Estimation and Inference of Quantile Regression for Survival Data under Biased Sampling

Supplementary Materials: Proofs of the Main Results

S1 Verification of the weight function $v_i(t)$ for the length-biased sampling scheme

Here, we provide the justification of (11) which is introduced in Section 3. Observe that

$$
\begin{aligned}
f_{\tilde{T},\Delta}(t, \Delta = 1 \mid A, \mathbf{Z}) &= P\{\tilde{T} \in (t, t+dt), \Delta = 1 \mid A, \mathbf{Z}\}/dt \\
&= I(t \geq A) f_{A,V}(A, t-A \mid \mathbf{Z}) S_c(t-A \mid \mathbf{Z}) \\
&= I(t \geq A) f(t \mid \mathbf{Z}) S_c(t-A \mid \mathbf{Z}) \times \frac{1}{\mu(\mathbf{Z})} \\
&= I(t \geq A) P(T^* \in (t, t+dt), C^* > t \mid \mathbf{Z})/dt \times \frac{1}{\mu(\mathbf{Z})} \\
&= I(t \geq A) f_{\tilde{T}^*,\Delta^*}(t, \Delta = 1 \mid \mathbf{Z}) \times \frac{1}{\mu(\mathbf{Z})}, \\
f_{\tilde{T},\Delta}(t, \Delta = 0 \mid A, \mathbf{Z}) &= P(\tilde{T} \in (t, t+dt), \Delta = 0 \mid A, \mathbf{Z})/dt \\
&= I(t \geq A) \int_t^\infty f_{A,V}(A, s-A|\mathbf{Z}) ds \times g_c(t-A|\mathbf{Z}) \\
&= I(t \geq A) \int_t^\infty f(s \mid \mathbf{Z}) ds \times g_c(t-A \mid \mathbf{Z}) \frac{1}{\mu(\mathbf{Z})} \\
&= I(t \geq A) P\{T^* > t, C^* \in (t, t+dt) \mid \mathbf{Z}\}/dt \times \frac{1}{\mu(\mathbf{Z})} \\
&= I(t \geq A) f_{\tilde{T}^*,\Delta^*}(t, \Delta = 0 \mid \mathbf{Z}) \times \frac{1}{\mu(\mathbf{Z})}.
\end{aligned}
$$

It follows that that

$$v_i(t) = \frac{f_{\tilde{T}^*,\Delta^*}(\tilde{T}_i, \Delta_i \mid \mathbf{Z}_i)}{f_{\tilde{T},\Delta}(\tilde{T}_i, \Delta_i \mid \mathbf{Z}_i)} \times \frac{f_{\tilde{T},\Delta}(t, 1 \mid \mathbf{Z}_i)}{f_{\tilde{T}^*,\Delta^*}(t, 1 \mid \mathbf{Z}_i)} = I(A_i \leq t).$$

S2 Proof of Theorems 1 and 2

For the biased samplings introduced in Sections 2.2 and 3, let $f_{\tilde{T}}(t \mid \mathbf{Z})$ be the conditional density function of $\tilde{T}$, i.e., $f_{\tilde{T}}(t \mid \mathbf{Z}) = \sum_{\delta \in \{0,1\}} f_{\tilde{T},\Delta}(t, \delta \mid \mathbf{Z})$, where $f_{\tilde{T},\Delta}(t, \delta \mid \mathbf{Z})$ is defined in (8) of Section 2.2. For a vector a, let $a^{\otimes 2}$ denote $aa^\top$, and $\|a\|$ denote the Euclidean norm of a. For $\mathbf{b} \in \mathbb{R}^p$, define

$$\begin{aligned}
m(\mathbf{b}) &= E\left\{\mathbf{Z}N(e^{\mathbf{Z}^\top \mathbf{b}})\right\}, \\
\tilde{m}(\mathbf{b}) &= E\left\{\mathbf{Z}v(e^{\mathbf{Z}^\top \mathbf{b}})Y(e^{\mathbf{Z}^\top \mathbf{b}})\right\}, \\
m_n(\mathbf{b}) &= \frac{1}{n}\sum_{i=1}^{n}\left\{\mathbf{Z}_i N_i(e^{\mathbf{Z}_i^\top \mathbf{b}})\right\}, \\
\tilde{m}_n(\mathbf{b}) &= \frac{1}{n}\sum_{i=1}^{n}\left\{\mathbf{Z}_i v_i(e^{\mathbf{Z}_i^\top \mathbf{b}})Y_i(e^{\mathbf{Z}_i^\top \mathbf{b}})\right\}, \\
\mathbf{B}(\mathbf{b}) &= E\left\{\mathbf{Z}^{\otimes 2} f_{\tilde{T},\Delta}(e^{\mathbf{Z}^\top \mathbf{b}}, 1 \mid \mathbf{Z}) \exp(\mathbf{Z}^\top \mathbf{b})\right\}, \\
\mathbf{J}(\mathbf{b}) &= -E\left\{\mathbf{Z}^{\otimes 2} v(e^{\mathbf{Z}^\top \mathbf{b}}) f_{\tilde{T}}(e^{\mathbf{Z}^\top \mathbf{b}} \mid \mathbf{Z}) \exp(\mathbf{Z}^\top \mathbf{b})\right\}.
\end{aligned}$$

For $d > 0$, define the set $\mathcal{B}(d)$ as

$$\mathcal{B}(d) = \left\{\mathbf{b} \in \mathbb{R}^p : \inf_{\tau \in (0, \tau_u]} \|m(\mathbf{b}) - m(\boldsymbol{\beta}_0(\tau))\| \leq d\right\},$$

where $\boldsymbol{\beta}_0(\tau)$ is the true parameter value, $\tau \in (0, \tau_u]$; and τ_u satisfies Condition C4 below.

Furthermore, consider the setting in Section 3.2. For $\mathbf{b} \in \mathbb{R}^p$, define

