

convergence of $s \{\hat{\beta}(r_0), \alpha_0, r_0, \tau\}$ to $s(\beta(r_0), \alpha_0, r_0, \tau)$ for $\tau \in [\nu_1, \tau_{U,1}]$ in (A4), and the uniform convergence of $\mu \{\hat{\alpha}(\tau)\}$ to $\mu \{\alpha_0(\tau)\}$ for $\tau \in (0, \tau_{U,2}]$ in Peng and Huang (2008), we can use Lemma 2 to show

$$\begin{aligned} n^{1/2} [S_n(\beta_0, \hat{\alpha}, r_0, \tau) - S_n(\beta_0, \alpha_0, r_0, \tau)] &= n^{1/2} [s(\beta_0, \hat{\alpha}, r_0, \tau) - s(\beta_0, \alpha_0, r_0, \tau)] + o_p^{\tau \in [\nu_1, \tau_{U,1}]}(1), \\ n^{1/2} [S_n \{\hat{\beta}(r_0), \alpha_0, r_0, \tau\} - S_n(\beta_0, \alpha_0, r_0, \tau)] &= n^{1/2} [s \{\hat{\beta}(r_0), \alpha_0, r_0, \tau\} - s(\beta_0, \alpha_0, r_0, \tau)] + o_p^{\tau \in [\nu_1, \tau_{U,1}]}(1). \end{aligned} \quad (\text{A10})$$

Here and in the sequel, $o_p^s(1)$ means convergence to 0 in probability uniformly on set S .

Similarly, we use $o_p^s(n^{-1/2})$ to denote uniform root n convergence to 0 in probability on set S .

According to Peng and Huang (2008),

$$n^{1/2} [\mu \{\hat{\alpha}(u)\} - \mu \{\alpha_0(u)\}] = -n^{1/2} \sum_{i=1}^n \phi \left\{ Z_i R_i^Y(\alpha_0, u) \right\} + o_p^{u \in (0, \tau_{U,2}]}(1), \quad (\text{A11})$$

where

$$R_i^Y(\alpha, u) = I \left[Y_i \leq \exp \left\{ Z_i^T \alpha(u) \right\}, \eta = 1 \right] - \int_0^u I \left[Y_i \geq \exp \left\{ Z_i^T \alpha(t) \right\} \right] dH(t), \quad H(t) = -\log(1-t)$$

for $0 < t < 1$, and ϕ is a continuous linear map that involves product integration. The detailed form of ϕ can be found in Appendix B of Peng and Huang (2008).

By Lemma 1 and the Lipschitz continuity of $K_A(u, v, \theta)$ in v , we can use Taylor expansion to see

$$\begin{aligned} n^{1/2} \{s(\beta_0, \hat{\alpha}, r_0, \tau) - s(\beta_0, \alpha_0, r_0, \tau)\} \\ = -n^{1/2} E \left\{ Z^T I \left\{ \log Y > Z^T \beta_0(\tau) \right\} \times d_A \left\{ \tau, F_2 \left[\exp \left\{ Z^T \beta_0(\tau) \right\} | Z \right], g(Z^T r_0) \right\} \times \left[\int_0^{\tau_{U,2}} I \left\{ Z^T \beta_0(\tau) \geq Z^T \hat{\alpha}(u) \right\} du \right] \right\} \\ + o_p^{\tau \in [\nu_1, \tau_{U,1}]}(1), \end{aligned} \quad (\text{A12})$$

where $d_A(u, v, \theta) = K_A(u, v, \theta) / v$. In the following we will show that the left hand side (LHS) of (A12) has a uniform i.i.d. representation.

We first consider the one-sample case, where $\{\exp \{\hat{\alpha}(u)\}, u \in (0, \tau_{U,2}]\}$ is asymptotically equivalent to the quantile function of the Nelson-Aalen estimator (Peng and Huang, 2008) of

T_2 's distribution, denoted $\hat{F}_2^{NA}(\cdot)$. Specifically, it can be shown that

$\hat{F}_2^{NA}[\exp\{\hat{\alpha}(u)\}] = u + o_p\left(\int_0^{\tau_{U,2}} \left(n^{-1/2}\right) du\right)$. By this result and C5(i), we get

$$\int_0^{\tau_{U,2}} I\{\beta_0(\tau) \geq \hat{\alpha}(u)\} du = \int_0^{\tau_{U,2}} I\left(\hat{F}_2^{NA}[\exp\{\beta_0(\tau)\}] \geq u\right) du + o_p\left(n^{-1/2}\right) = \hat{F}_2^{NA}[\exp\{\beta_0(\tau)\}] + o_p\left(n^{-1/2}\right).$$

Since $n^{1/2}\{\hat{F}_2^{NA}(t) - F_2(t)\}$ is asymptotically Gaussian and can be written as

$n^{-1/2}\sum_{i=1}^n \pi_i(t) + o(1)$, where $\{\pi_i(t)\}_{i=1}^n$ are i.i.d. and form a Donsker's class with mean 0 (Kosorok, 2008), we can see the right hand side (RHS) of (A12) equals

$$-n^{-1/2}\sum_{i=1}^n d_A\{\tau, F_2[\exp\{\beta_0(\tau)\}], g(\mathbf{r}_0)\} \pi_i[\exp\{\beta_0(\tau)\}] Pr\{\log Y > \beta_0(\tau)\} + o_p^{\tau \in [\nu_1, \tau_{U,1}]}(1). \quad (A13)$$

The above arguments for the one sample case can be easily extended to the K-sample case, which covers situations where all covariates are discrete. Specifically, suppose

