ANALYSIS OF DEPENDENTLY CENSORED DATA BASED ON QUANTILE REGRESSION

Shuang Ji[1], Limin Peng[2], Ruosha Li[3], and Michael J. Lynn[2]

[1]*Division of Biostatistics, University of Texas Health Science Center at Houston*
[2]*Department of Biostatistics and Bioinformatics, Emory University*
[3]*Department of Biostatistics, University of Pittsburgh*

Supplementary Material

S1 Appendix A: Regularity Conditions

Before stating the regularity conditions, we introduce some necessary notation. For a random variable W, define $F_W(t|\boldsymbol{Z}) = \Pr(W \le t|\boldsymbol{Z})$ and $f_W(t|\boldsymbol{Z}) = dF_W(t|\boldsymbol{Z})/dt$. Define $\tilde{F}_T(t|\boldsymbol{Z}) = \Pr(T \le t, \delta = 1|\boldsymbol{Z})$ and $\tilde{F}_D(t|\boldsymbol{Z}) = \Pr(D \le t, \delta = 1|\boldsymbol{Z})$, and $\tilde{f}_T(t)$ and $\tilde{f}_D(t)$ as the derivatives of $\tilde{F}_T(t|\boldsymbol{Z})$ and $\tilde{F}_D(t|\boldsymbol{Z})$ with respect to t, respectively. For a vector $\boldsymbol{r}$, let $\boldsymbol{r}^{\otimes 2}$ denote $\boldsymbol{r}\boldsymbol{r}^T$ and $\|\boldsymbol{r}\|$ denote the Euclidean norm of $\boldsymbol{r}$. Define $\boldsymbol{s}^{(k)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) = E\{\boldsymbol{S}_n^{(k)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau)\}$ for $k = 1, 2$, $\boldsymbol{\mu}_1(\boldsymbol{b}) = E[\boldsymbol{Z}N_1\{g(\boldsymbol{Z}^T\boldsymbol{b})\}]$, $\boldsymbol{\mu}_2(\boldsymbol{a}) = E[\boldsymbol{Z}N_2\{g(\boldsymbol{Z}^T\boldsymbol{a})\}]$, $\boldsymbol{J}_1(\boldsymbol{b}) = E[\boldsymbol{Z}^{\otimes 2}\tilde{f}_T\{g(\boldsymbol{Z}^T\boldsymbol{b})|\boldsymbol{Z}\}g'(\boldsymbol{Z}^T\boldsymbol{b})]$, $\boldsymbol{J}_2(\boldsymbol{a}) = E[\boldsymbol{Z}^{\otimes 2}\tilde{f}_D\{g(\boldsymbol{Z}^T\boldsymbol{a})|\boldsymbol{Z}\}g'(\boldsymbol{Z}^T\boldsymbol{a})]$, $\tilde{\boldsymbol{\mu}}_1(\boldsymbol{b}, \boldsymbol{\alpha}) = E[\boldsymbol{Z}Y\{g(\boldsymbol{Z}^T\boldsymbol{b})\}\times\phi_1(1-u, 1-\int_0^{\tau_{U,2}} I\{\boldsymbol{Z}^T\boldsymbol{\alpha}(v) \le \boldsymbol{Z}^T\boldsymbol{b}\}\,\mathrm{d}v)]$, $\tilde{\boldsymbol{\mu}}_2(\boldsymbol{a}, \boldsymbol{\beta}) = E[\boldsymbol{Z}Y\{g(\boldsymbol{Z}^T\boldsymbol{a})\}\times \phi_2(1-\int_0^{\tau_{U,1}} I\{\boldsymbol{Z}^T\boldsymbol{\beta}(v) \le \boldsymbol{Z}^T\boldsymbol{a}\}\,\mathrm{d}v, 1-u)]$, $\tilde{\boldsymbol{J}}_1(\boldsymbol{b}, \boldsymbol{\alpha}) = \partial\tilde{\boldsymbol{\mu}}_1(\boldsymbol{b}, \boldsymbol{\alpha})/\partial\boldsymbol{b}$, $\tilde{\boldsymbol{J}}_2(\boldsymbol{a}, \boldsymbol{\beta}) = \partial\tilde{\boldsymbol{\mu}}_2(\boldsymbol{a}, \boldsymbol{\beta})/\partial\boldsymbol{a}$.

For any constant d, define $\mathscr{B}(d) = \{(\boldsymbol{b}^T, \boldsymbol{a}^T)^T : \boldsymbol{b} \in \mathbb{R}^{p+1}, \boldsymbol{a} \in \mathbb{R}^{p+1}, \inf_{\tau\in(0,\tau_U]} \|\boldsymbol{\mu}_1(\boldsymbol{b}) - \boldsymbol{\mu}_1\{\boldsymbol{\beta}_0(\tau)\}\| \le d, \inf_{\tau\in(0,\tau_U]} \|\boldsymbol{\mu}_2(\boldsymbol{a}) - \boldsymbol{\mu}_2\{\boldsymbol{\alpha}_0(\tau)\}\| \le d\}$. Let $\mathcal{D}$ denote a function space that contains all continuous functions mapping $[0, 1]$ to $\mathbb{R}^{2p+2}$, and $\mathcal{F} = \{c(\boldsymbol{G}_1 - \boldsymbol{G}_2) : c \in \mathbb{R}, \boldsymbol{G}_j \in \mathcal{D}, j = 1, 2\}$.

The regularity conditions are:

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

C2. $f_T(t|\boldsymbol{z})$, $f_D(t|\boldsymbol{z})$, $\tilde{f}_T(t|\boldsymbol{z})$ and $\tilde{f}_D(t|\boldsymbol{z})$ are bounded above uniformly in t and $\boldsymbol{z}$.

C3. (a) $\tilde{f}_T\{g(\boldsymbol{Z}^T\boldsymbol{b})|\boldsymbol{Z}\} > 0$ and $\tilde{f}_D\{g(\boldsymbol{Z}^T\boldsymbol{a})|\boldsymbol{Z}\} > 0$ for all $(\boldsymbol{b}^T, \boldsymbol{a}^T)^T \in \mathscr{B}(d_0)$, where d_0 is a constant; (b) $E(\boldsymbol{Z}^{\otimes 2}) > 0$; (c) $\inf_{\tau\in[\nu_1,\tau_{U,1}]} eigmin(\boldsymbol{J}_1\{\boldsymbol{\beta}_0(\tau)\}) > 0$ and $\inf_{\tau\in[\nu_2,\tau_{U,2}]} eigmin(\boldsymbol{J}_2\{\boldsymbol{\alpha}_0(\tau)\}) > 0$ for any $\nu_1 \in (0, \tau_{U,1}]$ and $\nu_2 \in (0, \tau_{U,2}]$, where $eigmin(\cdot)$ denotes the minimum eigenvalue of a matrix.

