ADDITIVE MEAN RESIDUAL LIFE MODEL WITH LATENT VARIABLES UNDER RIGHT CENSORING

Haijin He[1], Deng Pan[2], Xinyuan Song[3] and Liuquan Sun[4]

[1] *Shenzhen University*

[2] *Huazhong University of Science and Technology*

[3] *The Chinese University of Hong Kong*

[4] *Chinese Academy of Sciences*

Supplementary Material

S1 Asymptotic results of the ADF-GLS estimator

To study the asymptotic properties of the ADF-GLS estimator, we need the following regularity condition:

(C1) $\mathbf{\Pi}(\boldsymbol{\theta}_0)$ is positive definite; all partial derivatives of the first three orders of $\mathbf{\Pi}(\boldsymbol{\theta})$ with respect to the elements of $\boldsymbol{\theta}$ are continuous and bounded in a neighborhood of $\boldsymbol{\theta}_0$; $\dot{\mathbf{\Pi}}(\boldsymbol{\theta})$ is of full rank in a neighborhood of $\boldsymbol{\theta}_0$; the eighth-order moments of $\mathbf{V}_i$ are finite.

We first define

$$
\begin{aligned}
\mathbf{K}(\boldsymbol{\theta}) &= \left[\dot{\mathbf{\Pi}}(\boldsymbol{\theta})\mathbf{\Sigma}^*(\boldsymbol{\theta})^{-1}\dot{\mathbf{\Pi}}(\boldsymbol{\theta})^T\right]^{-1}\dot{\mathbf{\Pi}}(\boldsymbol{\theta})\mathbf{\Sigma}^*(\boldsymbol{\theta})^{-1}, \\
\mathbf{R}_i(\boldsymbol{\theta}) &= \left[\dot{\mathbf{\Gamma}}_1(\boldsymbol{\theta})\mathbf{K}(\boldsymbol{\theta})\mathrm{vec}\{\mathbf{V}_i^{\otimes 2}-\mathbf{\Pi}(\boldsymbol{\theta})\},\cdots,\dot{\mathbf{\Gamma}}_p(\boldsymbol{\theta})\mathbf{K}(\boldsymbol{\theta})\mathrm{vec}\{\mathbf{V}_i^{\otimes 2}-\mathbf{\Pi}(\boldsymbol{\theta})\}\right], \\
\mathbf{P}_i(\boldsymbol{\theta}) &= \left[\dot{\mathbf{D}}_1(\boldsymbol{\theta})\mathbf{K}(\boldsymbol{\theta})\mathrm{vec}\{\mathbf{V}_i^{\otimes 2}-\mathbf{\Pi}(\boldsymbol{\theta})\},\cdots,\dot{\mathbf{D}}_q(\boldsymbol{\theta})\mathbf{K}(\boldsymbol{\theta})\mathrm{vec}\{\mathbf{V}_i^{\otimes 2}-\mathbf{\Pi}(\boldsymbol{\theta})\}\right],
\end{aligned}
$$

where $\dot{\mathbf{\Gamma}}_s(\boldsymbol{\theta})$ $(s=1,\cdots,p)$ and $\dot{\mathbf{D}}_r(\boldsymbol{\theta})$ $(r=1,\cdots,q)$ denote the derivatives of the sth column of $\mathbf{\Gamma}(\boldsymbol{\theta})$ and the rth column of $\mathbf{D}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}^T$, respectively.

$$
\begin{aligned}
\mathbf{K}^\dagger(\widehat{\boldsymbol{\theta}}) &= \left[\dot{\mathbf{\Pi}}(\widehat{\boldsymbol{\theta}})\mathbf{W}^{-1}\dot{\mathbf{\Pi}}(\widehat{\boldsymbol{\theta}})^T\right]^{-1}\dot{\mathbf{\Pi}}(\widehat{\boldsymbol{\theta}})\mathbf{W}^{-1}, \\
\mathbf{R}_i^\dagger(\widehat{\boldsymbol{\theta}}) &= \left[\dot{\mathbf{\Gamma}}_1(\widehat{\boldsymbol{\theta}})\mathbf{K}^\dagger(\widehat{\boldsymbol{\theta}})\mathrm{vec}\{\mathbf{V}_i^{\otimes 2}-\mathbf{\Pi}(\widehat{\boldsymbol{\theta}})\},\cdots,\dot{\mathbf{\Gamma}}_p(\widehat{\boldsymbol{\theta}})\mathbf{K}^\dagger(\widehat{\boldsymbol{\theta}})\mathrm{vec}\{\mathbf{V}_i^{\otimes 2}-\mathbf{\Pi}(\widehat{\boldsymbol{\theta}})\}\right], \\
\mathbf{P}_i^\dagger(\widehat{\boldsymbol{\theta}}) &= \left[\dot{\mathbf{D}}_1(\widehat{\boldsymbol{\theta}})\mathbf{K}^\dagger(\widehat{\boldsymbol{\theta}})\mathrm{vec}\{\mathbf{V}_i^{\otimes 2}-\mathbf{\Pi}(\widehat{\boldsymbol{\theta}})\},\cdots,\dot{\mathbf{D}}_q(\widehat{\boldsymbol{\theta}})\mathbf{K}^\dagger(\widehat{\boldsymbol{\theta}})\mathrm{vec}\{\mathbf{V}_i^{\otimes 2}-\mathbf{\Pi}(\widehat{\boldsymbol{\theta}})\}\right],
\end{aligned}
$$

are estimates of $\mathbf{K}(\boldsymbol{\theta}_0)$, $\mathbf{R}_i(\boldsymbol{\theta}_0)$ and $\mathbf{P}_i(\boldsymbol{\theta}_0)$, respectively.

(S1.1) in the following Lemma is basically an adoption of Browne (1984), while we adjust the statement for our purpose. (S1.2) and (S1.3) can be regarded as an application of the delta-method.

Lemma 1. *Under the condition (C1), $\mathbf{\Gamma}(\widehat{\boldsymbol{\theta}})$ and $\mathbf{D}(\widehat{\boldsymbol{\theta}})$ are consistent to $\mathbf{\Gamma}(\boldsymbol{\theta}_0)$ and $\mathbf{D}(\boldsymbol{\theta}_0)$, respectively. Furthermore,*

$$\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0=\mathbf{K}(\boldsymbol{\theta}_0)\mathrm{vec}\{\mathbf{S}^*-\mathbf{\Pi}(\boldsymbol{\theta}_0)\}+o_p(n^{-1/2}), \tag{S1.1}$$

$$\mathbf{\Gamma}(\widehat{\boldsymbol{\theta}})-\mathbf{\Gamma}(\boldsymbol{\theta}_0)=n^{-1}\sum_{i=1}^n\mathbf{R}_i(\boldsymbol{\theta}_0)+o_p(n^{-1/2}), \tag{S1.2}$$

$$\mathbf{D}(\widehat{\boldsymbol{\theta}}) - \mathbf{D}(\boldsymbol{\theta}_0) = n^{-1}\sum_{i=1}^{n} \mathbf{P}_i(\boldsymbol{\theta}_0) + o_p(n^{-1/2}), \tag{S1.3}$$

where $\mathbf{S}^* = n^{-1}\sum_{i=1}^{n} \mathbf{V}_i^{\otimes 2}$.