$\mathcal{Z} = \{z_1, z_2, \dots, z_K\}$ and $\mathbf{Z}_i = \mathbf{z}_k$ if and only if observation i belongs to the k_{th} group. We can show that RHS of (A12) equals

$$-n^{-1/2}\sum_{i=1}^n \sum_{k=1}^K \left[I(Z_i = z_k) z_k d_A\left\{\tau, F_2\left[\exp\left\{z_k^T \beta_0(\tau)\right\} | z_k\right], g\left(z_k^T \mathbf{r}_0\right)\right\} \pi_i\left[\exp\left\{z_k^T \beta_0(\tau)\right\}\right] \times Pr\left\{\log Y > z_k^T \beta_0(\tau) | z_k\right\} \right] + o_p^{\tau \in [\nu_1, \tau_{U,1}]}(1). \quad (A14)$$

By the boundedness of $d_A(\cdot)$, $\mathcal{Z}$ and $Pr\{\log Y > z_k^T \beta_0(\tau) | z_k\}$, we can show that

$-\sum_{k=1}^K z_k I(Z_i = z_k) d_A\left\{\tau, F_2\left[\exp\left\{z_k^T \beta_0(\tau)\right\} | z_k\right], g\left(z_k^T \mathbf{r}_0\right)\right\} \times Pr\left\{\log Y > z_k^T \beta_0(\tau) | z_k\right\} \pi_{k,i}\left[\exp\left\{z_k^T \beta_0(\tau)\right\}\right]$ is mean 0 and form a Donsker's class, as the Donsker property is preserved under Lipschitz transformations. Therefore, (A14) converges weakly to a mean 0 Gaussian process.

Next we consider the case when $\mathbf{Z}$ involves some continuous components. Define $\mathbf{v}_A(\mathbf{a}, \tau) = EV_A(\mathbf{a}, \tau)$. By C2(iii), $\mathbf{J}_A(\mathbf{a}, \tau) \equiv \mathbf{v}_A(\mathbf{a}, \tau)/\mathbf{a}$ exists, and each component of

$\mathbf{J}_A\{\alpha_0(u), \tau\} \mathbf{B}_a^{-1}\{\alpha_0(u)\}$ is uniformly bounded for any $u \in (0, \tau_{U,2})$ and $\tau \in [\nu_1, \tau_{U,1}]$.

Therefore, Taylor expansion, coupled with (A11) and (A12), shows that

$$\begin{aligned} n^{1/2}\{s(\beta_0, \hat{\alpha}, \mathbf{r}_0, \tau) - s(\beta_0, \alpha_0, \mathbf{r}_0, \tau)\} &= \int_0^{\tau_{U,2}} n^{1/2} [v_A\{\hat{\alpha}(u), \tau\} - v_A\{\alpha_0(u), \tau\}] du + o_p^{\tau \in [\nu_1, \tau_{U,1}]}(1) \\ &= -n^{-1/2} \sum_{i=1}^n \int_0^{\tau_{U,2}} \mathbf{J}_A\{\alpha_0(u), \tau\} \mathbf{B}_a^{-1}\{\alpha_0(u)\} \phi\{\mathbf{Z}_i R_i^Y(\alpha_0, u)\} du + o_p^{\tau \in [\nu_1, \tau_{U,1}]}(1). \end{aligned} \quad (A15)$$

It is seen from (A14) and (A15) that $n^{1/2}[s(\beta_0, \hat{\alpha}, \mathbf{r}_0, \tau) - s(\beta_0, \alpha_0, \mathbf{r}_0, \tau)]$ is asymptotically equivalent to a sum of i.i.d. influence functions, no matter whether $\mathbf{Z}$ includes continuous covariates or not. In the sequel we unify (A14) and (A15) by writing

$$n^{1/2} \{s(\beta_0, \hat{\alpha}, r_0, \tau) - s(\beta_0, \alpha_0, r_0, \tau)\} = n^{-1/2} \sum_{i=1}^n \rho_i(\beta_0, \alpha_0, r_0, \tau) + o_p^{\tau \in [\nu_1, \tau_{U,1}]}(1). \quad (A16)$$

Note that $S_n \{ \hat{\beta}(r_0), \hat{\alpha}, r_0, \tau \} = o_p^{\tau \in [\nu_1, \tau_{U,1}]}(n^{-1/2})$, (A9) and (A10) imply that

$$\begin{aligned} & n^{1/2} [s\{\hat{\beta}(r_0), \alpha_0, r_0, \tau\} - s(\beta_0, \alpha_0, r_0, \tau)] \\ &= -n^{1/2} S_n(\beta_0, \alpha_0, r_0, \tau) \\ & \quad - n^{1/2} \{s(\beta_0, \hat{\alpha}, r_0, \tau) - s(\beta_0, \alpha_0, r_0, \tau)\} \\ & \quad + o_p^{\tau \in [\nu_1, \tau_{U,1}]}(1). \end{aligned}$$

By (A16), we then get

$$n^{1/2} [s\{\hat{\beta}(r_0), \alpha_0, r_0, \tau\} - s(\beta_0, \alpha_0, r_0, \tau)] = n^{-1/2} \sum_{i=1}^n \chi_i(\beta_0, \alpha_0, r_0, \tau) + o_p^{\tau \in [\nu_1, \tau_{U,1}]}(1), \quad (A17)$$

where $\chi_i(\beta_0, \alpha_0, r_0, \tau) = -Z_i P_i(\beta_0, \alpha_0, r_0, \tau) - \rho_i(\beta_0, \alpha_0, r_0, \tau)$. This uniform i.i.d. representation implies the weak convergence of $n^{1/2} [s\{\hat{\beta}(r_0), \alpha_0, r_0, \tau\} - s(\beta_0, \alpha_0, r_0, \tau)]$ to a mean zero Gaussian process by an application of Donker theorem.