C4. (a) Each component of $\tilde{\boldsymbol{\mu}}_1\{\boldsymbol{\beta}_0(\tau), \boldsymbol{\alpha}_0\}$ and $\tilde{\boldsymbol{\mu}}_2\{\boldsymbol{\alpha}_0(\tau), \boldsymbol{\beta}_0\}$ is a Lipschitz function of τ; (b) $\phi_1(u, v)$ and $\phi_2(u, v)$ are differentiable with respect to u and v, and furthermore, each component of $E\big[\boldsymbol{Z}Y\{g(\boldsymbol{Z}^T\boldsymbol{b})\} \times \phi_{12}(1-u, 1-\int_0^{\tau_{U,2}} I\{\boldsymbol{Z}^T\boldsymbol{\alpha}_0(v) \le \boldsymbol{Z}^T\boldsymbol{b}\}\,\mathrm{d}v)\big]$ and $E\big[\boldsymbol{Z}Y\{g(\boldsymbol{Z}^T\boldsymbol{a})\} \times \phi_{21}(1-\int_0^{\tau_{U,1}} I\{\boldsymbol{Z}^T\boldsymbol{\beta}_0(v) \le \boldsymbol{Z}^T\boldsymbol{a}\}\,\mathrm{d}v, 1-u)\big]$ is bounded uniformly in $(\boldsymbol{b}^T, \boldsymbol{a}^T)^T \in \mathscr{B}(d_0)$, where $\phi_{12}(u, v) = \partial\phi_1(u, v)/\partial v$ and $\phi_{21}(u, v) = \partial\phi_2(u, v)/\partial u$; (c) $\boldsymbol{w}_1(\boldsymbol{b}, \boldsymbol{a}) = E\big[\boldsymbol{Z}Y\{g(\boldsymbol{Z}^T\boldsymbol{b})\}$
$\times\ \phi_{12}(1-u, 1-\int_0^{\tau_{U,2}} I\{\boldsymbol{Z}^T\boldsymbol{\alpha}_0(v) \le \boldsymbol{Z}^T\boldsymbol{b}\}\,\mathrm{d}v) \times I(\boldsymbol{Z}^T\boldsymbol{b} \ge \boldsymbol{Z}^T\boldsymbol{a})\big]$ and $\boldsymbol{w}_2(\boldsymbol{b}, \boldsymbol{a}) = E\big[\boldsymbol{Z}Y\{g(\boldsymbol{Z}^T\boldsymbol{a})\}$
$\times\,\phi_{21}(1-\int_0^{\tau_{U,1}} I\{\boldsymbol{Z}^T\boldsymbol{\beta}_0(v) \le \boldsymbol{Z}^T\boldsymbol{a}\}\,\mathrm{d}v, 1-u) \times I(\boldsymbol{Z}^T\boldsymbol{a} \ge \boldsymbol{Z}^T\boldsymbol{b})\big]$ are differentiable with respect to $\boldsymbol{b}$ and $\boldsymbol{a}$, and furthermore, each component of $\boldsymbol{w}_{12}\{\boldsymbol{b}, \boldsymbol{\alpha}_0(\tau)\}$ and $\boldsymbol{w}_{21}\{\boldsymbol{\beta}_0(\tau), \boldsymbol{a}\}$ is bounded uniformly in $(\boldsymbol{b}^T, \boldsymbol{a}^T)^T \in \mathscr{B}(d_0)$ and $\tau \in (0, \max\{\tau_{U,1}, \tau_{U,2}\}]$, where $\boldsymbol{w}_{12}(\boldsymbol{b}, \boldsymbol{a}) = \partial\boldsymbol{w}_1(\boldsymbol{b}, \boldsymbol{a})/\partial\boldsymbol{a}$ and $\boldsymbol{w}_{21}(\boldsymbol{b}, \boldsymbol{a}) = \partial\boldsymbol{w}_2(\boldsymbol{b}, \boldsymbol{a})/\partial\boldsymbol{b}$.

C5. (a) For any fixed τ, $\boldsymbol{\rho}(\boldsymbol{\theta}, \tau)$, as a functional of $\boldsymbol{\theta}(\cdot)$ defined on $\mathcal{D}$, is Gâteaux differentiable at $\boldsymbol{\theta}_0(\cdot)$ with derivative $\boldsymbol{\rho}'_{\boldsymbol{\theta}_0}$, where $\boldsymbol{\theta}_0(u) = \big(\boldsymbol{\mu}_1\{\boldsymbol{\beta}_0(u)\}^T, \boldsymbol{\mu}_2\{\boldsymbol{\alpha}_0(u)\}^T\big)^T$; (b) $\|\boldsymbol{\rho}'_{\boldsymbol{\theta}_0}(\boldsymbol{h})\| > 0$ for any $\boldsymbol{h} \in \mathcal{F}$ such that $\sup_{\tau\in(0,1)} \|\boldsymbol{h}(\tau)\| \ne 0$.

The boundedness of covariates and density functions imposed by Conditions C1 and C2 are quite realistic. Condition C3(a)-(b) assume positive density functions and positive definiteness of $E(\boldsymbol{Z}^{\otimes 2})$, which are common requirements in quantile regression literature. Condition C3(c) is the technical assumption for ensuring the identifiability of $\{\boldsymbol{\beta}_0(\tau), \tau \in (0, \tau_{U,1}]\}$ and $\{\boldsymbol{\alpha}_0(\tau), \tau \in (0, \tau_{U,2}]\}$. Conditions C4 and C5 essentially require the smoothness of regression quantiles and the limits of the estimating equations, which are reasonable in many applications.

S2 Appendix B: Proof of Theorem 2.1

For simplicity, we assume $\tau_{U,1} = \tau_{U,2} = \tau_U$. We present the proof of both theorems based on the estimating equation (2.9), which can be adapted to the proof based on the estimating equation (2.10) with minor modification. Let $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\alpha}}$ be abbreviations for functions $\boldsymbol{\beta}(\cdot)$, $\boldsymbol{\alpha}(\cdot)$, $\hat{\boldsymbol{\beta}}(\cdot)$ and $\hat{\boldsymbol{\alpha}}(\cdot)$. With fixed $\boldsymbol{\alpha}$, define $\hat{\boldsymbol{\beta}}(\boldsymbol{\alpha}, \tau)$ and $\tilde{\boldsymbol{\beta}}(\boldsymbol{\alpha}, \tau)$ as the solutions for $\boldsymbol{\beta}$ to $\boldsymbol{S}_n^{(1)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) = 0$ and $\boldsymbol{s}^{(1)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) = 0$ respectively. Similarly, with fixed $\boldsymbol{\beta}$, define $\hat{\boldsymbol{\alpha}}(\boldsymbol{\beta}, \tau)$ and $\tilde{\boldsymbol{\alpha}}(\boldsymbol{\beta}, \tau)$ as the solutions for $\boldsymbol{\alpha}$ to $\boldsymbol{S}_n^{(2)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) = 0$ and $\boldsymbol{s}^{(2)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) = 0$ respectively. It is easy to see that $\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}}, \tau) = \hat{\boldsymbol{\beta}}(\tau)$, $\hat{\boldsymbol{\alpha}}(\hat{\boldsymbol{\beta}}, \tau) = \hat{\boldsymbol{\alpha}}(\tau)$, $\tilde{\boldsymbol{\beta}}(\boldsymbol{\alpha}_0, \tau) = \boldsymbol{\beta}_0(\tau)$ and $\tilde{\boldsymbol{\alpha}}(\boldsymbol{\beta}_0, \tau) = \boldsymbol{\alpha}_0(\tau)$ for $\tau \in (0, \tau_U]$.

Using the Glivenko-Cantelli Theorem (van der Vaart and Wellner 1996), we can show

$$\sup_{\tau \in (0, \tau_U]} \|\boldsymbol{S}_n^{(k)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) - \boldsymbol{s}^{(k)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau)\| \xrightarrow{p} 0, k = 1, 2. \tag{S2.1}$$

Together with the facts that $\boldsymbol{S}_n^{(k)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) = 0$ and $\boldsymbol{s}^{(k)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau) = 0$, (S2.1) implies

$$\sup_{\tau \in (0, \tau_U]} \|\boldsymbol{s}^{(1)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{s}^{(1)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)\| \xrightarrow{p} 0,$$

$$\sup_{\tau \in (0, \tau_U]} \|\boldsymbol{s}^{(2)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{s}^{(2)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)\| \xrightarrow{p} 0.$$

Let

$$\boldsymbol{r} = \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{a} \end{pmatrix}, \quad \boldsymbol{\mu}(\boldsymbol{r}) = \begin{pmatrix} \boldsymbol{\mu}_1(\boldsymbol{b}) \\ \boldsymbol{\mu}_2(\boldsymbol{a}) \end{pmatrix} \quad \text{and} \quad \mathscr{A}(d) = \{\boldsymbol{\mu}(\boldsymbol{r}) : \boldsymbol{r} \in \mathscr{B}(d)\}.$$

By condition C3(a)-(b), we can show that $\boldsymbol{\mu}(\cdot)$ is a 1-1 mapping from $\mathscr{B}(d_0)$ to $\mathscr{A}(d_0)$. Hence, there exists an inverse function of $\boldsymbol{\mu}(\cdot)$, denoted by $\boldsymbol{\kappa}(\cdot)$, mapping $\mathscr{A}(d_0)$ to $\mathscr{B}(d_0)$, such that $\boldsymbol{\kappa}\{\boldsymbol{\mu}(\boldsymbol{r})\} = \boldsymbol{r}$ for any $\boldsymbol{r} \in \mathscr{B}(d_0)$.