Proof. To prove the consistency of $\widehat{\boldsymbol{\theta}}$, we define

$$F^*(\boldsymbol{\theta}) = \frac{1}{2}\{\text{vec}(\boldsymbol{\Pi}(\boldsymbol{\theta}_0) - \boldsymbol{\Pi}(\boldsymbol{\theta}))\}^T \boldsymbol{\Sigma}^*(\boldsymbol{\theta}_0)^{-1}\{\text{vec}(\boldsymbol{\Pi}(\boldsymbol{\theta}_0) - \boldsymbol{\Pi}(\boldsymbol{\theta}))\}.$$

Given that model (2.1) of the main text is identified and $\boldsymbol{\Sigma}^*(\boldsymbol{\theta}_0)$ is positive definite, $F^*(\boldsymbol{\theta})$ has its unique minimum 0 at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Since $\mathbf{S}$ and $\mathbf{W}$ converge in probability to $\boldsymbol{\Pi}(\boldsymbol{\theta}_0)$ and $\boldsymbol{\Sigma}^*(\boldsymbol{\theta}_0)$, respectively, and $\boldsymbol{\Pi}(\boldsymbol{\theta})$ is bounded in a neighborhood of $\boldsymbol{\theta}_0$, $F(\boldsymbol{\theta})$ converges in probability to $F^*(\boldsymbol{\theta})$ in a neighborhood of $\boldsymbol{\theta}_0$, then the unique minimizer $\widehat{\boldsymbol{\theta}}$ of $F(\boldsymbol{\theta})$ converges in probability to the unique minimizer $\boldsymbol{\theta}_0$ of $F^*(\boldsymbol{\theta})$, this consistency proof is an adoption of Browne (1974, 1984).

By the uniform convergence theorem, $\ddot{F}(\boldsymbol{\theta})$ converges to a deterministic function uniformly in a neighborhood of $\boldsymbol{\theta}_0$. In particular, $\ddot{F}(\boldsymbol{\theta}_0)$ converges in probability to $\dot{\boldsymbol{\Pi}}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}^*(\boldsymbol{\theta}_0)^{-1}\dot{\boldsymbol{\Pi}}(\boldsymbol{\theta}_0)^T$, which is nonsingular by regularity condition (C1). Since $\dot{F}(\widehat{\boldsymbol{\theta}}) = 0$, apply the Taylor expansion to each element of $\dot{F}(\widehat{\boldsymbol{\theta}})$, and use the consistency of $\widehat{\boldsymbol{\theta}}$, we have

$$-\dot{F}(\boldsymbol{\theta}_0) = \ddot{F}(\boldsymbol{\theta}_0)(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + o_p(n^{-1/2}).$$

Note that by the multivariate central limit theorem, $\dot{F}(\boldsymbol{\theta}_0) = O_p(n^{-1/2})$. Using the convergence of $\ddot{F}(\boldsymbol{\theta}_0)$, we obtain

$$\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 = \mathbf{K}(\boldsymbol{\theta}_0)\mathrm{vec}\{\mathbf{S} - \boldsymbol{\Pi}(\boldsymbol{\theta}_0)\} + o_p(n^{-1/2}),$$

since we can show

$$\mathbf{S} - \mathbf{S}^* = (n-1)^{-1}\mathbf{S}^* - (n-1)^{-1}n\bar{\mathbf{V}}\bar{\mathbf{V}}^T = o_p(n^{-1/2}),$$

(S1.1) follows. Similarly, by the Taylor expansion,

$$\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}}) - \boldsymbol{\Gamma}(\boldsymbol{\theta}_0) = \left[\dot{\boldsymbol{\Gamma}}_1(\boldsymbol{\theta}_0)(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0), \cdots, \dot{\boldsymbol{\Gamma}}_p(\boldsymbol{\theta}_0)(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)\right] + o_p(n^{-1/2}),$$

which together with (S1.1) gives (S1.2). In a similar manner, we obtain (S1.3).

S2 Proofs of Asymptotic Results of CEE1 estimators

Expressions in $\widehat{\boldsymbol{\alpha}}_a$.

$$\begin{aligned}\frac{\partial \mathbf{U}_a(\boldsymbol{\alpha};\boldsymbol{\theta})}{\partial \boldsymbol{\alpha}^T} &= \sum_{i=1}^n \int_0^\tau \begin{pmatrix} \mathbf{Z}_i - \bar{\mathbf{Z}}_a(t) \\ \widehat{\boldsymbol{\xi}}_i(\boldsymbol{\theta}) - \bar{\boldsymbol{\xi}}_a^*(t;\boldsymbol{\theta}) \end{pmatrix} \left[\frac{\partial \widehat{m}_{a0}(t;\boldsymbol{\alpha},\boldsymbol{\theta})}{\partial \boldsymbol{\alpha}^T} + \begin{pmatrix} \mathbf{Z}_i \\ \widehat{\boldsymbol{\xi}}_i(\boldsymbol{\theta}) \end{pmatrix}^T\right] dN_i(t) \\ &\quad - \begin{bmatrix} \mathbf{0}_{s\times s} & \mathbf{0}_{s\times q} \\ \mathbf{0}_{q\times s} & \mathbf{D}(\boldsymbol{\theta}) \end{bmatrix} \sum_{i=1}^n \int_0^\tau dN_i(t), \end{aligned} \tag{S2.1}$$

and

$$\mathbf{U}_a(0;\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^n \int_0^\tau \begin{pmatrix} \mathbf{Z}_i - \bar{\mathbf{Z}}_a(t) \\ \widehat{\boldsymbol{\xi}}_i(\widehat{\boldsymbol{\theta}}) - \bar{\boldsymbol{\xi}}_a^*(t;\widehat{\boldsymbol{\theta}}) \end{pmatrix} \left(\widehat{S}_{NA}(t)^{-1}\int_t^\tau \widehat{S}_{NA}(u)du\right) dN_i(t).$$

Notations in Theorem 1.