Weak Convergence of $n^{1/2} W_n \{ \hat{\beta}(r_0), \hat{\alpha}, r_0 \}$

Using similar arguments in (A9) and Lemma 2, we

$$\begin{aligned} n^{1/2} [W_n \{ \hat{\beta}(r_0), \hat{\alpha}, r_0 \} - W_n(\beta_0, \alpha_0, r_0)] &= n^{1/2} [W_n(\beta_0, \hat{\alpha}, r_0) - W_n(\beta_0, \alpha_0, r_0)] + n^{1/2} [W_n \{ \hat{\beta}(r_0), \alpha_0, r_0 \} - W_n(\beta_0, \\ &= n^{1/2} [w(\beta_0, \hat{\alpha}, r_0) - w(\beta_0, \alpha_0, r_0)] + n^{1/2} [w\{\hat{\beta}(r_0), \alpha_0, r_0\} - w(\beta_0, \alpha_0, r_0)] \end{aligned} \quad (A18)$$

where

$$\varepsilon_W = n^{1/2} [W_n \{ \hat{\beta}(r_0), \hat{\alpha}, r_0 \} - W_n(\beta_0, \hat{\alpha}, r_0)] - n^{1/2} [W_n \{ \hat{\beta}(r_0), \alpha_0, r_0 \} - W_n(\beta_0, \alpha_0, r_0)]$$

can also be shown to be $o_p(1)$. It is easy to see $n^{1/2} W_n(\beta_0, \alpha_0, r_0)$ converges weakly to a Gaussian process by the Donsker property. Moreover, we can follow the arguments for (A14) and (A15) to show

$$n^{1/2} \{w(\beta_0, \hat{\alpha}, r_0) - w(\beta_0, \alpha_0, r_0)\} = n^{-1/2} \sum_{i=1}^n \kappa_i(\beta_0, \alpha_0, r_0) + o_p(1). \quad (A19)$$

The i.i.d. influence functions $\kappa_i(\beta_0, \alpha_0, r_0)$ take the form

$$-\sum_{k=1}^K \int_{\tau_a}^{\tau_b} \int_{t \in (0, \infty)} [\bar{z}_k I(Z_i = z_k) I\{z_k^T \beta_0(\tau) < \log t\} Pr(Y > t | z_k) d_B \{\tau, F_2(t | z_k), g(z_k^T r_0)\} \times \pi_{k,i}(t) d\tau dt], \quad (A20)$$

in the K-sample case, and equals

$$-\int_0^{\tau_{U,2}} \mathbf{J}_B \{ \boldsymbol{\alpha}_0(u) \} \mathbf{B}_a^{-1} \{ \boldsymbol{\alpha}_0(u) \} \phi \left\{ \mathbf{Z}_i R_i^Y(\boldsymbol{\alpha}_0, u) \right\} du \quad (\text{A21})$$

when $\mathbf{Z}$ contains continuous components, where $\mathbf{J}_B(\mathbf{a}) \equiv d\mathbf{v}_B(\mathbf{a})/d\mathbf{a}$, and $\mathbf{v}_B(\mathbf{a}) = E\mathbf{V}_B(\mathbf{a})$. It then follows from the central limit theory and condition C2 that

$n^{1/2} \{ \mathbf{w}(\boldsymbol{\beta}_0, \hat{\boldsymbol{\alpha}}, \mathbf{r}_0) - \mathbf{w}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0) \}$ converges to a mean 0 normal distribution.

We also show by C4(iv) and a Taylor expansion that

$$\begin{aligned} & n^{1/2} \left[\mathbf{w}(\hat{\boldsymbol{\beta}}(\mathbf{r}_0), \boldsymbol{\alpha}_0, \mathbf{r}_0) - \mathbf{w}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0) \right] \\ &= \int_{\tau_a}^{\tau_b} \left[\mathbf{L}_b \{ \boldsymbol{\beta}_0(\tau), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} \mathbf{B}_b \{ \boldsymbol{\beta}_0(\tau), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \}^{-1} \times n^{1/2} \left[\mathbf{s} \{ \hat{\boldsymbol{\beta}}(\mathbf{r}_0), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} - \mathbf{s}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau) \right] \right] d\tau + o_p(1). \end{aligned} \quad (\text{A22})$$

Combining Equation (A17), (A18), (A19) and (A22), we then have

$$n^{1/2} \mathbf{W}_n \{ \hat{\boldsymbol{\beta}}(\mathbf{r}_0), \hat{\boldsymbol{\alpha}}, \mathbf{r}_0 \} = n^{-1/2} \sum_{i=1}^n \boldsymbol{\iota}_i(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0) + o_p(1), \quad (\text{A23})$$

where

$$\boldsymbol{\iota}_i(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0) = \bar{Z}_i Q_i(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0) + \boldsymbol{\kappa}_i(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0) + \int_{\tau_a}^{\tau_b} \mathbf{L}_b \{ \boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} \mathbf{B}_b \{ \boldsymbol{\beta}_0(\tau), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \}^{-1} \chi_i(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau)$$

. The weak convergence of $n^{1/2} \mathbf{W}_n \{ \hat{\boldsymbol{\beta}}(\mathbf{r}_0), \hat{\boldsymbol{\alpha}}, \mathbf{r}_0 \}$ follows from the Donsker Theorem.