Consider the following equalities:

$$\begin{aligned}\boldsymbol{\mu}_1\{\hat{\boldsymbol{\beta}}(\tau)\} - \boldsymbol{\mu}_1\{\boldsymbol{\beta}_0(\tau)\} &= \boldsymbol{\mu}_1\{\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}},\tau)\} - \boldsymbol{\mu}_1\{\tilde{\boldsymbol{\beta}}(\boldsymbol{\alpha}_0,\tau)\} \\ &= \boldsymbol{\mu}_1\{\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}},\tau)\} - \boldsymbol{\mu}_1\{\tilde{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}},\tau)\} + \boldsymbol{\mu}_1\{\tilde{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}},\tau)\} - \boldsymbol{\mu}_1\{\tilde{\boldsymbol{\beta}}(\boldsymbol{\alpha}_0,\tau)\}, \\ \boldsymbol{\mu}_2\{\hat{\boldsymbol{\alpha}}(\tau)\} - \boldsymbol{\mu}_2\{\boldsymbol{\alpha}_0(\tau)\} &= \boldsymbol{\mu}_2\{\hat{\boldsymbol{\alpha}}(\hat{\boldsymbol{\beta}},\tau)\} - \boldsymbol{\mu}_2\{\tilde{\boldsymbol{\alpha}}(\boldsymbol{\beta}_0,\tau)\} \\ &= \boldsymbol{\mu}_2\{\hat{\boldsymbol{\alpha}}(\hat{\boldsymbol{\beta}},\tau)\} - \boldsymbol{\mu}_2\{\tilde{\boldsymbol{\alpha}}(\hat{\boldsymbol{\beta}},\tau)\} + \boldsymbol{\mu}_2\{\tilde{\boldsymbol{\alpha}}(\hat{\boldsymbol{\beta}},\tau)\} - \boldsymbol{\mu}_2\{\tilde{\boldsymbol{\alpha}}(\boldsymbol{\beta}_0,\tau)\},\end{aligned} \tag{S2.2}$$

Following Peng and Huang (2008), we can show that $\sup_{\tau\in(0,\tau_U]}\|\boldsymbol{\mu}_1\{\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}},\tau)\} - \boldsymbol{\mu}_1\{\tilde{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}},\tau)\}\| = \boldsymbol{o}_p(1)$ and $\sup_{\tau\in(0,\tau_U]}\|\boldsymbol{\mu}_2\{\hat{\boldsymbol{\alpha}}(\hat{\boldsymbol{\beta}},\tau)\} - \boldsymbol{\mu}_2\{\tilde{\boldsymbol{\alpha}}(\hat{\boldsymbol{\beta}},\tau)\}\| = \boldsymbol{o}_p(1)$. Then (S2.2) can be rewritten as

$$\begin{aligned}\boldsymbol{\mu}_1\{\hat{\boldsymbol{\beta}}(\tau)\} - \boldsymbol{\mu}_1\{\boldsymbol{\beta}_0(\tau)\} &= \boldsymbol{\mu}_1\{\tilde{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}},\tau)\} - \boldsymbol{\mu}_1\{\tilde{\boldsymbol{\beta}}(\boldsymbol{\alpha}_0,\tau)\} + \boldsymbol{o}_{(0,\tau_U]}(1), \\ \boldsymbol{\mu}_2\{\hat{\boldsymbol{\alpha}}(\tau)\} - \boldsymbol{\mu}_2\{\boldsymbol{\alpha}_0(\tau)\} &= \boldsymbol{\mu}_2\{\tilde{\boldsymbol{\alpha}}(\hat{\boldsymbol{\beta}},\tau)\} - \boldsymbol{\mu}_2\{\tilde{\boldsymbol{\alpha}}(\boldsymbol{\beta}_0,\tau)\} + \boldsymbol{o}_{(0,\tau_U]}(1),\end{aligned} \tag{S2.3}$$

where $\boldsymbol{o}_I(1)$ denotes a term that converges to 0 in probability uniformly on the interval I.

Define

$$\boldsymbol{\gamma}(\tau) = \begin{pmatrix}\boldsymbol{\beta}(\tau)\\ \boldsymbol{\alpha}(\tau)\end{pmatrix} \text{ and } \tilde{\boldsymbol{g}}(\boldsymbol{\gamma},\tau) = \begin{pmatrix}\tilde{\boldsymbol{\beta}}(\boldsymbol{\alpha},\tau)\\ \tilde{\boldsymbol{\alpha}}(\boldsymbol{\beta},\tau)\end{pmatrix}.$$

Note that $\tilde{\boldsymbol{g}}(\boldsymbol{\gamma},\tau)$ can be viewed as a functional of $\boldsymbol{\gamma}$ with parameter τ. We further simplify (S2.3) as

$$\boldsymbol{\mu}\{\hat{\gamma}(\tau)\} - \boldsymbol{\mu}\{\boldsymbol{\gamma}_0(\tau)\} = \boldsymbol{\mu}\{\tilde{\boldsymbol{g}}(\hat{\gamma},\tau)\} - \boldsymbol{\mu}\{\tilde{\boldsymbol{g}}(\boldsymbol{\gamma}_0,\tau)\} + \boldsymbol{o}_{(0,\tau_U]}(1). \tag{S2.4}$$

Let $\boldsymbol{\theta}(\tau) = \boldsymbol{\mu}\{\boldsymbol{\gamma}(\tau)\}$ and $\boldsymbol{\rho}(\boldsymbol{\theta}, \tau) = \boldsymbol{\theta}(\tau) - \boldsymbol{\mu}\big(\tilde{\boldsymbol{g}}[\boldsymbol{\kappa}\{\boldsymbol{\theta}(\tau)\}]\big)$. Then (S2.4) becomes

$$\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}, \tau) - \boldsymbol{\rho}(\boldsymbol{\theta}_0, \tau) = \boldsymbol{o}_{(0,\tau_U]}(1). \tag{S2.5}$$

By viewing the parameter τ as fixed and dropping it from the notation of $\boldsymbol{\rho}(\boldsymbol{\theta}, \tau)$ for brevity, we have

$$\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}) - \boldsymbol{\rho}(\boldsymbol{\theta}_0) = \boldsymbol{o}_{(0,\tau_U]}(1). \tag{S2.6}$$

Note that $\boldsymbol{\rho}(\boldsymbol{\theta})$ is a functional of $\boldsymbol{\theta}$. By Condition C5(a), $\boldsymbol{\rho}$ is Gâteaux differentiable at $\boldsymbol{\theta}_0$, that is, for any direction $\boldsymbol{h} \in \mathcal{F}$ and $\boldsymbol{\theta}_0 + t\boldsymbol{h} \in \mathcal{D}$, there is a linear map $\boldsymbol{\rho}'_{\boldsymbol{\theta}_0}$ such that

$$\frac{\boldsymbol{\rho}(\boldsymbol{\theta}_0 + t\boldsymbol{h}) - \boldsymbol{\rho}(\boldsymbol{\theta}_0)}{t} \to \boldsymbol{\rho}'_{\boldsymbol{\theta}_0}(\boldsymbol{h}) \text{ as } t \to 0. \tag{S2.7}$$