$$
\begin{aligned}
\bar{\mathbf{V}}_a(t) &= \sum_{i=1}^n Y_i(t)\mathbf{V}_i / \sum_{i=1}^n Y_i(t), \\
\mathbf{W}_z(t) &= \frac{\widehat{S}_{NA}(t)}{\sum_{j=1}^n Y_j(t)} \int_0^t \widehat{S}_{NA}(u)^{-1} \sum_{j=1}^n \{\mathbf{Z}_j - \bar{\mathbf{Z}}_a(u)\} dN_j(u), \\
\mathbf{W}_\xi(t;\boldsymbol{\theta}) &= \frac{\widehat{S}_{NA}(t)}{\sum_{j=1}^n Y_j(t)} \int_0^t \widehat{S}_{NA}(u)^{-1} \sum_{j=1}^n \{\boldsymbol{\Gamma}(\boldsymbol{\theta})\mathbf{V}_j - \bar{\boldsymbol{\xi}}_a^*(u;\boldsymbol{\theta})\} dN_j(u), \\
d\widehat{\Omega}_i(t) &= \{\widehat{m}_{a0}(t) + \widehat{\boldsymbol{\beta}}_a^T \mathbf{Z}_i + \widehat{\boldsymbol{\gamma}}_a^T \boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i\} dN_i(t) - Y_i(t) d\{\widehat{m}_{a0}(t) + t\}, \\
\widehat{\mathbf{Q}}_{a1} &= n^{-1} \sum_{i=1}^n \int_0^\tau \{\mathbf{Z}_i - \bar{\mathbf{Z}}_a(t) - \mathbf{W}_z(t)\} \mathbf{V}_i^T dN_i(t), \\
\widehat{\mathbf{Q}}_{a2} &= n^{-1} \sum_{i=1}^n \int_0^\tau \{\mathbf{V}_i - \bar{\mathbf{V}}_a(t)\}\{\widehat{m}_{a0}(t) + \widehat{\boldsymbol{\beta}}_a^T \mathbf{Z}_i + \widehat{\boldsymbol{\gamma}}_a^T \boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i\} dN_i(t), \\
\widehat{\mathbf{Q}}_{a3} &= n^{-1} \sum_{i=1}^n \int_0^\tau \{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i - \bar{\boldsymbol{\xi}}_a^*(t;\widehat{\boldsymbol{\theta}}) - \mathbf{W}_\xi(t;\widehat{\boldsymbol{\theta}})\} \mathbf{V}_i^T dN_i(t), \\
\widehat{Q}_{a4} &= n^{-1} \sum_{i=1}^n \int_0^\tau dN_i(t), \quad \widehat{\boldsymbol{\Sigma}}_a = n^{-1} \sum_{i=1}^n \widehat{\mathbf{U}}_{ai}^{\otimes 2}, \\
\widehat{\mathbf{U}}_{ai1} &= \int_0^\tau \{\mathbf{Z}_i - \bar{\mathbf{Z}}_a(t) - \mathbf{W}_z(t)\} d\widehat{\Omega}_i(t) + \widehat{\mathbf{Q}}_{a1} \mathbf{R}_i^\dagger(\widehat{\boldsymbol{\theta}})^T \widehat{\boldsymbol{\gamma}}_a, \\
\widehat{\mathbf{U}}_{ai2} &= \int_0^\tau \{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i - \bar{\boldsymbol{\xi}}_a^*(t;\widehat{\boldsymbol{\theta}}) - \mathbf{W}_\xi(t;\widehat{\boldsymbol{\theta}})\} d\widehat{\Omega}_i(t) - \int_0^\tau \mathbf{D}(\widehat{\boldsymbol{\theta}})\widehat{\boldsymbol{\gamma}}_a dN_i(t) \\
&\quad + \mathbf{R}_i^\dagger(\widehat{\boldsymbol{\theta}})\widehat{\mathbf{Q}}_{a2} + \widehat{\mathbf{Q}}_{a3} \mathbf{R}_i^\dagger(\widehat{\boldsymbol{\theta}})^T \widehat{\boldsymbol{\gamma}}_a - \mathbf{P}_i^\dagger(\widehat{\boldsymbol{\theta}})\widehat{\boldsymbol{\gamma}}_a \widehat{Q}_{a4}, \\
\widehat{\mathbf{U}}_{ai} &= (\widehat{\mathbf{U}}_{ai1}^T, \widehat{\mathbf{U}}_{ai2}^T)^T, \quad \widehat{\mathbf{A}}_a = \frac{1}{n} \frac{\partial \mathbf{U}_a(\boldsymbol{\alpha};\widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\alpha}^T},
\end{aligned}
$$

in which $\mathbf{R}_i^\dagger(\widehat{\boldsymbol{\theta}})$ and $\mathbf{P}_i^\dagger(\widehat{\boldsymbol{\theta}})$ are defined in Supplementary Material S1. The explicit expression of $\widehat{\mathbf{A}}_a$ can be obtained from (3.8) in Section 3.1 of the main text and (S2.1).

Conditions for CEE1 estimators.

To study the asymptotic properties of the proposed CEE1 estimators, in addition to condition (C1), we need the following regularity conditions:

(C2) The true value $(\boldsymbol{\theta}_0^T, \boldsymbol{\alpha}_0^T)^T$ of $(\boldsymbol{\theta}^T, \boldsymbol{\alpha}^T)^T$ lies in the interior of a compact set $\boldsymbol{\Theta}$.

(C3) $\mathbf{Z}_i$ is bounded almost surely; $m_0(t)$ is continuously differentiable on $[0, \tau]$.

(C4) The limiting matrix of $\widehat{\mathbf{A}}_a$, denoted by $\mathbf{A}_a$, is nonsingular.

Proof of Theorem 1.

Denote $d\Omega_i(t) = d\Omega_i(t; \boldsymbol{\alpha}_0, \boldsymbol{\theta}_0)$ for simplicity. We have the following decomposition

$$\begin{aligned} & \mathbf{U}_{a1}(\boldsymbol{\alpha}_0; \boldsymbol{\theta}_0) \\ = & \sum_{i=1}^n \int_0^\tau \{\mathbf{Z}_i - \bar{\mathbf{Z}}_a(t)\}\{\widehat{m}_{a0}(t; \boldsymbol{\alpha}_0, \boldsymbol{\theta}_0) - m_0(t)\} dN_i(t) \\ & + \sum_{i=1}^n \int_0^\tau \{\mathbf{Z}_i - \bar{\mathbf{Z}}_a(t)\}\{m_0(t) + \boldsymbol{\beta}_0^T \mathbf{Z}_i + \boldsymbol{\gamma}_0^T \widehat{\boldsymbol{\xi}}_i(\boldsymbol{\theta}_0)\} dN_i(t). \quad \text{(S2.2)} \end{aligned}$$

Using (3.7) of the main text, and interchanging the order of integration with respect to t and u, the first term of (S2.2) becomes

$$-\sum_{i=1}^{n}\int_{0}^{\tau}\mathbf{W}_{z}(t)d\Omega_{i}(t). \tag{S2.3}$$

Some algebraic manipulations yield that the second term of (S2.2) equals

$$\sum_{i=1}^{n}\int_{0}^{\tau}\{\mathbf{Z}_{i}-\bar{\mathbf{Z}}_{a}(t)\}d\Omega_{i}(t). \tag{S2.4}$$

By (S2.2), (S2.3) and (S2.4), and Lemma 1 of Lin et al. (2000), we have

$$\begin{aligned}\mathbf{U}_{a1}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0) &= \sum_{i=1}^{n}\int_{0}^{\tau}\{\mathbf{Z}_{i}-\bar{\mathbf{Z}}_{a}(t)-\mathbf{W}_{z}(t)\}d\Omega_{i}(t)\\ &= \sum_{i=1}^{n}\int_{0}^{\tau}\{\mathbf{Z}_{i}-\mathbf{e}_{az}(t)-\mathbf{w}_{z}(t)\}d\Omega_{i}(t)+o_p(n^{1/2}),\end{aligned} \tag{S2.5}$$

where $\mathbf{e}_{az}(t)$ and $\mathbf{w}_{z}(t)$ are the limits of $\bar{\mathbf{Z}}_{a}(t)$ and $\mathbf{W}_{z}(t)$, respectively. A direct calculation shows

$$\begin{aligned}&\mathbf{U}_{a1}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}})-\mathbf{U}_{a1}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0)\\ =\ &\sum_{i=1}^{n}\int_{0}^{\tau}\{\mathbf{Z}_{i}-\bar{\mathbf{Z}}_{a}(t)\}\{\widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\widehat{\boldsymbol{\theta}})-\widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\boldsymbol{\theta}_0)\}dN_i(t)\\ &+\sum_{i=1}^{n}\int_{0}^{\tau}\{\mathbf{Z}_{i}-\bar{\mathbf{Z}}_{a}(t)\}\mathbf{V}_i^T dN_i(t)\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})-\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\}^T\boldsymbol{\gamma}_0.\end{aligned} \tag{S2.6}$$

Similar to the derivation of (S2.3), the first term in (S2.6) equals

$$-\sum_{i=1}^{n}\int_{0}^{\tau}\mathbf{W}_{z}(t)\mathbf{V}_i^T dN_i(t)\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})-\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\}^T\boldsymbol{\gamma}_0.$$