$$
\begin{aligned}
m^*(\mathbf{b}) &= E\left\{\mathbf{Z} \otimes \boldsymbol{\psi}(\tilde{T}) N(e^{\mathbf{Z}^\top \mathbf{b}})\right\}, \\
\tilde{m}^*(\mathbf{b}) &= E\left\{\mathbf{Z} \otimes \boldsymbol{\psi}(e^{\mathbf{Z}^\top \mathbf{b}}) v(e^{\mathbf{Z}^\top \mathbf{b}}) Y(e^{\mathbf{Z}^\top \mathbf{b}})\right\}, \\
m_n^*(\mathbf{b}) &= \frac{1}{n}\sum_{i=1}^n \mathbf{Z}_i \otimes \boldsymbol{\psi}(\tilde{T}_i) N_i(e^{\mathbf{Z}_i^\top \mathbf{b}}), \\
\tilde{m}_n^*(\mathbf{b}) &= \frac{1}{n}\sum_{i=1}^n \mathbf{Z}_i \otimes \boldsymbol{\psi}(e^{\mathbf{Z}_i^\top \mathbf{b}}) v_i(e^{\mathbf{Z}_i^\top \mathbf{b}}) Y_i(e^{\mathbf{Z}_i^\top \mathbf{b}}), \\
\mathbf{B}^*(\mathbf{b}) &= E\left\{[\mathbf{Z} \otimes \boldsymbol{\psi}(e^{\mathbf{Z}^\top \mathbf{b}})]\mathbf{Z}^\top f_{\tilde{T},\Delta}(e^{\mathbf{Z}^\top \mathbf{b}}, 1 \mid \mathbf{Z}) \exp(\mathbf{Z}^\top \mathbf{b})\right\}, \\
\mathbf{J}^*(\mathbf{b}) &= -E\left\{[\mathbf{Z} \otimes \boldsymbol{\psi}(e^{\mathbf{Z}^\top \mathbf{b}})]\mathbf{Z}^\top v(e^{\mathbf{Z}^\top \mathbf{b}}) f_{\tilde{T}}(e^{\mathbf{Z}^\top \mathbf{b}} \mid \mathbf{Z}) \exp(\mathbf{Z}^\top \mathbf{b})\right\}.
\end{aligned}
$$

We assume the following conditions for theoretical derivation.

C1: $\mathbf{Z}$ is uniformly bounded, i.e., $\sup_i \|\mathbf{Z}_i\| < \infty$ and the matrix $E(\mathbf{Z}^{\otimes 2})$ is positive definite.

C2: $m\{\boldsymbol{\beta}(\tau)\}$ is a Lipschitz continuous function of $\tau \in (0, \tau_u]$, and $f_{\tilde{T}}(t \mid \mathbf{z})$ and $f_{\tilde{T},\Delta}(t, 1 \mid \mathbf{z})$ are bounded above and continuous uniformly in t and $\mathbf{z}$.

C3: There exists $d_0 > 0$ such that for $\mathbf{b} \in \mathcal{B}(d_0)$ and any $\mathbf{Z}$, $f_{\tilde{T},\Delta}(e^{\mathbf{Z}^\top \mathbf{b}}, 1 \mid \mathbf{Z}) > 0$ and $\|\mathbf{J}(\mathbf{b})\mathbf{B}(\mathbf{b})^{-1}\|$ is uniformly bounded.

C4: $\inf_{\tau \in [\tau_l, \tau_u]} \operatorname{eigmin} \mathbf{B}\{\boldsymbol{\beta}_0(\tau)\} > 0$ for any $\tau_l \in (0, \tau_u]$, where $\operatorname{eigmin}(\cdot)$ denotes the minimum eigenvalue of a matrix.

C5: The class of weight functions $\{(Z, A, \tilde{T}, \Delta) \to v(e^{\mathbf{Z}^\top \mathbf{b}}, A, \tilde{T}, \Delta, Z); \mathbf{b} \in \mathbb{R}^d\}$ is Glivenko-Cantelli and uniformly bounded; and $\{(Z, A, \tilde{T}, \Delta) \to v(e^{\mathbf{Z}^\top \mathbf{b}}, A, \tilde{T}, \Delta, Z); \mathbf{b} \in \mathcal{B}(d_0)\}$ belongs to a Donsker class.

C6: The weight function $\boldsymbol{\psi}(t)$ is positive, bounded and differentiable; $\{(Z, A, \tilde{T}, \Delta) \to \boldsymbol{\psi}(e^{\mathbf{Z}^\top \mathbf{b}}, A, \tilde{T}, \Delta, Z); \mathbf{b} \in \mathbb{R}^d\}$ is Glivenko-Cantelli and $\{(Z, A, \tilde{T}, \Delta) \to \boldsymbol{\psi}(e^{\mathbf{Z}^\top \mathbf{b}}, A, \tilde{T}, \Delta, Z);$ $\mathbf{b} \in \mathcal{B}(d_0)\}$ belongs to a Donsker class. Matrix $W(\boldsymbol{\beta}_0, \tau)$ is nonsingular, $\inf_{t \in [\tau_l, \tau]} \operatorname{eigmin} \mathbf{B}^*\{\boldsymbol{\beta}_0(t)\}^\top$

$W(\boldsymbol{\beta}_0,\tau)^{-1}\mathbf{B}^*\{\boldsymbol{\beta}_0(t)\} > 0$ for any $\tau_l \in (0,\tau]$ and $\|\mathbf{J}^*(\mathbf{b})\{\mathbf{B}^*(\mathbf{b})^\top W(\boldsymbol{\beta}_0,\tau)^{-1}\mathbf{B}^*(\mathbf{b})\}^{-1}\|$ is uniformly bounded for $\mathbf{b} \in \mathcal{B}(d_0)$.

Conditions C1–C4 are mild assumptions concerning the covariates $\mathbf{Z}$, the underlying regression quantile parameter process $\boldsymbol{\beta}(\tau)$, and the density functions associated with the observed data $(\tilde{T},\Delta)$. The boundedness assumption of covariates in C1 are often assumed in survival models although it can be further relaxed with extra technical complexity. For Conditions C2 and C3, the boundedness and uniform continuity assumption of $f_{\tilde{T}}(t \mid \mathbf{z})$ and $f_{\tilde{T},\Delta}(t,1 \mid \mathbf{z})$ is reasonable in many biased sampling problems. For instance, it is satisfied for case-cohort study if the density functions of survival and censoring times are bounded and uniformly continuous. Similarly, for the length-biased sampling introduced in Section 3, this is satisfied if the density functions of T and $\tilde{C}$ are bounded and uniformly continuous. Under these conditions, it can been verified that $m(\mathbf{b})$ and $\tilde{m}(\mathbf{b})$ are differentiable and $\mathbf{B}(\mathbf{b}) = \partial m(\mathbf{b})/\partial \mathbf{b}$ and $\mathbf{J}(\mathbf{b}) = \partial \tilde{m}(\mathbf{b})/\partial \mathbf{b}$ are well defined. As in Peng and Huang (2008), Condition C4 is assumed to ensure the identifiability of $\boldsymbol{\beta}_0(\tau)$. The regularity of the weight function assumed in Condition C5 is satisfied for all the sampling schemes introduced in Section 3. In particular, for the case-cohort type designs (Examples 3 and 4), the weight function v does not depend on $\boldsymbol{\beta}$ and satisfies C5. For length-biased sampling with weight function in the form (14), $\{v(e^{\mathbf{Z}^\top \mathbf{b}}), \mathbf{b} \in \mathcal{B}(d_0)\}$ is a VC class, which implies C5 (Theorems 2.6.7 and 2.5.2 in van der Vaart and Wellner, 1996). Condition C6 is similar as conditions C3-C5 and is satisfied for many weight functions. It is assumed to ensure the asymptotic normality of the efficient estimator.