Let $\boldsymbol{h} = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)/t$. By (S2.7) we have

$$\{\boldsymbol{\rho}(\hat{\boldsymbol{\theta}}) - \boldsymbol{\rho}(\boldsymbol{\theta}_0)\} - t\boldsymbol{\rho}'_{\boldsymbol{\theta}_0}\{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)/t\} \to 0 \text{ as } t \to 0. \tag{S2.8}$$

By (S2.6), (S2.8), and the linearity of $\boldsymbol{\rho}'_{\boldsymbol{\theta}_0}$, we immediately have

$$\boldsymbol{\rho}'_{\boldsymbol{\theta}_0}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = \boldsymbol{o}_{(0,\tau_U]}(1). \tag{S2.9}$$

Since $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_0$ are bounded on $(0, \tau_U]$, by condition C5(b) and a subsequence argument, (S2.9) implies

$$\sup_{u \in (0,\tau_U]} \|\hat{\boldsymbol{\theta}}(u) - \boldsymbol{\theta}_0(u)\| = \boldsymbol{o}_p(1). \tag{S2.10}$$

Recall that $\boldsymbol{\kappa}\{\hat{\boldsymbol{\theta}}(u)\} = \hat{\boldsymbol{\gamma}}(u)$ and $\boldsymbol{\kappa}\{\boldsymbol{\theta}_0(u)\} = \boldsymbol{\gamma}_0(u)$. By a Taylor expansion of $\boldsymbol{\kappa}\{\hat{\boldsymbol{\theta}}(\tau)\}$ around $\boldsymbol{\theta}_0(\tau)$ for $\tau \in [\nu, \tau_U]$, together with (S2.10) and condition C3(c), we can show

$$\sup_{u \in [\nu,\tau_U]} \|\hat{\boldsymbol{\gamma}}(u) - \boldsymbol{\gamma}_0(u)\| = \boldsymbol{o}_p(1),$$

which implies

$$\sup_{u\in[\nu,\tau_U]} \|\hat{\boldsymbol{\beta}}(u) - \boldsymbol{\beta}_0(u)\| = \boldsymbol{o}_p(1),$$
$$\sup_{u\in[\nu,\tau_U]} \|\hat{\boldsymbol{\alpha}}(u) - \boldsymbol{\alpha}_0(u)\| = \boldsymbol{o}_p(1), \tag{S2.11}$$

and thus complete the proof for Theorem 2.1.

S3 Appendix C: Proof of Theorem 2.2

Having the uniform consistency of $\hat{\boldsymbol{\beta}}(\tau)$ and $\hat{\boldsymbol{\alpha}}(\tau)$ on $\tau \in [\nu, \tau_U]$, by following the proof of Lemma B.1. in Peng and Huang (2008), we can show that

$$\sup_{\tau\in[\nu,\tau_U]} n^{1/2}\|\{\boldsymbol{S}_n^{(k)}(\hat{\boldsymbol{\beta}}, \boldsymbol{\alpha}_0, \tau) - \boldsymbol{S}_n^{(k)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)\} - \{\boldsymbol{s}^{(k)}(\hat{\boldsymbol{\beta}}, \boldsymbol{\alpha}_0, \tau) - \boldsymbol{s}^{(k)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)\}\| \xrightarrow{p} 0,$$

and

$$\sup_{\tau\in[\nu,\tau_U]} n^{1/2}\|\{\boldsymbol{S}_n^{(k)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{S}_n^{(k)}(\hat{\boldsymbol{\beta}}, \boldsymbol{\alpha}_0, \tau)\} - \{\boldsymbol{s}^{(k)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{s}^{(k)}(\hat{\boldsymbol{\beta}}, \boldsymbol{\alpha}_0, \tau)\}\| \xrightarrow{p} 0. \tag{S3.1}$$

From the above results, we get

$$\begin{aligned}
&\sup_{\tau\in[\nu,\tau_U]} n^{1/2}\|\{\boldsymbol{S}_n^{(k)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{S}_n^{(k)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)\} - \{\boldsymbol{s}^{(k)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{s}^{(k)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)\}\| \\
&\le \sup_{\tau\in[\nu,\tau_U]} n^{1/2}\|\{\boldsymbol{S}_n^{(k)}(\hat{\boldsymbol{\beta}}, \boldsymbol{\alpha}_0, \tau) - \boldsymbol{S}_n^{(k)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)\} - \{\boldsymbol{s}^{(k)}(\hat{\boldsymbol{\beta}}, \boldsymbol{\alpha}_0, \tau) - \boldsymbol{s}^{(k)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)\}\| \\
&+ \sup_{\tau\in[\nu,\tau_U]} n^{1/2}\|\{\boldsymbol{S}_n^{(k)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{S}_n^{(k)}(\hat{\boldsymbol{\beta}}, \boldsymbol{\alpha}_0, \tau)\} - \{\boldsymbol{s}^{(k)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{s}^{(k)}(\hat{\boldsymbol{\beta}}, \boldsymbol{\alpha}_0, \tau)\}\| \xrightarrow{p} 0.
\end{aligned} \tag{S3.2}$$

Simple algebra shows that

$$\boldsymbol{s}^{(1)}(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\alpha}},\tau)-\boldsymbol{s}^{(1)}(\boldsymbol{\beta}_0,\boldsymbol{\alpha}_0,\tau)=\boldsymbol{\mu}_1(\hat{\boldsymbol{\beta}},\tau)-\boldsymbol{\mu}_1(\boldsymbol{\beta}_0,\tau)-\int_0^{\tau}\Big([\tilde{\boldsymbol{\mu}}_1\{\hat{\boldsymbol{\beta}}(u),\hat{\boldsymbol{\alpha}}\}-\tilde{\boldsymbol{\mu}}_1\{\hat{\boldsymbol{\beta}}(u),\boldsymbol{\alpha}_0\}] + [\tilde{\boldsymbol{\mu}}_1\{\hat{\boldsymbol{\beta}}(u),\boldsymbol{\alpha}_0\}-\tilde{\boldsymbol{\mu}}_1\{\boldsymbol{\beta}_0(u),\boldsymbol{\alpha}_0\}]\Big)\,\mathrm{d}u. \tag{S3.3}$$

For any $\nu\in(0,\tau_U]$ and any fixed $u\in[\nu,\tau_U]$, given the uniform consistency of $\hat{\boldsymbol{\alpha}}(\cdot)$ and condition C4(b)-(c), we have

$$\begin{aligned}
\tilde{\boldsymbol{\mu}}_1\{\hat{\boldsymbol{\beta}}(u),\hat{\boldsymbol{\alpha}}\}-\tilde{\boldsymbol{\mu}}_1\{\hat{\boldsymbol{\beta}}(u),\boldsymbol{\alpha}_0\} &= E\Big(\boldsymbol{Z}Y[g\{\boldsymbol{Z}^T\hat{\boldsymbol{\beta}}(u)\}][\phi_1(1-u,1-\int_0^{\tau_U}I\{\boldsymbol{Z}^T\hat{\boldsymbol{\alpha}}(v)\le\boldsymbol{Z}^T\hat{\boldsymbol{\beta}}(u)\}\,\mathrm{d}v)\\
&\qquad -\phi_1(1-u,1-\int_0^{\tau_U}I\{\boldsymbol{Z}^T\boldsymbol{\alpha}_0(v)\le\boldsymbol{Z}^T\hat{\boldsymbol{\beta}}(u)\}\,\mathrm{d}v)]\Big)\\
&\approx E\Big(\boldsymbol{Z}Y[g\{\boldsymbol{Z}^T\hat{\boldsymbol{\beta}}(u)\}]\phi_{12}(1-u,1-\int_0^{\tau_U}I\{\boldsymbol{Z}^T\boldsymbol{\alpha}_0(v)\le\boldsymbol{Z}^T\hat{\boldsymbol{\beta}}(u)\}\,\mathrm{d}v)\\
&\qquad \times\int_0^{\tau_U}[I\{\boldsymbol{Z}^T\boldsymbol{\alpha}_0(v)\le\boldsymbol{Z}^T\hat{\boldsymbol{\beta}}(u)\}-I\{\boldsymbol{Z}^T\hat{\boldsymbol{\alpha}}(v)\le\boldsymbol{Z}^T\hat{\boldsymbol{\beta}}(u)\}]\,\mathrm{d}v\Big)\\
&=\int_0^{\tau_U}[\boldsymbol{w}_1\{\hat{\boldsymbol{\beta}}(u),\boldsymbol{\alpha}_0(v)\}-\boldsymbol{w}_1\{\hat{\boldsymbol{\beta}}(u),\hat{\boldsymbol{\alpha}}(v)\}]\,\mathrm{d}v\\
&\approx\int_0^{\tau_U}-\boldsymbol{w}_{12}\{\boldsymbol{\beta}_0(u),\boldsymbol{\alpha}_0(v)\}\{\hat{\boldsymbol{\alpha}}(v)-\boldsymbol{\alpha}_0(v)\}\,\mathrm{d}v,
\end{aligned}\tag{S3.4}$$

where $\approx$ indicates that the difference converges uniformly to 0 for on $[\nu,\tau_U]$.