By Lemma 1 in Supplementary Material S1, (S2.6) equals

$$n\widehat{\mathbf{Q}}_{a1}\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})-\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\}^T\boldsymbol{\gamma}_0=\sum_{i=1}^{n}\mathbf{Q}_{a1}\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0+o_p(n^{1/2}), \tag{S2.7}$$

where $\mathbf{Q}_{a1}$ is the limit of $\widehat{\mathbf{Q}}_{a1}$. Then by (S2.5), (S2.6) and (S2.7), we have

$$\begin{aligned}\mathbf{U}_{a1}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}) &= \mathbf{U}_{a1}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0)+\{\mathbf{U}_{a1}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}})-\mathbf{U}_{a1}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0)\}\\ &= \sum_{i=1}^n \mathbf{U}_{ai1}+o_p(n^{1/2}),\end{aligned} \tag{S2.8}$$

where

$$\mathbf{U}_{ai1}=\int_0^\tau\{\mathbf{Z}_i-\mathbf{e}_{az}(t)-\mathbf{w}_z(t)\}d\Omega_i(t)+\mathbf{Q}_{a1}\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0.$$

Similar to the derivation of (S2.5), we have

$$\begin{aligned}\mathbf{U}_{a2}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0) &= \sum_{i=1}^n\int_0^\tau\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i-\bar{\boldsymbol{\xi}}_a^*(t;\boldsymbol{\theta}_0)-\mathbf{W}_\xi(t;\boldsymbol{\theta}_0)\}d\Omega_i(t)\\ &\quad -\sum_{i=1}^n\int_0^\tau\mathbf{D}(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0 dN_i(t)\\ &= \sum_{i=1}^n\int_0^\tau\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i-\mathbf{e}_{a\xi}(t)-\mathbf{w}_\xi(t)\}d\Omega_i(t)\\ &\quad -\sum_{i=1}^n\int_0^\tau\mathbf{D}(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0 dN_i(t)+o_p(n^{1/2}),\end{aligned} \tag{S2.9}$$

where $\mathbf{e}_{a\xi}(t)$ and $\mathbf{w}_\xi(t)$ are the limits of $\bar{\boldsymbol{\xi}}_a^*(t;\boldsymbol{\theta}_0)$ and $\mathbf{W}_\xi(t;\boldsymbol{\theta}_0)$, respectively.

Some algebra yields

$$\begin{aligned}&\mathbf{U}_{a2}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}})-\mathbf{U}_{a2}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0)\\ =\ & n\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})-\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\}\widehat{\mathbf{Q}}_{a2}^*+n\widehat{\mathbf{Q}}_{a3}^*\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})-\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\}^T\boldsymbol{\gamma}_0\\ & -n\{\mathbf{D}(\widehat{\boldsymbol{\theta}})-\mathbf{D}(\boldsymbol{\theta}_0)\}\boldsymbol{\gamma}_0\widehat{Q}_{a4},\end{aligned} \tag{S2.10}$$

where

$$\widehat{\mathbf{Q}}_{a2}^* = n^{-1}\sum_{i=1}^n\int_0^\tau\{\mathbf{V}_i-\bar{\mathbf{V}}_a(t)\}\{\widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\widehat{\boldsymbol{\theta}})+\boldsymbol{\beta}_0^T\mathbf{Z}_i+\boldsymbol{\gamma}_0^T\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i\}dN_i(t),$$

$$\widehat{\mathbf{Q}}^*_{a3} = n^{-1}\sum_{i=1}^{n}\int_0^\tau\{\mathbf{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i - \bar{\boldsymbol{\xi}}^*_a(t;\boldsymbol{\theta}_0) - \mathbf{W}_\xi(t;\boldsymbol{\theta}_0)\}\mathbf{V}_i^T dN_i(t).$$

Let $\mathbf{Q}_{a2}$, $\mathbf{Q}_{a3}$ and Q_{a4} be the limits of $\widehat{\mathbf{Q}}^*_{a2}$, $\widehat{\mathbf{Q}}^*_{a3}$ and $\widehat{Q}_{a4}$, respectively, then (S2.10) becomes

$$\sum_{i=1}^{n}\mathbf{R}_i(\boldsymbol{\theta}_0)\mathbf{Q}_{a2} + \sum_{i=1}^{n}\mathbf{Q}_{a3}\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0 - \sum_{i=1}^{n}\mathbf{P}_i(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0 Q_{a4} + o_p(n^{1/2}) \tag{S2.11}$$

By (S2.9) and (S2.11), we have

$$\begin{aligned}\mathbf{U}_{a2}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}) &= \mathbf{U}_{a2}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0) + \{\mathbf{U}_{a2}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}) - \mathbf{U}_{a2}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0)\} \\ &= \sum_{i=1}^{n}\mathbf{U}_{ai2} + o_p(n^{1/2}), \qquad \text{(S2.12)}\end{aligned}$$

where

$$\begin{aligned}\mathbf{U}_{ai2} &= \int_0^\tau\{\mathbf{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i - \mathbf{e}_{a\xi}(t) - \mathbf{w}_\xi(t)\}d\Omega_i(t) - \int_0^\tau \mathbf{D}(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0 dN_i(t) \\ &\quad + \mathbf{R}_i(\boldsymbol{\theta}_0)\mathbf{Q}_{a2} + \mathbf{Q}_{a3}\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0 - \mathbf{P}_i(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0 Q_{a4}.\end{aligned}$$

Let $\mathbf{U}_{ai} = (\mathbf{U}_{ai1}^T, \mathbf{U}_{ai2}^T)^T$. Then, it follows from (S2.8) and (S2.12) that

$$\mathbf{U}_a(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^{n}\mathbf{U}_{ai} + o_p(n^{1/2}), \tag{S2.13}$$

which is a sum of n iid zero-mean random vectors plus an asymptotically negligible term. The law of large numbers shows that $n^{-1}\mathbf{U}_a(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}) \to 0$ in probability, and the multivariate central limit theorem shows that $n^{-1/2}\mathbf{U}_a(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}})$ converges in distribution to a normal random vector with mean zero and covariance matrix $\mathbf{\Sigma}_a = E\{\mathbf{U}_{ai}^{\otimes 2}\}$. Note that

$$\widehat{\boldsymbol{\alpha}}_a - \boldsymbol{\alpha}_0 = -n^{-1}\widehat{\mathbf{A}}_a^{-1}\ \mathbf{U}_a(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}), \tag{S2.14}$$

and $\widehat{\mathbf{A}}_a \to \mathbf{A}_a$ in probability by the consistency of $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})$ and $\mathbf{D}(\widehat{\boldsymbol{\theta}})$. Then, based on (S2.14), $\widehat{\boldsymbol{\alpha}}_a$ converges in probability to $\boldsymbol{\alpha}_0$, and $n^{1/2}(\widehat{\boldsymbol{\alpha}}_a - \boldsymbol{\alpha}_0)$ is asymptotically normal with mean zero and covariance matrix $\mathbf{A}_a^{-1}\boldsymbol{\Sigma}_a\mathbf{A}_a^{-1T}$, which can be consistently estimated by $\widehat{\mathbf{A}}_a^{-1}\widehat{\boldsymbol{\Sigma}}_a\widehat{\mathbf{A}}_a^{-1T}$.