Proof of Theorem 1 The proof follows closely that in Peng and Huang (2008) and we only present the key steps. We should note that, however, their results cannot be directly applied due to the associated random weight functions in the estimating equation. Let $\boldsymbol{\alpha}_0(\tau) = m\{\boldsymbol{\beta}_0(\tau)\}$, $\hat{\boldsymbol{\alpha}}(\tau) = m\{\hat{\boldsymbol{\beta}}(\tau)\}$, and $\mathcal{A}(d) = \{m(\mathbf{b}) : \mathbf{b} \in \mathcal{B}(d)\}$. Following the argument in Peng and Huang (2008), m is a one-to-one mapping from $\mathcal{B}(d_0)$ to $\mathcal{A}(d_0)$. There exists an inverse function $\kappa : \mathcal{B}(d_0) \to \mathcal{A}(d_0)$ such that $\kappa\{m(\mathbf{b})\} = \mathbf{b}$, for any $\mathbf{b} \in \mathcal{B}(d_0)$.

Uniform Convergence Let $\xi_{n,k} = m_n\{\hat{\boldsymbol\beta}(\tau_k)\} - \int_0^{\tau_k} \tilde{m}_n\{\hat{\boldsymbol\beta}(u)\}dH(u)$. According to the estimating equation taking the general form of (10), we have $\sup_k \|\xi_{n,k}\| = O(1)\sup_i \|\mathbf{Z}_i\|/n = O(n^{-1})$. From $m\{\boldsymbol\beta_0(\tau_k)\} - \int_0^{\tau_k} \tilde{m}\{\boldsymbol\beta_0(s)\}dH(s) = 0$, we have the following decomposition:

$$
\begin{aligned}
& m\{\hat{\boldsymbol\beta}(\tau_k)\} - m\{\boldsymbol\beta_0(\tau_k)\} \\
= & -\Big[m_n\{\hat{\boldsymbol\beta}(\tau_k)\} - m\{\hat{\boldsymbol\beta}(\tau_k)\}\Big] + \sum_{j=0}^{k-1}\int_{\tau_j}^{\tau_{j+1}} \Big[\tilde{m}\{\hat{\boldsymbol\beta}(s)\} - \tilde{m}\{\boldsymbol\beta_0(s)\}\Big] dH(s) \\
& + \int_0^{\tau_k} \Big[\tilde{m}_n\{\hat{\boldsymbol\beta}(s)\} - \tilde{m}\{\hat{\boldsymbol\beta}(s)\}\Big] dH(s) + \xi_{n,k}. \qquad \text{(S1)}
\end{aligned}
$$

Under the condition that $\mathbf{Z}$ is uniformly bounded, $\{\mathbf{Z}_i N(e^{\mathbf{Z}_i^\top \mathbf{b}}); \mathbf{b} \in \mathbb{R}^p\}$ and $\{\mathbf{Z}_i Y_i(e^{\mathbf{Z}_i^\top \mathbf{b}}); \mathbf{b} \in \mathbb{R}^p\}$ are Glivenko-Cantelli. By Condition C5 and Theorem 3 in van der Vaart and Wellner (2000), $\{\mathbf{Z}_i v_i(e^{\mathbf{Z}_i^\top \mathbf{b}})Y_i(e^{\mathbf{Z}_i^\top \mathbf{b}}); \mathbf{b} \in \mathbb{R}^p\}$ is also Glivenko-Cantelli. Then, we have the following results: $\sup_{\mathbf{b}\in\mathbb{R}^p} \|m(\mathbf{b}) - m_n(\mathbf{b})\| \to 0$ and $\sup_{\mathbf{b}\in\mathbb{R}^p} \|\tilde{m}(\mathbf{b}) - \tilde{m}_n(\mathbf{b})\| \to 0$ almost surely. This implies that the first and third terms in (S1) are ignorable. Denote $c_{n,0} := \sup_k \| -[m_n\{\hat{\boldsymbol\beta}(\tau_k)\} - m\{\hat{\boldsymbol\beta}(\tau_k)\}] + \int_0^{\tau_k}[\tilde{m}_n\{\hat{\boldsymbol\beta}(s)\} - \tilde{m}\{\hat{\boldsymbol\beta}(s)\}]dH(s)\|$, and it follows that $c_{n,0} = o_p(1)$.

Under condition C2, there exists c_1 such that $\|m\{\boldsymbol\beta_0(\tau)\} - m\{\boldsymbol\beta_0(\tau')\}\| < c_1|\tau - \tau'|$ for any $\tau, \tau' \in (0, \tau_u]$. For $0 = \tau_0 \le \tau < \tau_1$, since $m\{\hat{\boldsymbol\beta}(0)\} = 0$, we have $\sup_{\tau_0\le\tau<\tau_1} \|m\{\hat{\boldsymbol\beta}(\tau)\} - m\{\boldsymbol\beta_0(\tau)\}\| = \sup_{\tau_0\le\tau<\tau_1} \|m\{\boldsymbol\beta_0(\tau)\}\| \le b_{n,0} := c_1\|\mathcal{S}_{L(n)}\|$. Then for large enough n, we know $b_{n,0} < d_0$ and $\hat{\boldsymbol\beta}_n(\tau) \in \mathcal{B}(d_0)$ for $\tau \in [\tau_0, \tau_1)$. Thus, we can write $\sup_{\tau_0\le\tau<\tau_1} \|\tilde{m}\{\hat{\boldsymbol\beta}(\tau)\} - \tilde{m}\{\boldsymbol\beta_0(\tau)\}\| = \sup_{\tau_0\le\tau<\tau_1} \|\tilde{m}[\kappa\{\hat{\boldsymbol\alpha}(\tau)\}] - \tilde{m}[\kappa\{\boldsymbol\alpha(\tau)\}]\|$. Under Condition C3, there exists $\tilde{\boldsymbol\alpha}(\tau)$ between $\boldsymbol\alpha_0(\tau)$ and $\hat{\boldsymbol\alpha}(\tau)$ such that

$$
\sup_{\tau_0\le\tau<\tau_1} \|\tilde{m}\{\hat{\boldsymbol\beta}(\tau)\} - \tilde{m}\{\boldsymbol\beta_0(\tau)\}\| = \sup_{\tau_0\le\tau<\tau_1} \|J[\kappa\{\tilde{\boldsymbol\alpha}(\tau)\}]B[\kappa\{\tilde{\boldsymbol\alpha}(\tau)\}]^{-1}\{\tilde{\boldsymbol\alpha}(\tau) - \boldsymbol\alpha_0(\tau)\}\| \le c_2 b_{n,0},
$$

where c_2 is some constant. Next consider $\tau_1 \leq \tau < \tau_2$. We have

$$\sup_{\tau_1 \leq \tau < \tau_2} \|m\{\hat{\boldsymbol{\beta}}(\tau)\} - m\{\boldsymbol{\beta}_0(\tau)\}\| \leq \|m\{\hat{\boldsymbol{\beta}}(\tau_1)\} - m\{\boldsymbol{\beta}_0(\tau_1)\}\| + \sup_{\tau_1 \leq \tau < \tau_2} \|m\{\boldsymbol{\beta}_0(\tau_1)\} - m\{\boldsymbol{\beta}_0(\tau)\}\|.$$