With Taylor expansion,we can show that

$$\tilde{\boldsymbol{\mu}}_1\{\hat{\boldsymbol{\beta}}(u),\boldsymbol{\alpha}_0\}-\tilde{\boldsymbol{\mu}}_1\{\boldsymbol{\beta}_0(u),\boldsymbol{\alpha}_0\}\approx\tilde{\boldsymbol{J}}_1\{\boldsymbol{\beta}_0(u),\boldsymbol{\alpha}_0\}\{\hat{\boldsymbol{\beta}}(u)-\boldsymbol{\beta}_0(u)\}. \tag{S3.5}$$

and we also have

$$\boldsymbol{\mu}_1\{\hat{\boldsymbol{\beta}}(\tau)\}-\boldsymbol{\mu}_1\{\boldsymbol{\beta}_0(\tau)\}\approx\boldsymbol{J}_1\{\boldsymbol{\beta}_0(\tau)\}\{\hat{\boldsymbol{\beta}}(\tau)-\boldsymbol{\beta}_0(\tau)\}. \tag{S3.6}$$

Let $\boldsymbol{\psi}_1(\tau)=\hat{\boldsymbol{\beta}}(\tau)-\boldsymbol{\beta}_0(\tau)$ and $\boldsymbol{\psi}_2(\tau)=\hat{\boldsymbol{\alpha}}(\tau)-\boldsymbol{\alpha}_0(\tau)$. From (S3.3),(S3.4),(S3.5) and (S3.6) we can see that

$$\boldsymbol{s}^{(1)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{s}^{(1)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau) \approx \boldsymbol{A}_{01}(\tau)\boldsymbol{\psi}_1(\tau) - \int_0^{\tau}\int_0^{\tau_U} \boldsymbol{B}_{11}(u,v)\boldsymbol{\psi}_2(v)\,\mathrm{d}v\,\mathrm{d}u - \int_0^{\tau_U} \boldsymbol{B}_{21}(v)\boldsymbol{\psi}_1(v)\,\mathrm{d}v, \tag{S3.7}$$

where $\boldsymbol{A}_{01}(\tau) = \boldsymbol{J}_1\{\boldsymbol{\beta}_0(\tau)\}$, $\boldsymbol{B}_{11}(u,v) = -\boldsymbol{w}_{12}\{\boldsymbol{\beta}_0(u), \boldsymbol{\alpha}_0(v)\}$ and $\boldsymbol{B}_{21}(v) = \tilde{\boldsymbol{J}}_1\{\boldsymbol{\beta}_0(v), \boldsymbol{\alpha}_0\}$.

Similarly, we can show that

$$\boldsymbol{s}^{(2)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{s}^{(2)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau) \approx \boldsymbol{A}_{02}(\tau)\boldsymbol{\psi}_2(\tau) - \int_0^{\tau}\int_0^{\tau_U} \boldsymbol{B}_{12}(u,v)\boldsymbol{\psi}_1(v)\,\mathrm{d}v\,\mathrm{d}u - \int_0^{\tau_U} \boldsymbol{B}_{22}(v)\boldsymbol{\psi}_2(v)\,\mathrm{d}v, \tag{S3.8}$$

where $\boldsymbol{A}_{02}(\tau) = \boldsymbol{J}_2\{\boldsymbol{\alpha}_0(\tau)\}$, $\boldsymbol{B}_{12}(u,v) = -\boldsymbol{w}_{21}\{\boldsymbol{\beta}_0(v), \boldsymbol{\alpha}_0(u)\}$ and $\boldsymbol{B}_{22}(v) = \tilde{\boldsymbol{J}}_2\{\boldsymbol{\alpha}_0(v), \boldsymbol{\beta}_0\}$.

Let

$$\boldsymbol{\omega}(\tau) = \begin{pmatrix} \boldsymbol{s}^{(1)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{s}^{(1)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau) \\ \boldsymbol{s}^{(2)}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}, \tau) - \boldsymbol{s}^{(2)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau) \end{pmatrix}, \quad \boldsymbol{\psi}(\tau) = \begin{pmatrix} \boldsymbol{\psi}_1(\tau) \\ \boldsymbol{\psi}_2(\tau) \end{pmatrix},$$

then from (S3.7) and (S3.8) we have

$$\boldsymbol{\omega}(\tau) \approx \boldsymbol{A}_0(\tau)\boldsymbol{\psi}(\tau) - \int_0^{\tau_U} \boldsymbol{A}_1(\tau, v)\boldsymbol{\psi}(v)\,\mathrm{d}v, \tag{S3.9}$$

where

$$\boldsymbol{A}_0(\tau) = \begin{bmatrix} \boldsymbol{A}_{01}(\tau) & 0 \\ 0 & \boldsymbol{A}_{02}(\tau) \end{bmatrix} \text{ and } \boldsymbol{A}_1(\tau, v) = \begin{bmatrix} \boldsymbol{B}_{21}(v) & \int_0^{\tau} \boldsymbol{B}_{11}(u,v)\,\mathrm{d}u \\ \int_0^{\tau} \boldsymbol{B}_{12}(u,v)\,\mathrm{d}u & \boldsymbol{B}_{22}(v) \end{bmatrix},$$

Let

$$\boldsymbol{S}_n(\boldsymbol{b}, \boldsymbol{a}, \tau) = \begin{pmatrix} \boldsymbol{S}_n^{(1)}(\boldsymbol{b}, \boldsymbol{a}, \tau) \\ \boldsymbol{S}_n^{(2)}(\boldsymbol{b}, \boldsymbol{a}, \tau) \end{pmatrix},$$

from (S3.2) and (S3.9) we have

$$-n^{1/2}\boldsymbol{S}_n(\boldsymbol{\beta}_0,\boldsymbol{\alpha}_0,\tau) = \{\boldsymbol{A}_0(\tau)+\boldsymbol{o}_{[\nu,\tau_U]}(1)\}\times n^{1/2}\boldsymbol{\psi}(\tau) - \int_0^{\tau_U}\{\boldsymbol{A}_1(\tau,v)+\boldsymbol{o}_{[\nu,\tau_U]}(1)\}\times n^{1/2}\boldsymbol{\psi}(v)\,\mathrm{d}v. \tag{S3.10}$$

Equation (S3.10) can be viewed as a stochastic differential equation for $n^{1/2}\boldsymbol{\psi}(\tau)$. Specifically, it is a Fredholm integral equation of the second kind and the solution can be presented in the following form (Polyanin and Manzhirov, 2008):

$$n^{1/2}\boldsymbol{\psi}(\tau) = -\boldsymbol{A}_0(\tau)^{-1}\{n^{1/2}\boldsymbol{S}_n(\boldsymbol{\beta}_0,\boldsymbol{\alpha}_0,\tau) - \int_0^{\tau_U}\boldsymbol{R}(\tau,v)\times n^{1/2}\boldsymbol{S}_n(\boldsymbol{\beta}_0,\boldsymbol{\alpha}_0,v)\,\mathrm{d}v\} + \boldsymbol{o}_{[\nu,\tau_U]}(1), \tag{S3.11}$$

where $\boldsymbol{R}(\tau,v)$ is determined by $\boldsymbol{A}_1(\tau,v)$ and $\boldsymbol{A}_0(\tau)$, and independent of $\boldsymbol{\psi}(\tau)$. The detailed solution can be found in Polyanin and Manzhirov (2008) and thus omitted here.