Notations in Theorem 2.

$$\begin{aligned}\widehat{O}_{ai}(t) &= \widehat{\mathbf{L}}_{a1}(t)^T\widehat{\mathbf{A}}_a^{-1}\widehat{\mathbf{U}}_{ai} - \widehat{\mathbf{L}}_{a2}(t)^T\mathbf{R}_i^{\dagger}(\widehat{\boldsymbol{\theta}})^T\widehat{\boldsymbol{\gamma}}_a \\ &\quad -\widehat{S}_{NA}(t)^{-1}\int_t^\tau \frac{\widehat{S}_{NA}(u)}{n^{-1}\sum_{j=1}^n Y_j(u)}d\widehat{\Omega}_i(u),\end{aligned}$$

in which

$$\widehat{\mathbf{L}}_{a1}(t) = \widehat{S}_{NA}(t)^{-1}\int_t^\tau \widehat{S}_{NA}(u)\frac{\sum_{i=1}^n\left(\mathbf{Z}_i^T, \widehat{\boldsymbol{\xi}}_i(\widehat{\boldsymbol{\theta}})^T\right)^T dN_i(u)}{\sum_{i=1}^n Y_i(u)},$$

and

$$\widehat{\mathbf{L}}_{a2}(t) = \widehat{S}_{NA}(t)^{-1}\int_t^\tau \widehat{S}_{NA}(u)\frac{\sum_{i=1}^n \mathbf{V}_i dN_i(u)}{\sum_{i=1}^n Y_i(u)}.$$

Proof of Theorem 2.

By (3.7) in Section 3.1 of the main text, and the uniform law of large numbers, $\widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\boldsymbol{\theta}_0) - m_0(t)$ converges in probability to 0 uniformly in $t \in [0,\tau]$. Along with the consistency of $\widehat{\boldsymbol{\alpha}}_a$ and $\widehat{\boldsymbol{\theta}}$, we have that $\widehat{m}_{a0}(t)$ converges in probability to $m_0(t)$ uniformly in $t \in [0,\tau]$. Using (S2.13) and (S2.14), we have

$$\widehat{m}_{a0}(t;\widehat{\boldsymbol{\alpha}}_a,\widehat{\boldsymbol{\theta}}) - \widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\widehat{\boldsymbol{\theta}})$$

$$\begin{aligned} &= -\widehat{\mathbf{L}}_{a1}(t)^T(\widehat{\boldsymbol{\alpha}}_a - \boldsymbol{\alpha}_0) \\ &= n^{-1}\sum_{i=1}^{n} \mathbf{l}_{a1}(t)^T \mathbf{A}_a^{-1}\mathbf{U}_{ai} + o_p(n^{-1/2}), \end{aligned} \tag{S2.15}$$

and using (S1.2) of Lemma 1 of Supplementary Material S1,

$$\begin{aligned} &\widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\widehat{\boldsymbol{\theta}}) - \widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\boldsymbol{\theta}_0) \\ &= -\widehat{\mathbf{L}}_{a2}(t)^T\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}}) - \boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\}^T\boldsymbol{\gamma}_0 \\ &= -n^{-1}\sum_{i=1}^{n} \mathbf{l}_{a2}(t)^T\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0 + o_p(n^{-1/2}), \end{aligned} \tag{S2.16}$$

where $\mathbf{l}_{a1}(t)$ and $\mathbf{l}_{a2}(t)$ are the limits of $\widehat{\mathbf{L}}_{a1}(t)$ and $\widehat{\mathbf{L}}_{a2}(t)$, respectively. We also have

$$\begin{aligned} &\widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\boldsymbol{\theta}_0) - m_0(t) \\ &= -n^{-1}\sum_{i=1}^{n} S_{NA}(t)^{-1}\int_t^{\tau}\frac{S_{NA}(u)}{E[Y_1(u)]}d\Omega_i(u) + o_p(n^{-1/2}), \end{aligned} \tag{S2.17}$$

where $S_{NA}(t)$ is the limit of $\widehat{S}_{NA}(t)$. Then by (S2.15)–(S2.17), we have

$$\begin{aligned} &n^{1/2}\{\widehat{m}_{a0}(t) - m_0(t)\} \\ &= n^{1/2}\{\widehat{m}_{a0}(t;\widehat{\boldsymbol{\alpha}},\widehat{\boldsymbol{\theta}}) - \widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\widehat{\boldsymbol{\theta}})\} \\ &\quad + n^{1/2}\{\widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\widehat{\boldsymbol{\theta}}) - \widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\boldsymbol{\theta}_0)\} \\ &\quad + n^{1/2}\{\widehat{m}_{a0}(t;\boldsymbol{\alpha}_0,\boldsymbol{\theta}_0) - m_0(t)\} \\ &= n^{-1/2}\sum_{i=1}^{n} O_{ai}(t) + o_p(1), \end{aligned} \tag{S2.18}$$

where

$$O_{ai}(t) = \mathbf{l}_{a1}(t)^T\mathbf{A}_a^{-1}\mathbf{U}_{ai} - \mathbf{l}_{a2}(t)^T\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0 - S_{NA}(t)^{-1}\int_t^{\tau}\frac{S_{NA}(u)}{E[Y_1(u)]}d\Omega_i(u).$$

By the multivariate central limit theorem, $n^{1/2}\{\widehat{m}_{a0}(t)-m_0(t)\}$ converges in finite-distribution to a zero-mean Gaussian process. Since each $O_{ai}(t)$ can be written as the sum or product of monotone functions of t and is thus managable, $O_{ai}(t)$ is tight, then $n^{1/2}\{\widehat{m}_{a0}(t)-m_0(t)\}$ is tight and converges weakly to a zero-mean Gaussian process with covariance function at (t,s) given by $\Upsilon_a(t,s) = E\{O_{ai}(t)O_{ai}(s)\}$, which can be consistently estimated by $\widehat{\Upsilon}_a(t,s)$.

S3 Corrected estimating equations under covariate-independent censoring (CEE2)

S3.1 The CEE2 Method

In this Supplementary Material S3, we develop another corrected estimating equation method under the assumption that the censoring time C_i is independent of T_i, $\mathbf{Z}_i$, and $\boldsymbol{\xi}_i$. Notably, the CEE1 method is also applicable to the case of independent censoring time. However, the simulation results show that CEE2 does attain a higher efficiency in certain situations. Thus, we present the CEE2 method for comparison.

Let $G(t)$ be the survivor function of C_i, and $\widehat{G}(t)$ be the Kaplan-Meier estimator of $G(t)$ based on data $\{X_i, 1-\Delta_i, i=1,\cdots,n\}$. Define

$$M_i^*(t) = \frac{\Delta_i I(X_i > t)}{G(X_i)}[(X_i - t) - m_0(t) - \boldsymbol{\beta}^T \mathbf{Z}_i - \boldsymbol{\gamma}^T \boldsymbol{\xi}_i], \quad i = 1, \cdots, n.$$

Under model (2.2) of the main text, $M_i^*(t)$ are mean-zero processes. If $\boldsymbol{\xi}_i$ are observable, for given $\boldsymbol{\alpha}$, $m_0(t)$ can be estimated by $\widehat{m}_{b0}(t;\boldsymbol{\alpha})$ satisfying the following estimating equation (Sun and Zhang, 2009):

$$\sum_{i=1}^n \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)}[(X_i - t) - \widehat{m}_{b0}(t;\boldsymbol{\alpha}) - \boldsymbol{\beta}^T \mathbf{Z}_i - \boldsymbol{\gamma}^T \boldsymbol{\xi}_i] = 0, \ 0 \le t \le \tau, \quad \text{(S3.1)}$$

The estimator is then given by

$$\widehat{m}_{b0}(t;\boldsymbol{\alpha}) = \frac{\sum_{i=1}^n \Delta_i I(X_i > t)\widehat{G}^{-1}(X_i)[(X_i - t) - \boldsymbol{\beta}^T \mathbf{Z}_i - \boldsymbol{\gamma}^T \boldsymbol{\xi}_i]}{\sum_{i=1}^n \Delta_i I(X_i > t)\widehat{G}^{-1}(X_i)}.$$

Without censoring, (S3.1) is similar to the efficient nonparametric estimation of the MRL function as discussed in Bickel et al. (1993).