We know $\sup_{\tau_1 \leq \tau < \tau_2} \|m\{\boldsymbol{\beta}_0(\tau_1)\} - m\{\boldsymbol{\beta}_0(\tau)\}\| \leq c_1 \|\mathcal{S}_{L(n)}\|$. Further from the decomposition (S1) that gives an upper bound for $\|m\{\hat{\boldsymbol{\beta}}(\tau_1)\} - m\{\boldsymbol{\beta}_0(\tau_1)\}\|$, we have:

$$\sup_{\tau_1 \leq \tau < \tau_2} \|m\{\hat{\boldsymbol{\beta}}(\tau)\} - m\{\boldsymbol{\beta}_0(\tau)\}\| \leq b_{n,1} := c_{n,0} + c_3 n^{-1} + c_2 b_{n,0} (1-\tau_u)^{-1} \mathcal{S}_{L(n)} + c_1 \|\mathcal{S}_{L(n)}\|,$$

where c_3 is some big constant. Note that $b_{1,n} \to 0$ and it is smaller than d_0 for large enough n. This implies $\hat{\boldsymbol{\beta}}_n(\tau) \in \mathcal{B}(d_0)$ for $\tau \in [\tau_1, \tau_2)$. Then an induction argument similarly to the Proof of Theorem 1 in Peng and Huang (2008) gives

$$\sup_{\tau_k \leq \tau \leq \tau_{k+1}} \|m\{\hat{\boldsymbol{\beta}}(\tau)\} - m\{\boldsymbol{\beta}_0(\tau)\}\| \leq b_{n,k} := c_{n,0} + c_3 n^{-1} + c_2 \sum_{i=1}^{k-1} b_{n,i} (1-\tau_u)^{-1} \mathcal{S}_{L(n)} + c_1 \|\mathcal{S}_{L(n)}\|.$$

This implies that $\sup_{\tau_0 \leq \tau \leq \tau_u} \|m\{\hat{\boldsymbol{\beta}}(\tau)\} - m\{\boldsymbol{\beta}_0(\tau)\}\| \xrightarrow{p} 0$ and for $\tau_0 \leq \tau \leq \tau_u$ and large enough n, $\hat{\boldsymbol{\beta}}(\tau) \in \mathcal{B}(d_0)$. Thus $\sup_\tau \|\hat{\boldsymbol{\alpha}}(\tau) - \boldsymbol{\alpha}_0(\tau)\| = \sup_\tau \|m\{\hat{\boldsymbol{\beta}}(\tau)\} - m\{\boldsymbol{\beta}_0(\tau)\}\| \xrightarrow{p} 0$. Take Taylor's expansion of $m\{\hat{\boldsymbol{\beta}}(\tau)\}$ around $\boldsymbol{\alpha}_0(\tau)$, and we have

$$\sup_{\tau \in [\tau_l, \tau_u]} \|\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\| \leq \sup_{\tau \in [\tau_l, \tau_u]} \|\mathbf{B}\{\boldsymbol{\beta}_0(\tau)\}^{-1}\{\hat{\boldsymbol{\alpha}}(\tau) - \boldsymbol{\alpha}_0(\tau)\}\| + \sup_{\tau \in [\tau_l, \tau_u]} \|\epsilon_n^*(\tau)\|,$$

where $\epsilon_n^*(\tau)$ is the remainder term of the Taylor expansion and $\sup_{\tau \in [\tau_l, \tau_u]} \|\epsilon_n^*(\tau)\| \to 0$. By Condition C4, we have $\sup_{\tau \in [\tau_l, \tau_u]} \|\mathbf{B}\{\boldsymbol{\beta}_0(\tau)\}^{-1}\{\hat{\boldsymbol{\alpha}}(\tau) - \boldsymbol{\alpha}_0(\tau)\}\| = O\{\hat{\boldsymbol{\alpha}}(\tau) - \boldsymbol{\alpha}_0(\tau)\}$ and this implies the uniform consistency.

Weak Convergence Following Lemma B.1 of Peng and Huang (2008), we have

$$\sup_{\tau\in(0,\tau_u]}\left\|n^{-1/2}\sum_{i=1}^{n}\mathbf{Z}_i\Big\{N_i(e^{\mathbf{Z}_i^\top\hat{\boldsymbol{\beta}}(\tau)})-N_i(e^{\mathbf{Z}_i^\top\boldsymbol{\beta}_0(\tau)})\Big\}-n^{1/2}\Big[m\{\hat{\boldsymbol{\beta}}(\tau)\}-m\{\boldsymbol{\beta}_0(\tau)\}\Big]\right\|=o_p(1) \tag{S2}$$

and

$$\sup_{\tau\in(0,\tau_u]}\Bigg\|n^{-1/2}\sum_{i=1}^{n}\mathbf{Z}_i\Big\{I\Big(\tilde{T}_i\geq e^{\mathbf{Z}_i^\top\hat{\boldsymbol{\beta}}(\tau)}\Big)-I\Big(\tilde{T}_i\geq e^{\mathbf{Z}_i^\top\boldsymbol{\beta}_0(\tau)}\Big)\Big\}$$
$$-n^{1/2}\Big[\tilde{m}\{\hat{\boldsymbol{\beta}}(\tau)\}-\tilde{m}\{\boldsymbol{\beta}_0(\tau)\}\Big]\Bigg\|=o_p(1). \tag{S3}$$

In addition, for the grid size chosen in Theorem 1, $S_n(\hat{\boldsymbol{\beta}},\tau)=o_p(1)$ uniformly in $\tau\in(0,\tau_u]$. Then, the following equations hold uniformly in $\tau\in(0,\tau_u]$,

$$\begin{aligned}-S_n(\boldsymbol{\beta}_0,\tau)=&n^{1/2}\Big[m\{\hat{\boldsymbol{\beta}}(\tau)\}-m\{\boldsymbol{\beta}_0(\tau)\}\Big]-\int_0^\tau n^{1/2}\Big[\tilde{m}\{\hat{\boldsymbol{\beta}}(u)\}-\tilde{m}\{\boldsymbol{\beta}_0(u)\}\Big]\,dH(u)+o_p(1)\\ =&n^{1/2}\Big[m\{\hat{\boldsymbol{\beta}}(\tau)\}-m\{\boldsymbol{\beta}_0(\tau)\}\Big]\\ &-\int_0^\tau n^{1/2}\left[\mathbf{J}\{\boldsymbol{\beta}_0(u)\}\mathbf{B}\{\boldsymbol{\beta}_0(u)\}^{-1}+o_p(1)\right]\Big[m\{\hat{\boldsymbol{\beta}}(u)\}-m\{\boldsymbol{\beta}_0(u)\}\Big]\,dH(u)+o_p(1).\end{aligned}$$