By observing (S3.11) we can see that, to show the weak convergence of $n^{1/2}\{\hat{\boldsymbol{\beta}}(\tau)-\boldsymbol{\beta}_0(\tau)\}$ and $n^{1/2}\{\hat{\boldsymbol{\alpha}}(\tau)-\boldsymbol{\alpha}_0(\tau)\}$, it suffices to show the weak convergence of $-n^{1/2}\boldsymbol{S}_n(\boldsymbol{\beta}_0,\boldsymbol{\alpha}_0,\tau)$.

Let

$$\boldsymbol{N}_i(\boldsymbol{\beta},\boldsymbol{\alpha},\tau) = \begin{pmatrix} N_{1i}[g\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)\}] \\ N_{2i}[g\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}(\tau)\}] \end{pmatrix}$$

and

$$\boldsymbol{K}_i(\boldsymbol{\beta},\boldsymbol{\alpha},\tau) = \begin{pmatrix} \int_0^\tau Y_i[g\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(u)\}]\times\phi_1(1-u, 1-\int_0^{\tau_U} I\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}(v)\le\boldsymbol{Z}_i^T\boldsymbol{\beta}(u)\}\,\mathrm{d}v)\,\mathrm{d}u \\ \int_0^\tau Y_i[g\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}(u)\}]\times\phi_2(1-\int_0^{\tau_U} I\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(v)\le\boldsymbol{Z}_i^T\boldsymbol{\alpha}(u)\}\,\mathrm{d}v, 1-u)\,\mathrm{d}u \end{pmatrix}.$$

First we note that $\{\boldsymbol{Z}_i\boldsymbol{N}_i(\boldsymbol{\beta}_0,\boldsymbol{\alpha}_0,\tau),\tau\in(0,\tau_U]\}$ is a VC-class (van der Vaart and Wellner 1996), and $\boldsymbol{K}_i(\boldsymbol{\beta}_0,\boldsymbol{\alpha}_0,\tau)$ is Lipschitz in τ. It then follows that $\Big\{\boldsymbol{Z}_i\{\boldsymbol{N}_i(\boldsymbol{\beta}_0,\boldsymbol{\alpha}_0,\tau)-\boldsymbol{K}_i(\boldsymbol{\beta}_0,\boldsymbol{\alpha}_0,\tau)\},\tau\in(0,\tau_U]\Big\}$ is a Donsker class by the permanence properties. By the Donsker theorem, $-n^{1/2}\boldsymbol{S}_n(\boldsymbol{\beta}_0,\boldsymbol{\alpha}_0,\tau)$ converges weakly to a Gaussian process, namely $\boldsymbol{G}(\tau)$, with mean 0 and covariance $\boldsymbol{\Sigma}(s,t)$ for $s,t\in(0,\tau_U]$, where $\boldsymbol{\Sigma}(s,t)=E\{\boldsymbol{\iota}_j(s)\boldsymbol{\iota}_j(t)^T\}$

with $\boldsymbol{\iota}_j(\tau) = \boldsymbol{Z}_j\{\boldsymbol{N}_j(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau) - \boldsymbol{K}_j(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)\}, \tau \in (0, \tau_U]$. By this fact, coupled with (S3.11) and condition C3(a), we can see that $n^{1/2}\{\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\}$ converges weakly to a Guassian process for $\tau \in [\nu_1, \tau_U]$, and $n^{1/2}\{\hat{\boldsymbol{\alpha}}(\tau) - \boldsymbol{\alpha}_0(\tau)\}$ also converges weakly to a Guassian process for $\tau \in [\nu_2, \tau_U]$, where $0 < \nu_1, \nu_2 < \tau_U$.

S4 Appendix D: Convergence Criteria for Computing Algorithms

In this section we provide detailed convergence criteria which is shared by the two algorithms A and B proposed in Section 2.3. First, we set the maximum number of iterations, denoted by M_0, and the tolerance level, denoted by tol. At the m–th iteration ($m \geq 1$), define $d_{\beta,q}^{[m]} = \frac{1}{\tau_{U,1}-\nu_1}\int_{\nu_1}^{\tau_{U,1}}[\beta^{[m+1]^{(q)}}(\tau) - \beta^{[m]^{(q)}}(\tau)]\,\mathrm{d}\tau$ and $\tilde{d}_{\beta,q}^{[m]} = \frac{1}{\tau_{U,1}-\nu_1}\int_{\nu_1}^{\tau_{U,1}}[\beta^{[m+1]^{(q)}}(\tau) - \beta^{[m-1]^{(q)}}(\tau)]\,\mathrm{d}\tau$ for $q = 0, \cdots, p$, recalling that $\beta^{(0)}(\tau)$ is the intercept and $\beta^{(q)}(\tau)$ is the coefficient corresponding to the q–th element of $\tilde{\boldsymbol{Z}}$ for $q = 1, \cdots, p$. Let $d_{\beta}^{[m]} = \max_q\{abs(d_{\beta,q}^{[m]})\}$ and $\tilde{d}_{\beta}^{[m]} = \max_q\{abs(\tilde{d}_{\beta,q}^{[m]})\}$, where $abs(\cdot)$ stands for the absolute value function. We also define $d_{\alpha}^{[m]}$ and $\tilde{d}_{\alpha}^{[m]}$ in the similar fashion. At the end of Step A2 (B2) of the m–th iteration, we carry out the following steps:

Step 0. Compare m with M_0. If $m < M_0$ then continue to the next step, otherwise stop and claim non-convergence.

Step 1. If $\max\{d_{\beta}^{[m]}, d_{\alpha}^{[m]}\} < tol$, then announce convergence. Let $\hat{\boldsymbol{\beta}}(\tau) = \hat{\boldsymbol{\beta}}^{[m+1]}(\tau)$ for $\tau \in (0, \tau_{U,1}]$ and $\hat{\boldsymbol{\alpha}}(\tau) = \hat{\boldsymbol{\alpha}}^{[m+1]}(\tau)$ for $\tau \in (0, \tau_{U,2}]$ and stop. Otherwise continue to the next step.

Step 2. If $\max\{\tilde{d}_{\beta}^{[m]}, \tilde{d}_{\alpha}^{[m]}\} < tol$, then announce convergence. Let $\hat{\boldsymbol{\beta}}(\tau) = \{\hat{\boldsymbol{\beta}}^{[m+1]}(\tau) + \hat{\boldsymbol{\beta}}^{[m-1]}(\tau)\}/2$ for $\tau \in (0, \tau_{U,1}]$ and $\hat{\boldsymbol{\alpha}}(\tau) = \{\hat{\boldsymbol{\alpha}}^{[m+1]}(\tau) + \hat{\boldsymbol{\alpha}}^{[m-1]}(\tau)\}/2$ for $\tau \in (0, \tau_{U,2}]$ and stop. Otherwise continue to Step A3 (B3).

In the simulation studies and data analysis reported in Sections 3 and 4, we chose $M_0 = 10$ and $tol = 10^{-2}$.

S5 Appendix E: Additional Tables and Figures

Table E1. Simulation Results on Parameter Estimation under the Frank copula. Bias: biases; AvgSD: average estimated resampling-based standard deviations; EmpSD: empirical standard deviations; Cov95: coverage rates of 95% Wald confidence intervals.