Asymptotic Normality of $n^{1/2}(\mathbf{r} - \mathbf{r}_0)$ and Weak Convergence of $n^{1/2} \{ \hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau) \}$ Using

$\mathbf{s} \{ \tilde{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \tau \} = \mathbf{s}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau) = 0$, we can obtain the following inequality

$$\begin{aligned} & \| \mathbf{s} \{ \hat{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} \\ & - \mathbf{s} \{ \hat{\boldsymbol{\beta}}(\mathbf{r}_0), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} \| \leq \left[\| \mathbf{s} \{ \hat{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} - \mathbf{s} \{ \hat{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \hat{\mathbf{r}}, \tau \} \| + \| \mathbf{s} \{ \hat{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \hat{\mathbf{r}}, \tau \} - \mathbf{s} \{ \tilde{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \hat{\mathbf{r}}, \tau \} \| + \| \mathbf{s} \{ \tilde{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \hat{\mathbf{r}}, \tau \} - \mathbf{s} \{ \tilde{\boldsymbol{\beta}}(\mathbf{r}_0), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} \| \right] \end{aligned}$$

By C3(ii), $\hat{\mathbf{r}} \xrightarrow{p} \mathbf{r}_0$ implies $\sup_{\tau \in [\nu_1, \tau_{U,1}]} \| \mathbf{s} \{ \hat{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} - \mathbf{s} \{ \hat{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \hat{\mathbf{r}}, \tau \} \| \xrightarrow{p} 0$.

This, coupled with the inequality above and (A4), implies

$$\sup_{\tau \in [\nu_1, \tau_{U,1}]} \| \mathbf{s} \{ \hat{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} - \mathbf{s} \{ \tilde{\boldsymbol{\beta}}(\mathbf{r}_0), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} \| \xrightarrow{p} 0. \text{ Hence, we have}$$

$$\sup_{z \in \mathcal{Z}, \tau \in [\nu_1, \tau_{U,1}]} \| \hat{\boldsymbol{\beta}}(\tau; \hat{\mathbf{r}}) - \hat{\boldsymbol{\beta}}(\tau; \mathbf{r}_0) \| \xrightarrow{p} 0 \text{ after using C4 (i) and Taylor expansions.}$$

Mimicking the arguments in Lemma 2, we can show with C3(ii) that

$$\sup_{\tau \in [\nu_1, \tau_{U,1}]} \| \mathbf{S}_n \{ \hat{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \hat{\mathbf{r}}, \tau \} - \mathbf{S}_n \{ \hat{\boldsymbol{\beta}}(\mathbf{r}_0), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} - \mathbf{s} \{ \hat{\boldsymbol{\beta}}(\hat{\mathbf{r}}), \boldsymbol{\alpha}_0, \hat{\mathbf{r}}, \tau \} + \mathbf{s} \{ \hat{\boldsymbol{\beta}}(\mathbf{r}_0), \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau \} \| = o_p(n^{-1/2}), \quad (\text{A24})$$

and similarly,

$$\sup_{\tau \in [\nu_1, \tau_{U,1}]} \|S_n \{\hat{\beta}(\hat{r}), \hat{\alpha}, \hat{r}, \tau\} - S_n \{\hat{\beta}(r_0), \hat{\alpha}, r_0, \tau\} - S_n \{\hat{\beta}(\hat{r}), \alpha_0, \hat{r}, \tau\} + S_n \{\hat{\beta}(r_0), \alpha_0, r_0, \tau\}\| = o_p(n^{-1/2}).$$

Note that $\sup_{\tau \in [\nu_1, \tau_{U,1}]} S_n \{\hat{\beta}(r), \hat{\alpha}, r, \tau\} = o(n^{-1/2})$, and thus

$\sup_{\tau \in [\nu_1, \tau_{U,1}]} \|S_n \{\hat{\beta}(\hat{r}), \alpha_0, \hat{r}, \tau\} - S_n \{\hat{\beta}(r_0), \alpha_0, r_0, \tau\}\| = o_p(n^{-1/2})$. This, combined with (A24), further implies

$$\sup_{\tau \in [\nu_1, \tau_{U,1}]} \|s \{\hat{\beta}(\hat{r}), \alpha_0, \hat{r}, \tau\} - s \{\hat{\beta}(r_0), \alpha_0, r_0, \tau\}\| = o_p(n^{-1/2}).$$

By the uniform convergence of $\hat{\beta}(\tau; r_0)$ to $\beta_0(\tau)$ in $\tau \in [\nu_1, \tau_{U,1}]$, we can use Taylor expansion to get

$$\sup_{\tau \in [\nu_1, \tau_{U,1}]} \|\hat{\beta}(\tau; \hat{r}) - \hat{\beta}(\tau; r_0) + B_b(\beta_0, \alpha_0, r_0, \tau)^{-1} B_r(\beta_0, \alpha_0, r_0, \tau) (\hat{r} - r_0)\| = o_p(n^{-1/2} \vee \|\hat{r} - r_0\|). \quad (\text{A25})$$