A discrete version of the above equation

$$-S_n(\boldsymbol{\beta}_0,\tau)=n^{1/2}\Big[m_n\{\hat{\boldsymbol{\beta}}(\tau)\}-m_n\{\boldsymbol{\beta}_0(\tau)\}\Big]-\int_0^\tau n^{1/2}\Big[\tilde{m}_n\{\hat{\boldsymbol{\beta}}(u)\}-\tilde{m}_n\{\boldsymbol{\beta}_0(u)\}\Big]\,dH(u)+o_p(1)$$

leads to the following recursive formula:

$$
\begin{aligned}
& - \{S_n(\boldsymbol{\beta}_0, \tau_k) - S_n(\boldsymbol{\beta}_0, \tau_{k-1})\} \\
= {} & n^{1/2} \left[m_n\{\hat{\boldsymbol{\beta}}(\tau_k)\} - m_n\{\boldsymbol{\beta}_0(\tau_k)\} \right] - n^{1/2} \left[\tilde{m}_n\{\hat{\boldsymbol{\beta}}(\tau_{k-1})\} - \tilde{m}_n\{\boldsymbol{\beta}_0(\tau_{k-1})\} \right] \{H(\tau_k) - H(\tau_{k-1})\} \\
& - n^{1/2} \left\{ m_n(\hat{\boldsymbol{\beta}}(\tau_{k-1})) - m_n(\boldsymbol{\beta}_0(\tau_{k-1})) \right\} \\
= {} & \mathbf{B}\{\boldsymbol{\beta}_0(\tau_k)\} n^{1/2} \{\hat{\boldsymbol{\beta}}(\tau_k) - \boldsymbol{\beta}_0(\tau_k)\} \\
& - [\mathbf{B}\{\boldsymbol{\beta}_0(\tau_{k-1})\} + \mathbf{J}\{\boldsymbol{\beta}_0(\tau_{k-1})\}\{H(\tau_k) - H(\tau_{k-1})\}] n^{1/2} \{\hat{\boldsymbol{\beta}}(\tau_{k-1}) - \boldsymbol{\beta}_0(\tau_{k-1})\} + o_p(1).
\end{aligned}
$$

The above recursive equation gives the following approximation result

$$
\begin{aligned}
& \mathbf{B}\{\boldsymbol{\beta}_0(\tau_k)\} n^{1/2} \{\hat{\boldsymbol{\beta}}(\tau_k) - \boldsymbol{\beta}_0(\tau_k)\} \\
= {} & - \{S_n(\boldsymbol{\beta}_0, \tau_k) - S_n(\boldsymbol{\beta}_0, \tau_{k-1})\} \\
& - [\mathbf{I} + \mathbf{J}\{\boldsymbol{\beta}_0(\tau_{k-1})\} \mathbf{B}^{-1}\{\boldsymbol{\beta}_0(\tau_{k-1})\}\{H(\tau_k) - H(\tau_{k-1})\}] \times \{S_n(\boldsymbol{\beta}_0, \tau_{k-1}) - S_n(\boldsymbol{\beta}_0, \tau_{k-2})\} \\
& - \cdots \\
& - \prod_{h=2}^{k} [\mathbf{I} + \mathbf{J}\{\boldsymbol{\beta}_0(\tau_{h-1})\} \mathbf{B}^{-1}\{\boldsymbol{\beta}_0(\tau_{h-1})\}\{H(\tau_h) - H(\tau_{h-1})\}] \times \{S_n(\boldsymbol{\beta}_0, \tau_1) - S_n(\boldsymbol{\beta}_0, \tau_0)\} \\
& + o_p(1). \qquad \text{(S4)}
\end{aligned}
$$

Using the product integration theory (Gill and Johansen, 1990; Andersen et al., 1993 II.6), we can write the above decomposition as

$$
n^{1/2}[m\{\hat{\boldsymbol{\beta}}(\tau)\} - m\{\boldsymbol{\beta}_0(\tau)\}] = \phi\{-S_n(\boldsymbol{\beta}_0, \tau)\} + o_p(1), \qquad \text{(S5)}
$$

where ϕ is functional defined as follows. For any $g \in \mathcal{G} = \{g : [0, \tau_u] \to \mathbb{R}^p, g$ is left-continuous with right limit, $g(0) = 0\}$ and product integral $\mathcal{I}(s,t) = \boldsymbol{\pi}_{u \in (s,t]}[I_p + \mathbf{J}\{\boldsymbol{\beta}_0(u)\}\mathbf{B}\{\boldsymbol{\beta}_0(u)\}^{-1}]dH(u)$

with $\boldsymbol{\pi}$ the product integral notation, $\phi(g)(\tau)$ is defined as

$$\phi(g)(\tau) = \int_0^\tau \mathcal{I}(s,\tau)dg(s). \tag{S6}$$

Note that $\{\mathbf{Z}_i N_i(e^{\mathbf{Z}_i \boldsymbol{\beta}_0(\tau)}), \tau \in [\tau_l, \tau_u]\}$ is a Donsker class and uniformly bounded. From Condition C5, the class $\{\mathbf{Z}_i v_i(e^{\mathbf{Z}_i^\top \boldsymbol{\beta}_0(s)})Y_i(e^{\mathbf{Z}_i^\top \boldsymbol{\beta}_0(s)}), s \in [\tau_l, \tau_u]\}$ is also Donsker (see Example 2.10.8 in Section 2.10 of van der Vaart and Wellner, 1996). Since $\int_0^\tau v_i(e^{\mathbf{Z}_i^\top \boldsymbol{\beta}_0(s)})Y_i(e^{\mathbf{Z}_i^\top \boldsymbol{\beta}_0(s)})dH(s)$ is Libschitz in τ, from the permanence properties of the Lipschitz transformation of the Donsker class (Theorem 2.10.6, van der Vaart and Wellner, 1996), we can show that

$$\left\{\mathbf{Z}_i N_i(e^{\mathbf{Z}_i \boldsymbol{\beta}_0(\tau)}) - \mathbf{Z}_i \int_0^\tau v_i(e^{\mathbf{Z}_i^\top \boldsymbol{\beta}_0(s)})Y_i(e^{\mathbf{Z}_i^\top \boldsymbol{\beta}_0(s)})dH(s), \tau \in [\tau_l, \tau_u]\right\}$$