τ		Bias	EmpSD	AvgSD	Cov95		Bias	EmpSD	AvgSD	Cov95
10% indep. censoring, 45% dep. censoring to T, 45% dep. censoring to D, model (2.2) for D										
0.1	$\hat{\beta}^{(0)}$	0.03	0.11	0.12	0.94	$\hat{\alpha}^{(0)}$	0.01	0.10	0.12	0.95
	$\hat{\beta}^{(1)}$	0.02	0.20	0.23	0.96	$\hat{\alpha}^{(1)}$	-0.02	0.16	0.19	0.96
	$\hat{\beta}^{(2)}$	-0.04	0.11	0.12	0.95	$\hat{\alpha}^{(2)}$	0.03	0.11	0.12	0.93
0.3	$\hat{\beta}^{(0)}$	0.01	0.08	0.09	0.95	$\hat{\alpha}^{(0)}$	0.01	0.08	0.10	0.93
	$\hat{\beta}^{(1)}$	0.01	0.16	0.19	0.96	$\hat{\alpha}^{(1)}$	-0.02	0.14	0.15	0.95
	$\hat{\beta}^{(2)}$	-0.01	0.08	0.09	0.95	$\hat{\alpha}^{(2)}$	0.02	0.08	0.09	0.96
0.5	$\hat{\beta}^{(0)}$	0.01	0.07	0.08	0.95	$\hat{\alpha}^{(0)}$	0.02	0.08	0.11	0.96
	$\hat{\beta}^{(1)}$	0.01	0.14	0.19	0.98	$\hat{\alpha}^{(1)}$	-0.03	0.12	0.16	0.97
	$\hat{\beta}^{(2)}$	-0.01	0.07	0.08	0.98	$\hat{\alpha}^{(2)}$	0.01	0.07	0.08	0.97
10% indep. censoring, 30% dep. censoring to T, 60% dep. censoring to D, AFT model for D										
0.1	$\hat{\beta}^{(0)}$	0.01	0.09	0.10	0.96	$\hat{\alpha}^{(0)}$	0.02	0.09	0.09	0.93
	$\hat{\beta}^{(1)}$	0.02	0.17	0.19	0.97	$\hat{\alpha}^{(1)}$	-0.02	0.13	0.14	0.94
	$\hat{\beta}^{(2)}$	-0.02	0.09	0.10	0.96	$\hat{\alpha}^{(2)}$	0.03	0.09	0.09	0.94
0.3	$\hat{\beta}^{(0)}$	0.01	0.07	0.08	0.94	$\hat{\alpha}^{(0)}$	0.01	0.09	0.09	0.94
	$\hat{\beta}^{(1)}$	0.00	0.14	0.15	0.95	$\hat{\alpha}^{(1)}$	-0.02	0.13	0.14	0.94
	$\hat{\beta}^{(2)}$	0.00	0.07	0.07	0.95	$\hat{\alpha}^{(2)}$	0.03	0.09	0.09	0.94
0.5	$\hat{\beta}^{(0)}$	0.00	0.06	0.07	0.96	$\hat{\alpha}^{(0)}$	0.01	0.09	0.09	0.95
	$\hat{\beta}^{(1)}$	0.00	0.12	0.13	0.96	$\hat{\alpha}^{(1)}$	-0.02	0.13	0.14	0.94
	$\hat{\beta}^{(2)}$	0.00	0.06	0.07	0.96	$\hat{\alpha}^{(2)}$	0.03	0.09	0.09	0.94
0.7	$\hat{\beta}^{(0)}$	0.00	0.07	0.08	0.95	$\hat{\alpha}^{(0)}$	0.02	0.09	0.10	0.95
	$\hat{\beta}^{(1)}$	0.00	0.12	0.17	0.97	$\hat{\alpha}^{(1)}$	-0.02	0.13	0.14	0.94
	$\hat{\beta}^{(2)}$	0.00	0.07	0.08	0.97	$\hat{\alpha}^{(2)}$	0.03	0.09	0.09	0.94

Table E2. Simulation Results on Parameter Estimation when the Copula Function is Misspecified. Bias: biases; AvgSD: average estimated resampling-based standard deviations; EmpSD: empirical standard deviations; Cov95: coverage rates of 95% Wald confidence intervals.

τ		Bias	EmpSD	AvgSD	Cov95		Bias	EmpSD	AvgSD	Cov95
	Underlying: Clayton, Kendall's tau=0.58; Assumed: Frank, Kendall's tau=0.58									
0.1	$\hat{\beta}^{(0)}$	-0.01	0.09	0.10	0.94	$\hat{\alpha}^{(0)}$	0.01	0.08	0.09	0.97
	$\hat{\beta}^{(1)}$	-0.01	0.17	0.18	0.95	$\hat{\alpha}^{(1)}$	-0.01	0.13	0.14	0.95
	$\hat{\beta}^{(2)}$	0.01	0.09	0.10	0.95	$\hat{\alpha}^{(2)}$	-0.02	0.09	0.10	0.94
0.3	$\hat{\beta}^{(0)}$	-0.01	0.07	0.08	0.95	$\hat{\alpha}^{(0)}$	0.00	0.08	0.09	0.96
	$\hat{\beta}^{(1)}$	-0.01	0.13	0.15	0.96	$\hat{\alpha}^{(1)}$	-0.01	0.13	0.14	0.95
	$\hat{\beta}^{(2)}$	0.02	0.07	0.08	0.95	$\hat{\alpha}^{(2)}$	-0.02	0.09	0.10	0.94
0.5	$\hat{\beta}^{(0)}$	0.00	0.07	0.08	0.95	$\hat{\alpha}^{(0)}$	0.01	0.08	0.10	0.97
	$\hat{\beta}^{(1)}$	0.01	0.13	0.15	0.96	$\hat{\alpha}^{(1)}$	-0.01	0.13	0.14	0.95
	$\hat{\beta}^{(2)}$	0.00	0.07	0.07	0.95	$\hat{\alpha}^{(2)}$	-0.02	0.09	0.10	0.94
0.7	$\hat{\beta}^{(0)}$	0.00	0.07	0.08	0.96	$\hat{\alpha}^{(0)}$	0.07	0.10	0.11	0.94
	$\hat{\beta}^{(1)}$	0.05	0.14	0.18	0.97	$\hat{\alpha}^{(1)}$	-0.01	0.13	0.14	0.95
	$\hat{\beta}^{(2)}$	-0.02	0.07	0.08	0.95	$\hat{\alpha}^{(2)}$	-0.02	0.09	0.10	0.94
	Underlying: Frank, Kendall's tau=0.58; Assumed: Clayton, Kendall's tau=0.58									
0.1	$\hat{\beta}^{(0)}$	0.04	0.09	0.10	0.93	$\hat{\alpha}^{(0)}$	0.04	0.09	0.10	0.92
	$\hat{\beta}^{(1)}$	0.03	0.16	0.19	0.96	$\hat{\alpha}^{(1)}$	-0.04	0.14	0.14	0.93
	$\hat{\beta}^{(2)}$	-0.05	0.09	0.10	0.93	$\hat{\alpha}^{(2)}$	0.07	0.08	0.09	0.86
0.3	$\hat{\beta}^{(0)}$	0.03	0.07	0.08	0.93	$\hat{\alpha}^{(0)}$	0.03	0.09	0.09	0.93
	$\hat{\beta}^{(1)}$	0.01	0.13	0.15	0.96	$\hat{\alpha}^{(1)}$	-0.04	0.14	0.14	0.93
	$\hat{\beta}^{(2)}$	-0.02	0.07	0.07	0.96	$\hat{\alpha}^{(2)}$	0.07	0.08	0.09	0.86
0.5	$\hat{\beta}^{(0)}$	0.01	0.07	0.07	0.95	$\hat{\alpha}^{(0)}$	0.01	0.09	0.09	0.95
	$\hat{\beta}^{(1)}$	-0.01	0.12	0.14	0.96	$\hat{\alpha}^{(1)}$	-0.04	0.14	0.14	0.93
	$\hat{\beta}^{(2)}$	0.00	0.06	0.07	0.96	$\hat{\alpha}^{(2)}$	0.07	0.08	0.09	0.86
0.7	$\hat{\beta}^{(0)}$	0.00	0.07	0.08	0.94	$\hat{\alpha}^{(0)}$	-0.01	0.09	0.10	0.95
	$\hat{\beta}^{(1)}$	-0.03	0.13	0.15	0.96	$\hat{\alpha}^{(1)}$	-0.04	0.14	0.14	0.93
	$\hat{\beta}^{(2)}$	0.02	0.07	0.08	0.96	$\hat{\alpha}^{(2)}$	0.07	0.08	0.09	0.86