To estimate $\boldsymbol{\alpha}$, inspired by the generalized estimating equation method (Liang and Zeger, 1986) and the inverse probability censoring weighting (IPCW) technique (Robins and Rotnitzky, 1992; Robins, et al., 1994; van der Laan and Robins, 2003, Sun and Zhang, 2009), we can obtain the estimating function

$$\mathbf{U}_b^*(\boldsymbol{\alpha}) = \sum_{i=1}^n \int_0^\tau \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)} \begin{pmatrix} \mathbf{Z}_i \\ \boldsymbol{\xi}_i \end{pmatrix} [(X_i - t) - \widehat{m}_{b0}(t;\boldsymbol{\alpha}) - \boldsymbol{\beta}^T \mathbf{Z}_i - \boldsymbol{\gamma}^T \boldsymbol{\xi}_i] dH(t),$$

where $H(t)$ is an increasing and known weight function on $[0, \tau]$. Substituting $\widehat{m}_{b0}(t;\boldsymbol{\alpha})$ to the above estimating function, we have

$$\mathbf{U}_b^*(\boldsymbol{\alpha}) = \sum_{i=1}^{n}\int_0^{\tau}\frac{\Delta_i I(X_i>t)}{\widehat{G}(X_i)}\begin{pmatrix}\mathbf{Z}_i-\bar{\mathbf{Z}}_b(t)\\ \\ \boldsymbol{\xi}_i-\bar{\boldsymbol{\xi}}_b(t)\end{pmatrix}[(X_i-t)-\boldsymbol{\beta}^T\mathbf{Z}_i-\boldsymbol{\gamma}^T\boldsymbol{\xi}_i]dH(t), \tag{S3.2}$$

where $H(t)$ is an increasing and known weight function on $[0,\tau]$,

$$\begin{aligned}\bar{\mathbf{Z}}_b(t) &= \frac{\sum_{i=1}^n \Delta_i I(X_i>t)\widehat{G}^{-1}(X_i)\mathbf{Z}_i}{\sum_{i=1}^n \Delta_i I(X_i>t)\widehat{G}^{-1}(X_i)},\\ \bar{\boldsymbol{\xi}}_b(t) &= \frac{\sum_{i=1}^n \Delta_i I(X_i>t)\widehat{G}^{-1}(X_i)\boldsymbol{\xi}_i}{\sum_{i=1}^n \Delta_i I(X_i>t)\widehat{G}^{-1}(X_i)}.\end{aligned}$$

Likewise, A simple replacement of $\boldsymbol{\xi}_i$ by $\widehat{\boldsymbol{\xi}}_i(\boldsymbol{\theta})$ in $\mathbf{U}_b^*(\boldsymbol{\alpha})$ would lead to a biased estimator. To reduce bias, for given $\boldsymbol{\theta}$, we propose the corrected estimating function $\mathbf{U}_b(\boldsymbol{\alpha};\boldsymbol{\theta}) = (\mathbf{U}_{b1}(\boldsymbol{\alpha};\boldsymbol{\theta})^T, \mathbf{U}_{b2}(\boldsymbol{\alpha};\boldsymbol{\theta})^T)^T$ with

$$\begin{aligned}\mathbf{U}_{b1}(\boldsymbol{\alpha};\boldsymbol{\theta}) &= \sum_{i=1}^{n}\int_0^{\tau}\frac{\Delta_i I(X_i>t)}{\widehat{G}(X_i)}\{\mathbf{Z}_i-\bar{\mathbf{Z}}_b(t)\}[(X_i-t)-\boldsymbol{\beta}^T\mathbf{Z}_i-\boldsymbol{\gamma}^T\widehat{\boldsymbol{\xi}}_i(\boldsymbol{\theta})]dH(t),\\ \mathbf{U}_{b2}(\boldsymbol{\alpha};\boldsymbol{\theta}) &= \sum_{i=1}^{n}\int_0^{\tau}\frac{\Delta_i I(X_i>t)}{\widehat{G}(X_i)}\{\widehat{\boldsymbol{\xi}}_i(\boldsymbol{\theta})-\bar{\boldsymbol{\xi}}_b^*(t;\boldsymbol{\theta})\}[(X_i-t)-\boldsymbol{\beta}^T\mathbf{Z}_i-\boldsymbol{\gamma}^T\widehat{\boldsymbol{\xi}}_i(\boldsymbol{\theta})]dH(t)\\ &\quad+\sum_{i=1}^{n}\int_0^{\tau}\frac{\Delta_i I(X_i>t)}{\widehat{G}(X_i)}\mathbf{D}(\boldsymbol{\theta})\boldsymbol{\gamma}dH(t),\end{aligned}$$

where

$$\bar{\boldsymbol{\xi}}_b^*(t;\boldsymbol{\theta}) = \frac{\sum_{i=1}^n \Delta_i I(X_i>t)\widehat{G}^{-1}(X_i)\widehat{\boldsymbol{\xi}}_i(\boldsymbol{\theta})}{\sum_{i=1}^n \Delta_i I(X_i>t)\widehat{G}^{-1}(X_i)}.$$

For $\widehat{\boldsymbol{\theta}}$, we use the estimating equation $\mathbf{U}_b(\boldsymbol{\alpha};\widehat{\boldsymbol{\theta}}) = 0$, which results in an explicit form estimator $\widehat{\boldsymbol{\alpha}}_b$ of $\boldsymbol{\alpha}$ as follows:

$$\begin{aligned}\widehat{\boldsymbol{\alpha}}_b \;=\; & \left(\sum_{i=1}^{n} \int_0^{\tau} \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)} \left\{ \begin{bmatrix} \mathbf{Z}_i - \bar{\mathbf{Z}}_b(t) \\ \boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i - \bar{\boldsymbol{\xi}}_b^*(t;\widehat{\boldsymbol{\theta}}) \end{bmatrix}^{\otimes 2} - \begin{bmatrix} \mathbf{0}_{s\times s} & \mathbf{0}_{s\times q} \\ \mathbf{0}_{q\times s} & \mathbf{D}(\widehat{\boldsymbol{\theta}}) \end{bmatrix} \right\} dH(t) \right)^{-1} \\ & \times \sum_{i=1}^{n} \int_0^{\tau} \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)} \begin{bmatrix} \mathbf{Z}_i - \bar{\mathbf{Z}}_b(t) \\ \boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i - \bar{\boldsymbol{\xi}}_b^*(t;\widehat{\boldsymbol{\theta}}) \end{bmatrix} (X_i - t) dH(t).\end{aligned}$$

The estimator of the baseline MRL function is given by $\widehat{m}_{b0}(t) = \widehat{m}_{b0}(t; \widehat{\boldsymbol{\alpha}}_b, \widehat{\boldsymbol{\theta}})$, $0 \le t \le \tau$, where

$$\widehat{m}_{b0}(t;\boldsymbol{\alpha},\boldsymbol{\theta}) = \frac{\sum_{i=1}^n \Delta_i I(X_i > t)\widehat{G}^{-1}(X_i)[(X_i - t) - \boldsymbol{\beta}^T\mathbf{Z}_i - \boldsymbol{\gamma}^T\boldsymbol{\Gamma}(\boldsymbol{\theta})\mathbf{V}_i]}{\sum_{i=1}^n \Delta_i I(X_i > t)\widehat{G}^{-1}(X_i)}.$$

The weight function $H(t)$ plays a similar role to that of the weighted log-rank statistic in survival analysis. Ideally, we would choose $H(t)$ to minimize the variance of $\widehat{\boldsymbol{\alpha}}_b$. However, it turns out to be very difficult to derive an optimal $H(t)$ analytically (Lin et al., 2001; Sun and Zhang, 2009). Thus, our choice of $H(t)$ is somewhat ad hoc. We will investigate the efficiency of $\widehat{\boldsymbol{\alpha}}_b$ using two different $H(t)$ functions in the simulation study.