is a Donsker class. Then the Donsker theorem implies that $S_n(\boldsymbol{\beta}_0, \tau)$ converges weakly to a Gaussian process $U(\tau)$ living on $\tau \in [\tau_l, \tau_u]$ with mean 0 and covariance matrix $\Sigma(s,t)$, where $\Sigma(s,t) = E\{u_i(s)u_i(t)^\top\}$ with $u_i(\tau) = \mathbf{Z}_i N_i(e^{\mathbf{Z}_i \boldsymbol{\beta}_0(\tau)}) - \mathbf{Z}_i \int_0^\tau v_i(e^{\mathbf{Z}_i^\top \boldsymbol{\beta}_0(s)})Y_i(e^{\mathbf{Z}_i^\top \boldsymbol{\beta}_0(s)})dH(s)$. Furthermore, since ϕ is a linear operator, $\phi\{-S_n(\boldsymbol{\beta}_0, \tau)\}$ converges weakly to $\phi\{-U(\tau)\}$ and $\phi\{-U(\tau)\}$ is also a Gaussian process (Römisch, 2005). From (S5), apply Taylor's expansion and we have

$$\mathbf{B}\{\boldsymbol{\beta}_0(\tau)\}n^{1/2}\{\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\} = \phi\{-S_n(\boldsymbol{\beta}_0, \tau)\} + o_p(1).$$

Then $n^{1/2}\{\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\}$ converges weakly to the Gaussian process $\mathbf{B}\{\boldsymbol{\beta}_0(\tau)\}^{-1}\phi\{-U(\tau)\}$ with covariance matrix $\mathbf{B}\{\boldsymbol{\beta}_0(\tau)\}^{-1}\Sigma^*[\mathbf{B}\{\boldsymbol{\beta}_0(\tau)\}^{-1}]^\top$, where $\Sigma^*(\tau)$ denotes the limiting covariance matrix of $\phi\{-S_n(\boldsymbol{\beta}_0, \tau)\}$.

Proof of Theorem 2 Next we prove the weak convergence of $\sqrt{n}\{\hat{\boldsymbol{\beta}}_{\text{eff}}(\tau) - \boldsymbol{\beta}_0(\tau)\}$. From the above definitions, we can write $\eta(\boldsymbol{\beta}, \tau)$ and $\eta^*(\boldsymbol{\beta}, \tau_k)$ in (18) and (20) as

$$\eta(\boldsymbol{\beta}, \tau) = n^{1/2}m_n^*\{\boldsymbol{\beta}(\tau)\} - \int_0^\tau n^{1/2}\tilde{m}_n^*\{\boldsymbol{\beta}(u)\}dH(u)$$

and

$$\eta^*(\boldsymbol{\beta}, \tau_k) \;\; = \;\; n^{1/2} m_n^*\{\boldsymbol{\beta}(\tau_k)\} - n^{1/2} \sum_{j=0}^{k-1} \tilde{m}_n^*\{\hat{\boldsymbol{\beta}}_{\text{eff}}(\tau_j)\}\{H(\tau_{j+1}) - H(\tau_j)\}.$$

By the proposed estimation method for $\hat{\boldsymbol{\beta}}_{\text{eff}}(\tau)$, the sequential estimator $\hat{\boldsymbol{\beta}}(\tau_k)$, $1 \le k \le L^*$, minimizes

$$n^{-1}\eta^*(\boldsymbol{\beta}, \tau_k)^\top W(\hat{\boldsymbol{\beta}}_{\text{int}}, \tau)^{-1}\eta^*(\boldsymbol{\beta}, \tau_k).$$

Since $\hat{\boldsymbol{\beta}}_{\text{int}}$ is taken as a consistent estimator of $\boldsymbol{\beta}_0$, it can be shown that $W(\hat{\boldsymbol{\beta}}_{\text{int}}, \tau)^{-1} \to W(\boldsymbol{\beta}_0, \tau)^{-1}$ in probability. Further note that

$$\mathbf{B}^*(\mathbf{b}) = \frac{\partial m^*(\mathbf{b})}{\partial \mathbf{b}^\top} \text{ and } \mathbf{J}^*(\mathbf{b}) = \frac{\partial \tilde{m}^*(\mathbf{b})}{\partial \mathbf{b}^\top}.$$

As in the proof of Theorem 1, $m_n^*(\mathbf{b})$ and $\tilde{m}_n^*(\mathbf{b})$ converge to $m^*(\mathbf{b})$ and $\tilde{m}^*(\mathbf{b})$ uniformly in probability; then the sequential estimators $\hat{\boldsymbol{\beta}}_{\text{eff}}(\tau_k)$ satisfy $n^{-1/2}\mathbf{B}^*\{\hat{\boldsymbol{\beta}}_{\text{eff}}(\tau_k)\}^\top W(\boldsymbol{\beta}_0, \tau)^{-1} \times \eta^*(\hat{\boldsymbol{\beta}}_{\text{eff}}, \tau_k) = o_p(1)$. Then from a similar argument as in the proof of Theorem 1 for the consistency of $\hat{\boldsymbol{\beta}}(\cdot)$, we have that $\hat{\boldsymbol{\beta}}_{\text{eff}}(t)$ is uniformly consistent in $0 \le t \le \tau$. Similarly as Lemma B.1 of Peng and Huang (2008), we have for $\tilde{\boldsymbol{\beta}}(t)$ such that $\sup_t \|\tilde{\boldsymbol{\beta}}(t) - \boldsymbol{\beta}_0(t)\| \to 0$ in probability,

$$\sup_{t\in(0,\tau]} \|[m_n^*\{\tilde{\boldsymbol{\beta}}(t)\} - m_n^*\{\boldsymbol{\beta}_0(t)\}] - [m^*\{\tilde{\boldsymbol{\beta}}(t)\} - m^*\{\boldsymbol{\beta}_0(t)\}]\| = o_p(-n^{1/2})$$

and

$$\sup_{t\in(0,\tau]} \|[\tilde{m}_n^*\{\tilde{\boldsymbol{\beta}}(t)\} - \tilde{m}_n^*\{\boldsymbol{\beta}_0(t)\}] - [\tilde{m}^*\{\tilde{\boldsymbol{\beta}}(t)\} - \tilde{m}^*\{\boldsymbol{\beta}_0(t)\}]\| = o_p(n^{-1/2}).$$