By the consistency of r , uniform convergence of $\hat{\beta}(\tau; r_0)$, and (A25), we see after an Taylor expansion that

$$\begin{aligned} & w \{\hat{\beta}(\hat{r}), \alpha_0, \hat{r}\} \\ & - w \{\hat{\beta}(r_0), \alpha_0, r_0\} \\ & = L_r(\beta_0, \alpha_0, r_0) (\hat{r} - r_0) \\ & + \int_{\tau_a}^{\tau_b} L_b(\beta_0, \alpha_0, r_0, \tau) \{\hat{\beta}(\tau; \hat{r}) - \hat{\beta}(\tau; r_0)\} d\tau \quad (\text{A26}) \\ & + o_p(n^{-1/2} \vee \|\hat{r} - r_0\|) \\ & = D_r(\beta_0, \alpha_0, r_0) \\ & \times (\hat{r} - r_0) + o_p(n^{-1/2} \vee \|\hat{r} - r_0\|), \end{aligned}$$

where

$$D_r(\beta_0, \alpha_0, r_0) = L_r(\beta_0, \alpha_0, r_0) - \int_{\tau_a}^{\tau_b} L_b(\beta_0, \alpha_0, r_0, \tau) B_b(\beta_0, \alpha_0, r_0, \tau)^{-1} \times B_r(\beta_0, \alpha_0, r_0, \tau) d\tau$$

Based on (A7), $D_r(\beta_0, \alpha_0, r_0)$ equals $\partial_w \{\tilde{\beta}(r), \alpha_0, r\} / \partial r|_{r=r_0}$.

Following the arguments for Lemma 2, we can also show that

$n^{1/2} [W_n \{\hat{\beta}(\hat{r}), \hat{\alpha}, \hat{r}\} - W_n \{\hat{\beta}(r_0), \hat{\alpha}, r_0\}]$ can be approximated by

$n^{1/2} [w \{\hat{\beta}(\hat{r}), \hat{\alpha}, \hat{r}\} - w \{\hat{\beta}(r_0), \hat{\alpha}, r_0\}]$, and furthermore by

$n^{1/2} [w \{\hat{\beta}(\hat{r}), \alpha_0, \hat{r}\} - w \{\hat{\beta}(r_0), \alpha_0, r_0\}]$. This fact, coupled with

$W_n \{\hat{\beta}(\hat{r}), \hat{\alpha}, \hat{r}\} = o(n^{-1/2})$ and (A26), gives

$$-n^{1/2} \mathbf{W}_n \left\{ \hat{\boldsymbol{\beta}}(\mathbf{r}_0), \hat{\boldsymbol{\alpha}}, \mathbf{r}_0 \right\} = \mathbf{D}_{\mathbf{r}} \times n^{1/2} (\hat{\mathbf{r}} - \mathbf{r}_0) + o_p \left(1 \vee n^{1/2} \|\hat{\mathbf{r}} - \mathbf{r}_0\| \right),$$

which, combined with (A23), further implies

$$n^{1/2} (\hat{\mathbf{r}} - \mathbf{r}_0) = -n^{-1/2} \sum_{i=1}^n \mathbf{D}_{\mathbf{r}}^{-1} \boldsymbol{\ell}_i(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0) + o_p(1). \quad (\text{A27})$$

Here, $\mathbf{D}_{\mathbf{r}}^{-1}$ is a shorthand notation for $\mathbf{D}_{\mathbf{r}}^{-1}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0)$. It then follows that $n^{1/2}(\hat{\mathbf{r}} - \mathbf{r}_0)$ converges to a normal distribution with mean 0.

Finally, by combining (A17) and (A25), we can see that

$$\begin{aligned} & n^{1/2} \left\{ \hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau) \right\} \\ &= n^{1/2} \left\{ \hat{\boldsymbol{\beta}}(\tau; \hat{\mathbf{r}}) - \hat{\boldsymbol{\beta}}(\tau; \mathbf{r}_0) \right\} \\ &+ n^{1/2} \left\{ \hat{\boldsymbol{\beta}}(\tau; \mathbf{r}_0) - \boldsymbol{\beta}_0(\tau) \right\} \\ &= n^{-1/2} \sum_{i=1}^n \left[\mathbf{B}_b(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau)^{-1} \mathbf{B}_r(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau) \mathbf{D}_{\mathbf{r}}^{-1} \boldsymbol{\ell}_i(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0) + \mathbf{B}_b(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau)^{-1} \boldsymbol{\chi}_i(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \mathbf{r}_0, \tau) \right] + o_p(1). \end{aligned}$$

This implies that $n^{1/2} \left\{ \hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau) \right\}$ converges weakly to a Gaussian process with mean 0 for $\tau \in [\nu_1, \tau_{U,1}]$. This completes the proof of Theorem 2.

APPENDIX D: Convergence criteria for the proposed algorithm

The algorithm proposed in Section 2.3 involves iterative procedures which are stopped by certain pre-specified convergence criteria. In the below, we elaborate the convergence criteria that were adopted in our numerical studies and are also expected to function well in other typical settings.

First, to assess the convergence status, the distance between two estimates for $\boldsymbol{\beta}_0(\cdot)$ is

defined as $\mathcal{D}_b(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = \max_{i=1, \dots, p+1} \int_{\nu_1}^{\tau_{U,1}} \|\boldsymbol{\beta}_2^{(i)}(\tau) - \boldsymbol{\beta}_1^{(i)}(\tau)\| d\tau$. The distance between two estimates for $\mathbf{r}_0$ is defined as $\mathcal{D}_r(\mathbf{r}_1, \mathbf{r}_2) = |Ken(\mathbf{r}_1) - Ken(\mathbf{r}_2)|$, where $Ken(\mathbf{r})$ denotes the Kendall's tau coefficient under the assumed model (2) with $\mathbf{r}_0 = \mathbf{r}$. We define the distance measure for $\mathbf{r}$ on the Kendall's tau's scale in order to have a unified criterion for different choices of copula functions.

For the iterations within Steps D.0–D.2, the convergence of the sequence,

$\left\{ \hat{\boldsymbol{\beta}}^{[k,m]}(\tau) : m=0, 1, \dots, \right\}$, is determined via the following procedure.