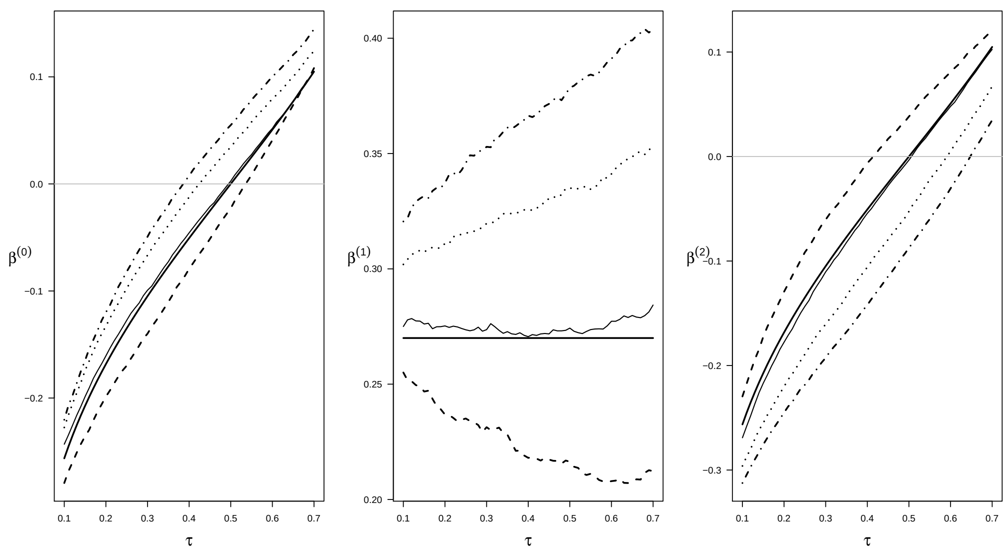

Figure E1. Simulation Results on Estimates for $\boldsymbol{\beta}_0(\tau)$ under the Correctly Specified Clayton Copula with Misspecified Association Parameters: Kendall's tau= 0.79 (Dashed Lines), Kendall's tau= 0.33 (dotted Lines), and Kendall's tau= 0.16 (Dotdash Lines), and the True Association Parameter: Kendall's tau= 0.58 (Solid Lines); and the True Coefficients $\boldsymbol{\beta}_0(\tau)$ (Bold Solid Lines).

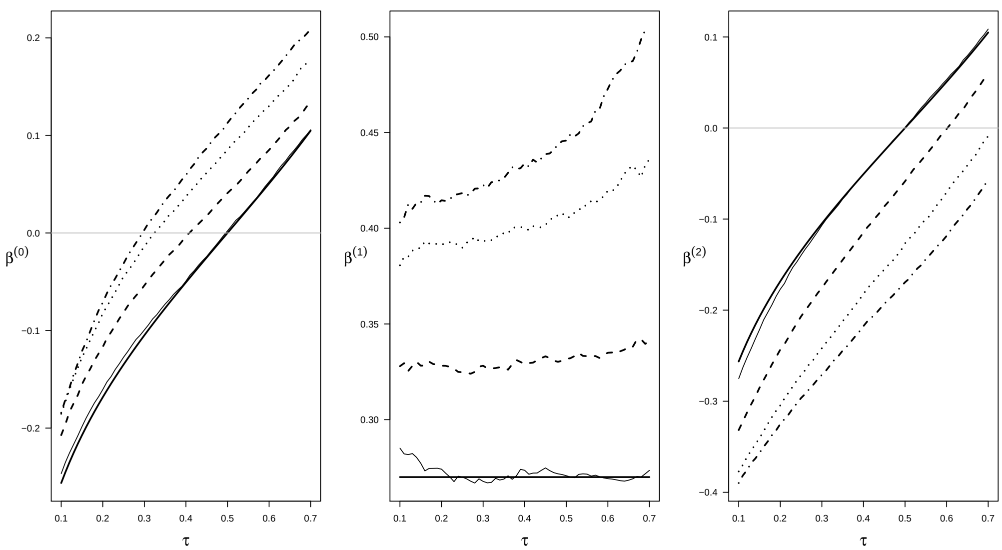

Figure E2. Simulation Results on Estimates for $\boldsymbol{\beta}_0(\tau)$ under the Correctly Specified Frank Copula with Misspecified Association Parameters: Kendall's tau= 0.26 (Dashed Lines), Kendall's tau= -0.12 (Dotted Lines), and Kendall's tau= -0.33 (Dotdash Lines), and the True Association Parameter: Kendall's tau= 0.58 (Solid Lines); and the True Coefficients $\boldsymbol{\beta}_0(\tau)$ (Bold Solid Lines).

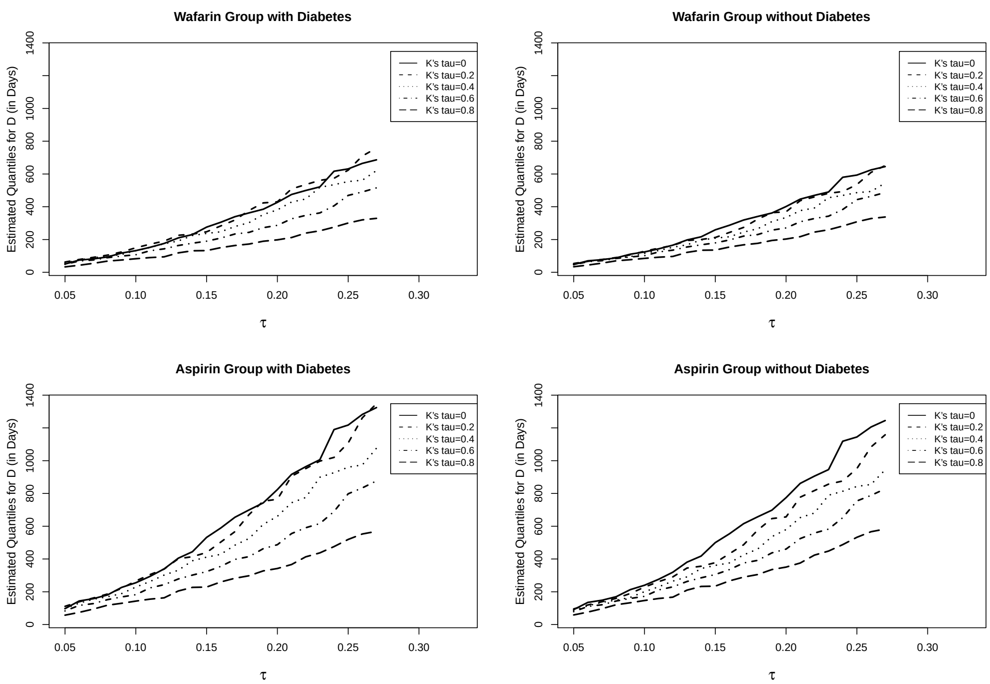

Figure E3. The WASID Example: Estimated Quantiles of Time to Early Termination of Study Medication under the Clayton Copula with Kendall's tau=0, 0.2, 0.4, 0.6 and 0.8, with the Stenosis Percentage Fixed at Its Mean (63.7%).