Then we can write

$$\begin{aligned}\eta^*(\tilde{\boldsymbol{\beta}}, \tau_k) \;\; &= \;\; n^{1/2}\mathbf{B}^*\{\boldsymbol{\beta}_0(\tau_k)\}\{\tilde{\boldsymbol{\beta}}(\tau_k) - \boldsymbol{\beta}_0(\tau_k)\} + o_p(1) \\ &\quad + n^{1/2} m_n^*\{\boldsymbol{\beta}_0(\tau_k)\} - n^{1/2} \sum_{j=0}^{k-1} \tilde{m}_n^*\{\hat{\boldsymbol{\beta}}_{\text{eff}}(\tau_j)\}\{H(\tau_{j+1}) - H(\tau_j)\}.\end{aligned}$$

This implies the sequential estimator $\hat{\boldsymbol{\beta}}_{\text{eff}}(\tau_k)$ satisfies

$$\mathbf{B}^*\{\boldsymbol{\beta}_0(\tau_k)\}^\top W(\boldsymbol{\beta}_0,\tau)^{-1}\eta^*(\hat{\boldsymbol{\beta}}_{\text{eff}},\tau_k) = o_p(1),$$

uniformly in probability, and furthermore, uniformly in t, $\mathbf{B}^*\{\boldsymbol{\beta}_0(t)\}W(\boldsymbol{\beta}_0,\tau)^{-1}\eta(\hat{\boldsymbol{\beta}}_{\text{eff}},t) = o_p(1)$. This equation plays a similar role in the current proof as that of the unweighted estimating equation $S_n(\hat{\boldsymbol{\beta}},t) = o_p(1)$ in the proof of Theorem 1 for the weak convergence of $\hat{\boldsymbol{\beta}}$. The weak convergence of $\hat{\boldsymbol{\beta}}_{\text{eff}}$ then follows from a similar argument.

S3 Resampling method

An alternative Resampling method A commonly used resampling method is the perturbation-based approach proposed by Jin et al. (2003) for the AFT model; its modified version for the quantile regression was adopted in Peng and Huang (2008) under unbiased sampling. This approach can be easily extended to the biased sampling case. In particular, consider the following stochastic perturbation of the observed estimating equations

$$\tilde{S}_n(\boldsymbol{\beta},\tau) = n^{-1/2}\sum_{i=1}^{n}\zeta_i\mathbf{Z}_i\left\{N_i(e^{\mathbf{Z}_i^\top\boldsymbol{\beta}(\tau)}) - \int_0^\tau v_i(e^{\mathbf{Z}_i^\top\boldsymbol{\beta}(s)})Y_i(e^{\mathbf{Z}_i^\top\boldsymbol{\beta}(s)})dH(s)\right\} = 0. \qquad \text{(S7)}$$

where ζ_i's are i.i.d. from a known nonnegative distribution with mean 1 and variance 1, such as the exponential distribution with rate 1. The above estimation can be implemented in the same fashion as discussed in Section 2.3. Conditional on the observed data, we independently generate the variates $(\zeta_1,\ldots,\zeta_n)$ M_b times, with M_b being a big number, and obtain the corresponding M_b estimates $\hat{\boldsymbol{\beta}}^*(\tau)$ by solving (S7).

The resampling method described above is consistent. Following the proof of Theorem 3 below, conditional on $\{\mathbf{Z}_i,\Delta_i,\tilde{T}_i\}_{i=1}^n$, we have $\tilde{S}_n(\hat{\boldsymbol{\beta}},\tau)$ has the same asymptotic distribution as $S_n(\boldsymbol{\beta}_0, \tau)$ in probability; moreover, for τ uniformly in $(0, \tau_u]$, we have

$$n^{1/2}\{\boldsymbol{\beta}^*(\tau) - \hat{\boldsymbol{\beta}}(\tau)\} = \mathbf{B}(\boldsymbol{\beta}_0(\tau))^{-1}\phi\{-\tilde{S}_n(\hat{\boldsymbol{\beta}}, \tau)\} + o_p(1), \tag{S8}$$

where function ϕ is defined as in (S6). Therefore, conditional on the observed data, $n^{1/2}\{\hat{\boldsymbol{\beta}}^*(\tau) - \hat{\boldsymbol{\beta}}(\tau)\}$ converges weakly to the same limiting process of $n^{1/2}\{\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\}$ for $\tau \in [\tau_\ell, \tau_u]$, where $\tau_\ell \in (0, \tau_u)$.

Proof of Theorem 3 We first show that, conditional on the observed data, $\tilde{S}_n(\hat{\boldsymbol{\beta}}, \tau)$ has the same asymptotic distribution as $S_n(\boldsymbol{\beta}_0, \tau)$ in probability. Consider the weighed summation defined in (S7)

$$\tilde{S}_n(\boldsymbol{\beta}, \tau) = n^{-1/2}\sum_{i=1}^{n}\zeta_i\mathbf{Z}_i\left\{N_i(e^{\mathbf{Z}_i^\top\boldsymbol{\beta}(\tau)}) - \int_0^\tau v_i(e^{\mathbf{Z}_i^\top\boldsymbol{\beta}(s)})Y_i(e^{\mathbf{Z}_i^\top\boldsymbol{\beta}(s)})dH(s)\right\}.$$

Since $E(\zeta_i) = 1$ and $Var(\zeta_i) = 1$, a similar argument as in the proof of Theorem 1 gives that

$$\sup_{\tau\in(0,\tau_u]}\tilde{S}_n(\hat{\boldsymbol{\beta}}^*, \tau) = o_p(1),$$

where $\hat{\boldsymbol{\beta}}^*$ is the solution of $\tilde{S}_n(\boldsymbol{\beta}, \tau) = 0$ using the proposed sequential algorithm. Because $\sup_{\tau\in(0,\tau_u]} S_n(\hat{\boldsymbol{\beta}}, \tau) = o_p(1)$ as shown in the proof of Theorem 1, for τ uniformly in $(0, \tau_u]$, $-\tilde{S}_n(\hat{\boldsymbol{\beta}}, \tau)$ can be expressed as

$$-\tilde{S}_n(\hat{\boldsymbol{\beta}}, \tau) = S_n(\hat{\boldsymbol{\beta}}, \tau) - \tilde{S}_n(\hat{\boldsymbol{\beta}}, \tau) + o_p(1) = n^{-1/2}\sum_{i=1}^{n}(1 - \zeta_i)\hat{u}_i(\tau) + o_p(1),$$

where

$$\hat{u}_i(\tau) = \mathbf{Z}_iN(e^{\mathbf{Z}_i\hat{\boldsymbol{\beta}}(\tau)}) - \mathbf{Z}_i\int_0^\tau v_i(e^{\mathbf{Z}_i^\top\hat{\boldsymbol{\beta}}(s)}, \tilde{T}_i, \Delta_i)Y_i(e^{\mathbf{Z}_i^\top\hat{\boldsymbol{\beta}}(s)})dH(s).$$