- a. Set a maximum number of iterations as $max.iter_0$, and choose convergence tolerance levels as $tol_{b,1}$ and $tol_{b,2}$.

- b. In the q th iteration ($q < \max.iter_0$), we regard the sequence as converged if

$$\mathcal{D}_b \left(\hat{\beta}^{[k,q]}, \hat{\beta}^{[k,q-1]} \right) \leq tol_{b,1} \text{ or } \mathcal{D}_b \left(\hat{\beta}^{[k,q]}, \hat{\beta}^{[k,q-2]} \right) \leq tol_{b,1}. \text{ The final estimate,}$$

$$\hat{\beta} \left(\tau; \mathbf{r}^{[k]} \right), \text{ is set as } \hat{\beta}^{[k,q]} \left(\tau \right) \text{ in the former case and as}$$

$$\left\{ \hat{\beta}^{[k,q]} \left(\tau \right) + \hat{\beta}^{[k,q-1]} \left(\tau \right) \right\} / 2, \text{ in the latter case.}$$

- c. When $q = \max.iter_0$, the sequence is concluded as convergent if

$$\mathcal{D}_b \left(\hat{\beta}^{[k,q]}, \hat{\beta}^{[k,q-1]} \right) \leq tol_{b,2} \text{ or } \mathcal{D}_b \left(\hat{\beta}^{[k,q]}, \hat{\beta}^{[k,q-2]} \right) \leq tol_{b,2}, \text{ and as non-convergent}$$

otherwise. In the convergent case, $\hat{\beta} \left(\tau; \mathbf{r}^{[k]} \right)$ is set in the same way as in 2.

The convergence criteria for the iterations within Steps A-D need to concern both

$$\left\{ \hat{\beta}^{[k]} \left(\tau \right) : k=0, 2, \dots \right\} \text{ and } \left\{ \mathbf{r}^{[k]} \left(\tau \right) : k=1, 2, \dots \right\}. \text{ We assess the convergence of}$$

$\left\{ \hat{\beta}^{[k]} \left(\tau \right) : k=1, 2, \dots \right\}$ in the same way as that for $\left\{ \hat{\beta}^{[k]} \left(\tau \right) : k=1, 2, \dots \right\}$, except that we may adopt a different maximum number of iterations, $\max.iter_1$. We stop the algorithm at the q th

iteration ($q = \max.iter_1$) if the sequence, $\left\{ \hat{\beta}^{[k]} \left(\tau \right) : k=1, 2, \dots \right\}$, reaches convergence and

$$\mathcal{D}_r \left(\mathbf{r}^{[q]}, \mathbf{r}^{[q-1]} \right) \leq tol_k, \text{ where } tol_k \text{ is a prespecified tolerance level.}$$

In the numerical studies reported in Section 3 and 4, we choose $tol_{b,1} = 5 \times 10^{-4}$ and $tol_{b,2} = tol_k = 0.005$, and set $\max.iter_0 = 10$ and $\max.iter_1 = 20$.

References

- Chen Y. Semiparametric marginal regression analysis for dependent competing risks under an assumed copula. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2010; 72(2):235–251.
- Chen Y. Maximum likelihood analysis of semicompeting risks data with semiparametric regression models. *Lifetime Data Analysis*. 2011:1–22. [PubMed: 20617379]
- Clayton D. A Model for Association in Bivariate Life Tables and its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence. *Biometrika*. 1978; 65(1):141–151.
- Copelan E, Biggs J, Thompson J, Crilley P, Szer J, Klein J, Kapoor N, Avalos B, Cunningham I, Atkinson K. Treatment for acute myelocytic leukemia with allogeneic bone marrow transplantation following preparation with bucy2. *Blood*. 1991; 78(3):838–843. [PubMed: 1859895]
- Ding A, Shi G, Wang W, Hsieh J. Marginal regression analysis for semi-competing risks data under dependent censoring. *Scandinavian Journal of Statistics*. 2009; 36(3):481–500.
- Efron B. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*. 1979; 7(1):1–26.
- Fine J, Jiang H, Chappell R. On Semi-competing Risks Data. *Biometrika*. 2001; 88(4):907–919.
- Fygenso M, Ritov Y. Monotone Estimating Equations for Censored Data. *The Annals of Statistics*. 1994; 22(2):732–746.
- Genest C. Frank's Family of Bivariate Distributions. *Biometrika*. 1987; 74(3):549–555.
- Hsieh J, Wang W, Adam Ding A. Regression Analysis based on Semicompeting Risks Data. *Journal of the Royal Statistical Society: Series B(Statistical Methodology)*. 2008; 70(1):3–20.
- Huang Y. Quantile Calculus and Censored Regression. *The Annals of Statistics*. 2010; 38(3):1607–1637.