S3.2 Asymptotic results of CEE2 estimator

To study the asymptotic properties of the proposed CEE2 estimators, in addition to conditions (C1) and (C2), we need the following regularity conditions:

(C5) $\mathbf{Z}_i$ is bounded almost surely; $H(t)$ converges almost surely to a non-random and bounded function $\tilde{H}(t)$ uniformly in $t \in [0,\tau]$; $G(t)$ is continuous; $m_0(t)$ is continuously differentiable on $[0,\tau]$.

(C6) $\mathbf{A}_b$ is nonsingular, and

$$\begin{aligned}
\mathbf{A}_b &= \begin{bmatrix} \mathbf{A}_{b11} & \mathbf{A}_{b12} \\ \mathbf{A}_{b12}^T & \mathbf{A}_{b22} \end{bmatrix}, \\
\mathbf{A}_{b11} &= E\left[\int_0^\tau I(T_i > t)\left\{\mathbf{Z}_i - \mathbf{e}_{bz}(t)\right\}^{\otimes 2} d\tilde{H}(t)\right], \\
\mathbf{A}_{b12} &= E\left[\int_0^\tau I(T_i > t)\left\{\mathbf{Z}_i - \mathbf{e}_{bz}(t)\right\}\left\{\boldsymbol{\xi}_i - \mathbf{e}_{\xi}(t)\right\}^T d\tilde{H}(t)\right], \\
\mathbf{A}_{b22} &= E\left[\int_0^\tau I(T_i > t)\left\{\boldsymbol{\xi}_i - \mathbf{e}_{b\xi}(t)\right\}^{\otimes 2} d\tilde{H}(t)\right],
\end{aligned}$$

where $\mathbf{e}_{bz}(t)$ and $\mathbf{e}_{b\xi}(t)$ are the limits of $\bar{\mathbf{Z}}_b(t)$ and $\bar{\boldsymbol{\xi}}_b(t)$, respectively.

Define

$$\begin{aligned}
N_i^c(t) &= I(X_i \le t, \Delta_i = 0), \\
\widehat{\pi}(t) &= n^{-1}\sum_{i=1}^n I(X_i \ge t), \\
\widehat{\Lambda}^c(t) &= n^{-1}\sum_{i=1}^n \int_0^t \frac{dN_i^c(u)}{\widehat{\pi}(u)}, \\
\widehat{M}_i^c(t) &= N_i^c(t) - \int_0^t I(X_i \ge u)\, d\widehat{\Lambda}^c(u), \\
S^{(0)}(t) &= n^{-1}\sum_{i=1}^n \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)}, \\
\bar{\mathbf{V}}_b(t) &= \frac{\sum_{i=1}^n \Delta_i I(X_i > t)\widehat{G}^{-1}(X_i)\mathbf{V}_i}{\sum_{i=1}^n \Delta_i I(X_i > t)\widehat{G}^{-1}(X_i)},
\end{aligned}$$

$$
\begin{aligned}
\widehat{M}_i(t) &= \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)}[(X_i - t) - \widehat{m}_{b0}(t) - \widehat{\boldsymbol{\beta}}_b^T \mathbf{Z}_i - \widehat{\boldsymbol{\gamma}}_b^T \boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i], \\
\widehat{\mathbf{S}}_1(t) &= n^{-1}\sum_{i=1}^n I(X_i \geq t)\int_0^\tau \widehat{M}_i(u)\{\mathbf{Z}_i - \bar{\mathbf{Z}}_b(u)\}dH(u), \\
\widehat{\mathbf{S}}_2(t) &= n^{-1}\sum_{i=1}^n I(X_i \geq t)\int_0^\tau \left[\widehat{M}_i(u)\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i - \bar{\boldsymbol{\xi}}_b^*(u;\widehat{\boldsymbol{\theta}})\} + \frac{\Delta_i I(X_i > u)}{\widehat{G}(X_i)}\mathbf{D}(\widehat{\boldsymbol{\theta}})\widehat{\boldsymbol{\gamma}}_b\right] dH(u), \\
\widehat{S}_3(t,u) &= n^{-1}\sum_{i=1}^n \widehat{M}_i(t) I(X_i \geq u), \\
\widehat{\mathbf{Q}}_{b1} &= n^{-1}\sum_{i=1}^n \int_0^\tau \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)}\{\mathbf{Z}_i - \bar{\mathbf{Z}}_b(t)\}\mathbf{V}_i^T dH(t), \\
\widehat{\mathbf{Q}}_{b2} &= n^{-1}\sum_{i=1}^n \int_0^\tau \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)}\{\mathbf{V}_i - \bar{\mathbf{V}}_b(t)\}[(X_i - t) - \widehat{\boldsymbol{\beta}}_b^T \mathbf{Z}_i - \widehat{\boldsymbol{\gamma}}_b^T \boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i] dH(t), \\
\widehat{\mathbf{Q}}_{b3} &= n^{-1}\sum_{i=1}^n \int_0^\tau \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)}\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i - \bar{\boldsymbol{\xi}}_b^*(t;\widehat{\boldsymbol{\theta}})\}\mathbf{V}_i^T dH(t), \\
\widehat{Q}_{b4} &= n^{-1}\sum_{i=1}^n \int_0^\tau \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)} dH(t).
\end{aligned}
$$

The following Theorems 3 and 4 establish the asymptotic properties of CEE2-based $\widehat{\boldsymbol{\alpha}}_b$ and $\widehat{m}_{b0}(t)$.

Theorem 3. *Under the regularity conditions (C1), (C2), (C5), and (C6), $\widehat{\boldsymbol{\alpha}}_b$ is consistent to $\boldsymbol{\alpha}_0$, and $n^{1/2}(\widehat{\boldsymbol{\alpha}}_b - \boldsymbol{\alpha}_0)$ has asymptotically a normal distribution with mean zero and covariance matrix that can be consistently estimated by $\widehat{\mathbf{A}}_b^{-1}\widehat{\boldsymbol{\Sigma}}_b\widehat{\mathbf{A}}_b^{-1}$, where $\widehat{\boldsymbol{\Sigma}}_b = n^{-1}\sum_{i=1}^n \widehat{\mathbf{U}}_{bi}^{\otimes 2}$, $\widehat{\mathbf{U}}_{bi} = (\widehat{\mathbf{U}}_{bi1}^T, \widehat{\mathbf{U}}_{bi2}^T)^T$,*

$$
\begin{aligned}
\widehat{\mathbf{U}}_{bi1} &= \int_0^\tau \widehat{M}_i(t)\{\mathbf{Z}_i - \bar{\mathbf{Z}}_b(t)\}dH(t) + \int_0^\tau \frac{\widehat{\mathbf{S}}_1(t)}{\widehat{\pi}(t)} d\widehat{M}_i^c(t) - \widehat{\mathbf{Q}}_{b1}\mathbf{R}_i^\dagger(\widehat{\boldsymbol{\theta}})^T\widehat{\boldsymbol{\gamma}}_b, \\
\widehat{\mathbf{U}}_{bi2} &= \int_0^\tau \left[\widehat{M}_i(t)\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i - \bar{\boldsymbol{\xi}}_b^*(t;\widehat{\boldsymbol{\theta}})\} + \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)}\mathbf{D}(\widehat{\boldsymbol{\theta}})\widehat{\boldsymbol{\gamma}}_b\right] dH(t)
\end{aligned}
$$