Then we have for any $s, t \in [\tau_l, \tau_u]$,

$$
\begin{aligned}
& E\left[\left\{n^{-1/2}\sum_{i=1}^{n}(1-\zeta_i)\hat{u}_i(s)\right\}\left\{n^{-1/2}\sum_{i=1}^{n}(1-\zeta_i)\hat{u}_i(t)\right\}^{\top}\Big|\{\mathbf{Z}_i, \Delta_i, \tilde{T}_i\}_{i=1}^n\right] \\
= \; & n^{-1}\sum_{i=1}^{n}\hat{u}_i(s)\hat{u}_i(t)^{\top} \to \Sigma(s,t),
\end{aligned}
$$

in probability as $n \to \infty$, where $\Sigma(\cdot,\cdot)$ is the corresponding limiting covariance function of $-S_n(\boldsymbol{\beta}_0, \cdot)$ as defined in the proof of Theorem 1. Therefore, by a similar argument as in Lin, Wei, and Ying (1993), conditional on $\{\mathbf{Z}_i, \Delta_i, \tilde{T}_i\}_{i=1}^n$, $\tilde{S}_n(\hat{\boldsymbol{\beta}}, \tau)$ has the same limiting covariance matrix and asymptotic distribution as $S_n(\boldsymbol{\beta}_0, \tau)$ in probability. From the decomposition in (S4), then we can show that conditional on the data, $\phi_n\{-\tilde{S}_n(\hat{\boldsymbol{\beta}}, \tau)\}$ has the same asymptotic distribution as $\phi_n(-S_n(\boldsymbol{\beta}_0, \tau))$ in probability. Furthermore, since $n^{1/2}[m\{\hat{\boldsymbol{\beta}}(\tau)\} - m\{\boldsymbol{\beta}_0(\tau)\}] = \phi_n\{-S_n(\boldsymbol{\beta}_0, \tau)\} + o_p(1)$, we obtain that conditional on the data, $\phi_n\{-\tilde{S}_n(\hat{\boldsymbol{\beta}}, \tau)\}$ converges weakly to the limiting distribution of $n^{1/2}[m\{\hat{\boldsymbol{\beta}}(\tau)\} - m\{\boldsymbol{\beta}_0(\tau)\}]$ in probability.

Following from Zeng and Lin (2008), we use the perturbation estimators $\hat{\mathbf{B}}$ and $\hat{\mathbf{J}}$ to estimate $\mathbf{B}$ and $\mathbf{J}$. The consistency property can be established as follows. Following Lemma B.1 of Peng and Huang (2008), for $\tilde{\boldsymbol{\beta}}(t)$ such that $\sup_t \|\tilde{\boldsymbol{\beta}}(t) - \boldsymbol{\beta}_0(t)\| \to 0$ in probability, the following result holds uniformly in τ

$$
\|n^{1/2}[m_n\{\tilde{\boldsymbol{\beta}}(\tau)\} - m_n\{\boldsymbol{\beta}_0(\tau)\}] - n^{1/2}[m\{\tilde{\boldsymbol{\beta}}(\tau)\} - m\{\boldsymbol{\beta}_0(\tau)\}]\| = o_p(1).
$$

This implies that

$$
n^{1/2}[m_n\{\tilde{\boldsymbol{\beta}}(\tau)\} - m_n\{\boldsymbol{\beta}_0(\tau)\}] = \mathbf{B}\{\boldsymbol{\beta}_0(\tau)\}n^{1/2}\{\tilde{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\} + o_p(1).
$$

Therefore, the above approximation holds for $\tilde{\boldsymbol{\beta}}(\tau) = \hat{\boldsymbol{\beta}}(\tau) + n^{-1/2}\gamma$, where γ follows a p-dimensional multivariate normal distribution with zero mean and identity covariance matrix.

For M generated γ's, let $\boldsymbol{\gamma}$ be the $M \times p$ matrix $(\gamma_1, \cdots, \gamma_M)^\top$ and $\tilde{\mathbf{m}}$ be the m-dimensional vector $[n^{1/2} m_n\{\hat{\boldsymbol{\beta}}(\tau) + n^{-1/2}\gamma_i\} - n^{1/2} m_n\{\hat{\boldsymbol{\beta}}(\tau)\}; i = 1, \cdots, M]^\top$. Then we have

$$\mathbf{B}\{\mathbf{b}_0(\tau)\} = (\boldsymbol{\gamma}^\top \boldsymbol{\gamma})^{-1} \boldsymbol{\gamma}^\top \tilde{\mathbf{m}} + o_p(\mathbf{1}).$$

Note that $(\boldsymbol{\gamma}^\top \boldsymbol{\gamma})^{-1}$ exist with probability 1 for $M > p$. Therefore, the slope matrix estimator given by regressing the perturbed values $n^{1/2} m_n\{\hat{\boldsymbol{\beta}}(\tau) + n^{-1/2}\gamma_i\}$ on γ_i, i.e., $\hat{\mathbf{B}}\{\hat{\boldsymbol{\beta}}(\tau)\} = (\boldsymbol{\gamma}^\top \boldsymbol{\gamma})^{-1} \boldsymbol{\gamma}^\top \tilde{\mathbf{m}}$, is consistent. Similarly we have the consistency of $\hat{\mathbf{J}}$. This implies that $\hat{\mathbf{B}}\{\hat{\boldsymbol{\beta}}(\tau)\}^{-1} \phi_n\{-\tilde{S}_n(\hat{\boldsymbol{\beta}}, \tau)\}$ has the same asymptotic distribution as $n^{1/2}\{\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\}$. This validates the proposed resampling method.

References

Andersen, P. K., Borgan, Ø., Gill, R. D., and Keiding, N. (1993), *Statistical Models Based on Counting Processes*, Springer, New York.

Gill, R. D. and Johansen, S. (1990), "A Survey of Product-Integration with a View Toward Application in Survival Analysis," *The Annals of Statistics*, 18, 1505–1555.

Jin, Z., Lin, D. Y., Wei, L. J., and Ying, Z. (2003), "Rank-Based Inference for the Accelerated Failure Time Model," *Biometrika*, 90, 341–353.

Lin, D. Y., Wei, L.-J., and Ying, Z. (1993), "Checking the Cox Model with Cumulative Sums of Martingale-Based Residuals," *Biometrika*, 80, 557–572.

Peng, L. and Huang, Y. (2008), "Survival Analysis with Quantile Regression Models," *Journal of the American Statistical Association*, 103, 637–649.

Römisch, W. (2005), "Delta Method, Infinite Dimensional," *Encyclopedia of Statistical Sciences*.

van der Vaart, A. and Wellner, J. A. (2000), "Preservation theorems for Glivenko-Cantelli and uniform Glivenko-Cantelli classes," in *High dimensional probability II*, Springer, pp. 115–133.

van der Vaart, A. W. and Wellner, J. A. (1996), *Weak Convergence*, Springer.