- Jiang H, Chappell R, Fine J. Estimating the distribution of nonterminal event time in the presence of mortality or informative dropout. *Controlled Clinical Trials*. 2003; 24:135–146. [PubMed: 12689735]
- Kendall, M.; Gibbons, J. Rank Correlation Methods. Charles Griffin; London: 1962.
- Klein, J.; Moeschberger, M. Survival analysis: techniques for censored and truncated data. Springer; New York: 2005.
- Koenker, R. Quantile Regression. Cambridge Univ Pr; 2005.
- Koenker R, Bassett G. Regression Quantiles. *Econometrica*. 1978; 46(1):33–50.
- Kosorok, M. Introduction to Empirical Processes and Semiparametric Inference. Springer Verlag; 2008.
- Lai T, Ying Z. Stochastic Integrals of Empirical-type Processes with Applications to Censored Regression. *Journal of Multivariate Analysis*. 1988; 27(2):334–358.
- Lin D, Robins J, Wei L. Comparing two Failure Time Distributions in the presence of Dependent Censoring. *Biometrika*. 1996; 83(2):381–393.
- Martinussen, T.; Scheike, TH. Dynamic Regression Models for Survival Data. Springer; New York: 2006.
- Nelsen, R. An Introduction to Copulas. Springer Us; 2006.
- Peng L, Fine J. Rank Estimation of Accelerated Lifetime Models with Dependent Censoring. *Journal of the American Statistical Association*. 2006; 101(475):1085–1093.
- Peng L, Fine J. Nonparametric Quantile Inference with Competing Risks Data. *Biometrika*. 2007a; 94:735–744.
- Peng L, Fine J. Competing Risks Quantile Regression. *Journal of the American Statistical Association*. 2009; 104(488):1440–1453.
- Peng L, Fine JP. Regression modeling of semi-competing risks data. *Biometrics*. 2007b; 63:96–108. [PubMed: 17447934]
- Peng L, Huang Y. Survival Analysis with Quantile Regression Models. *Journal of the American Statistical Association*. 2008; 103(482):637–649.
- Portnoy S. Censored Regression Quantiles. *Journal of the American Statistical Association*. 2003; 98(464):1001–1013.
- Powell J. Censored Regression Quantiles. *Journal of Econometrics*. 1986; 32(1):143–155.
- Tsiatis A. A Nonidentifiability Aspect of the Problem of Competing Risks. *Proceedings of the National Academy of Sciences*. 1975; 72(1):20–22.
- Van der Vaart, A.; Wellner, J. Weak Convergence and Empirical Processes: with Applications to Statistics. Springer Verlag; 1996.
- Wang H, Wang L. Locally weighted censored quantile regression. *Journal of the American Statistical Association*. 2009; 104(487):1117–1128.
- Ying Z, Jung S, Wei L. Survival Analysis with Median Regression Models. *Journal of the American Statistical Association*. 1995; 90:178–184.
- Zheng M, Klein J. Estimates of Marginal Survival for Dependent Competing Risks based on an Assumed Copula. *Biometrika*. 1995; 82(1):127–138.

Fig. 1.

Assessment of the robustness of $\hat{\beta}(\tau)$ with data generated from set-up S2.C (the first row) and from set-up S2.F (the second row). Solid lines represent true regression quantiles, dashed lines represent empirical averages of naive estimates, dotted lines represent empirical averages of the proposed estimates based on Clayton's copula, and dashed dotted lines represent empirical averages of the proposed estimates based on Frank's copula.

Fig. 2.

Estimated regression coefficients for the BMT dataset, where the x-axis represents τ and the y-axis represents the regression coefficients. Bold solid line represents the proposed regression coefficients as a function of τ , dashed line represents the estimator based on Peng and Huang's method, and dotted lines represent the percentile-based 95% intervals.

Fig. 3.

Estimated quantiles of time to the GVHD endpoint, where x-axis represents quantiles, and y-axis represents predicted quantiles in months. The solid line and dashed line correspond to the proposed estimator and Peng and Huang (2008)'s estimator, respectively.

Table 1

Summary of simulation setups, with $P_{00} = P(\delta = 0, \eta = 0)$, $P_{01} = P(\delta = 0, \eta = 1)$, $P_{10} = P(\delta = 1, \eta = 0)$ and $P_{11} = P(\delta = 1, \eta = 1)$.

Setup	Copula	μ_2	(a_1, a_2)	(b_1, b_2)	U_C	P_{00}	P_{01}	P_{10}	P_{11}
S1.C	Clayton	0.1	(0.4, 0.2)	(0, 0)	18	0.06	0.15	0.04	0.76
S1.F	Frank	0.1	(0.4, 0.2)	(0, 0)	18	0.06	0.15	0.04	0.76
S2.C	Clayton	0.0	(0.3, -0.25)	(-0.4, 0)	3.5	0.24	0.24	0.09	0.43
S2.F	Frank	0.0	(0.3, -0.25)	(-0.4, 0)	3.5	0.23	0.24	0.10	0.43

Table 2

Summary of simulation results on $\hat{\beta}(\tau)$, which includes the empirical biases ($\times 10^3$), empirical standard errors ($\times 10^3$), averages of resampling based standard error estimates ($\times 10^3$), and empirical coverages (%) of 95% wald-type and percentile-based confidence intervals.

τ	EBias($\times 10^3$)			ESD($\times 10^3$)			ASD($\times 10^3$)			ECP _w			ECP _p		
	$\hat{\beta}^{(0)}$	$\hat{\beta}^{(1)}$	$\hat{\beta}^{(2)}$	$\hat{\beta}^{(0)}$	$\hat{\beta}^{(1)}$	$\hat{\beta}^{(2)}$	$\hat{\beta}^{(0)}$	$\hat{\beta}^{(1)}$	$\hat{\beta}^{(2)}$	$\hat{\beta}^{(0)}$	$\hat{\beta}^{(1)}$	$\hat{\beta}^{(2)}$	$\hat{\beta}^{(0)}$	$\hat{\beta}^{(1)}$	$\hat{\beta}^{(2)}$
Set-up S1.C															
0.1	4	2	2	100	162	99	102	167	105	93.9	94.6	94.5	93.8	95.9	95.5
0.2	5	-4	1	89	143	86	94	151	92	93.9	94.5	95.0	94.5	95.9	95.7
0.3	4	-2	2	86	138	84	90	144	88	94.3	94.5	94.9	95.9	96.2	96.2
0.4	4	-3	2	83	133	82	87	140	87	95.1	95.7	95.5	95.5	95.9	96.2
0.5	4	-2	3	81	131	83	86	138	87	94.7	94.7	95.9	95.2	96.2	95.4
0.6	4	-1	3	81	131	86	86	139	88	94.3	94.9	95.4	95.0	95.6	95.0
0.7	5	-4	4	82	134	87	87	142	92	94.6	95.3	94.5	95.1	94.8	95.5
Set-up S1.F															
0.1	13	-10	4	101	159	99	102	164	101	93.4	95.0	93.9	94.1	96.5	95.3
0.2	11	-8	3	86	136	86	89	142	87	94.5	96.3	94.5	94.3	96.6	96.0
0.3	11	-9	4	75	121	78	81	131	81	95.1	96.5	94.2	95.2	96.2	95.2
0.4	10	-7	1	71	115	75	77	124	78	96.0	96.2	95.2	96.5	97.2	95.0
0.5	11	-9	2	69	111	74	74	121	78	95.3	96.4	95.7	95.7	97.1	95.3
0.6	11	-9	2	66	111	76	74	120	78	94.8	96.5	95.5	95.1	96.9	95.6
0.7	12	-11	2	70	117	79	75	124	81	94.3	94.6	95.4	94.7	95.1	95.7
Set-up S2.C															
0.1	9	-8	9	106	171	113	112	180	113	94.4	95.6	92.7	94.6	95.7	93.3
0.2	6	-1	5	97	157	106	110	171	110	96.3	95.9	94.9	96.3	97.3	96.2
0.3	10	-8	5	97	152	108	109	168	112	95.7	95.8	95.3	96.8	97.2	95.9
0.4	10	-9	6	100	156	110	107	167	117	96.3	95.7	95.8	95.9	97.3	96.0
0.5	13	-10	6	95	151	116	107	168	124	95.4	95.8	95.6	95.8	96.5	95.4
0.6	15	-14	13	97	154	126	107	170	131	95.7	95.5	94.7	95.4	96.2	95.6
Set-up S2.F															
0.1	26	-30	13	125	191	118	114	179	115	88.6	91.8	91.0	88.8	93.4	93.2
0.2	18	-21	18	108	164	111	108	167	111	92.6	94.8	93.4	92.3	94.8	94.0

τ	EBias($\times 10^3$)			ESD($\times 10^3$)			ASD($\times 10^3$)			ECP _W			ECP _P		
	$\beta^{(0)}$	$\beta^{(1)}$	$\beta^{(2)}$	$\beta^{(0)}$	$\beta^{(1)}$	$\beta^{(2)}$	$\beta^{(0)}$	$\beta^{(1)}$	$\beta^{(2)}$	$\beta^{(0)}$	$\beta^{(1)}$	$\beta^{(2)}$	$\beta^{(0)}$	$\beta^{(1)}$	$\beta^{(2)}$
0.3	21	-26	20	98	148	106	101	157	108	93.4	95.3	94.0	93.7	96.2	94.0
0.4	22	-27	20	91	137	102	95	149	106	93.8	95.2	94.2	92.3	95.5	93.3
0.5	16	-15	14	84	129	101	93	147	108	95.2	95.5	95.3	92.7	95.3	93.6
0.6	19	-21	20	84	133	107	93	149	112	95.5	96.1	93.7	94.8	95.6	93.9

Table 3

Summary of simulation results on $r, \hat{r}$ which include empirical biases ($\times 10^3$), empirical standard errors ($\times 10^3$), averages of resampling based standard error estimates ($\times 10^3$), empirical coverages (%) of 95% wald-type and percentile-based confidence intervals. The same sets of summary statistics are also presented for the corresponding Kendall's tau coefficient estimates, $\mathcal{K}(\hat{r})$.

		TRUE	EBias($\times 10^3$)	ESD($\times 10^3$)	ASD($\times 10^3$)	ECP _W	ECP _P
S1.C	r	0.693	-14	206	226	96.7	94.8
	$K(r)$	0.500	-4	51	56	96.1	94.8
S1.F	r	5.750	-215	852	885	94.5	92.5
	$K(r)$	0.500	-15	48	51	95.2	92.5
S2.C	r	0.693	-30	305	370	98.6	96.0
	$K(r)$	0.500	-7	74	90	97.7	96.0
S2.F	r	5.750	-356	1317	1371	94.8	92.3
	$K(r)$	0.500	-28	77	81	96.1	92.3

Table 4

Summary of simulation results on second stage inferences, which include empirical bias ($\times 10^3$), empirical standard error ($\times 10^3$) and average of resampling based standard error estimates ($\times 10^3$) of $\hat{\Phi}_1^{(q)}$, $q = 2, 3$, as well as the empirical rejection rates of H_{01} and H_{02} with Wald-type (ERR_W) and percentile-based (ERR_P) methods.

	q	$\hat{\Phi}_1^{(q)}$			H_{01}		H_{02}	
		EBias	ESD	ASD	ERR_W	ERR_P	ERR_W	ERR_P
S1.C	2	-2	108	109	0.056	0.062	0.049	0.038
	3	-1	68	68	0.129	0.129	0.936	0.946
S1.F	2	-9	93	96	0.046	0.056	0.043	0.033
	3	-1	61	61	0.162	0.178	0.950	0.967
S2.C	2	-9	123	128	0.886	0.930	0.026	0.017
	3	4	88	87	0.190	0.145	0.561	0.736
S2.F	2	-22	115	116	0.950	0.973	0.040	0.020
	3	15	83	80	0.167	0.136	0.614	0.750