$$+\int_0^\tau \frac{\widehat{\mathbf{S}}_2(t)}{\widehat{\pi}(t)} d\widehat{M}_i^c(t) + \mathbf{R}_i^\dagger(\widehat{\boldsymbol{\theta}})\widehat{\mathbf{Q}}_{b2} - \widehat{\mathbf{Q}}_{b3}\mathbf{R}_i^\dagger(\widehat{\boldsymbol{\theta}})^T\widehat{\boldsymbol{\gamma}}_b + \widehat{Q}_{b4}\mathbf{P}_i^\dagger(\widehat{\boldsymbol{\theta}})\widehat{\boldsymbol{\gamma}}_b,$$

$$\widehat{\mathbf{A}}_b = \begin{bmatrix} \widehat{\mathbf{A}}_{b11} & \widehat{\mathbf{A}}_{b12} \\ \widehat{\mathbf{A}}_{b12}^T & \widehat{\mathbf{A}}_{b22} \end{bmatrix},$$

$$\begin{aligned}
\widehat{\mathbf{A}}_{b11} &= n^{-1}\sum_{i=1}^n \int_0^\tau \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)} \{\mathbf{Z}_i - \bar{\mathbf{Z}}_b(t)\}^{\otimes 2} dH(t), \\
\widehat{\mathbf{A}}_{b12} &= n^{-1}\sum_{i=1}^n \int_0^\tau \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)} \{\mathbf{Z}_i - \bar{\mathbf{Z}}_b(t)\}\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i - \bar{\boldsymbol{\xi}}_b^*(t;\widehat{\boldsymbol{\theta}})\}^T dH(t), \\
\widehat{\mathbf{A}}_{b22} &= n^{-1}\sum_{i=1}^n \int_0^\tau \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)} \{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i - \bar{\boldsymbol{\xi}}_b^*(t;\widehat{\boldsymbol{\theta}})\}^{\otimes 2} dH(t) - \widehat{Q}_{b4}\mathbf{D}(\widehat{\boldsymbol{\theta}}).
\end{aligned}$$

Theorem 4. *Under the regularity conditions (C1), (C2), (C5), and (C6), $\widehat{m}_{b0}(t)$ converges in probability to $m_0(t)$ uniformly in $t \in [0,\tau]$, and $n^{1/2}\{\widehat{m}_{b0}(t) - m_0(t)\}$ converges weakly on $[0,\tau]$ to a zero-mean Gaussian process whose covariance function at (t,s) can be consistently estimated by $\widehat{\Upsilon}_b(t,s) = n^{-1}\sum_{i=1}^n \widehat{O}_{bi}(t)\widehat{O}_{bi}(s)$, where*

$$\widehat{O}_{bi}(t) = S^{(0)}(t)^{-1}\left[\widehat{M}_i(t) + \int_0^\tau \frac{\widehat{S}_3(t,u)}{\widehat{\pi}(u)} d\widehat{M}_i^c(u)\right] - \widehat{\mathbf{L}}_b(t)^T\widehat{\mathbf{A}}_b^{-1}\widehat{\mathbf{U}}_{bi} - \bar{\mathbf{V}}_b(t)^T\mathbf{R}_i^\dagger(\widehat{\boldsymbol{\theta}})^T\widehat{\boldsymbol{\gamma}}_b,$$

and

$$\widehat{\mathbf{L}}_b(t) = \left(\bar{\mathbf{Z}}_b(t)^T, \bar{\mathbf{V}}_b(t)^T\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})^T\right)^T.$$

Proof of Theorem 3.

We first define

$$M_i(t) = \frac{\Delta_i I(X_i > t)}{G(X_i)}[(X_i - t) - m_0(t) - \boldsymbol{\beta}_0^T \mathbf{Z}_i - \boldsymbol{\gamma}_0^T \boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i].$$

Since $E\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i|\boldsymbol{\xi}_i\} = \boldsymbol{\xi}_i$, then $M_i(t)$ are zero-mean stochastic processes, $i = 1, \cdots, n$, we can rewrite

$$\mathbf{U}_{b1}(\boldsymbol{\alpha}_0; \boldsymbol{\theta}_0) = \sum_{i=1}^{n} \int_0^{\tau} M_i(t) \frac{G(X_i)}{\widehat{G}(X_i)} \{\mathbf{Z}_i - \bar{\mathbf{Z}}_b(t)\} dH(t).$$

By the uniform consistency of $\widehat{G}(\cdot)$, and the uniform strong law of large numbers (USLLN), $\bar{\mathbf{Z}}_b(t)$ converges almost surely to a bounded $\mathbf{e}_{bz}(t)$ uniformly in $t \in [0, \tau]$. By the fact that $P(C \geq \tau) > 0$, we have

$$\sup_{0 \leq i \leq n} \left| \frac{G(X_i)}{\widehat{G}(X_i)} - 1 \right| = O_p(n^{-1/2}).$$

Since $M_i(t)$ can be written as the sums or products of monotone function of t, it is managable, then by the functional central limit theorem (Pollard, 1990, P53), $\sum_{i=1}^{n} M_i(t) = O_p(n^{1/2})$ uniformly in $t \in [0, \tau]$, thus we have the following

$$\begin{aligned}
&\left| \sum_{i=1}^{n} \int_0^{\tau} M_i(t) \{\bar{\mathbf{Z}}_b(t) - \mathbf{e}_{bz}(t)\} dH(t) \right| \\
&\leq \int_0^{\tau} |\bar{\mathbf{Z}}_b(t) - \mathbf{e}_{bz}(t)| |\sum_{i=1}^{n} M_i(t)| dH(t) \\
&\leq \sup_{0 \leq i \leq n} |\bar{\mathbf{Z}}_b(t) - \mathbf{e}_{bz}(t)| \int_0^{\tau} |\sum_{i=1}^{n} M_i(t)| dH(t) \\
&= o_p(1) O_p(n^{1/2}) = o_p(n^{1/2}),
\end{aligned}$$

and

$$
\begin{aligned}
&\left|\sum_{i=1}^{n}\int_0^{\tau} M_i(t)\{\bar{\mathbf{Z}}_b(t)-\mathbf{e}_{bz}(t)\}\left(\frac{G(X_i)}{\widehat{G}(X_i)}-1\right)dH(t)\right| \\
&\leq \sup_{0\leq i\leq n}|\bar{\mathbf{Z}}_b(t)-\mathbf{e}_{bz}(t)|\sup_{0\leq i\leq n}\left|\frac{G(X_i)}{\widehat{G}(X_i)}-1\right|\int_0^{\tau}|\sum_{i=1}^{n}M_i(t)|dH(t) \\
&= o_p(1)O_p(n^{-1/2})O_p(n^{1/2}) = o_p(n^{1/2}).
\end{aligned}
$$

Then,

$$
\begin{aligned}
&\sum_{i=1}^{n}\int_0^{\tau} M_i(t)\frac{G(X_i)}{\widehat{G}(X_i)}\{\bar{\mathbf{Z}}_b(t)-\mathbf{e}_{bz}(t)\}dH(t) \\
= &\sum_{i=1}^{n}\int_0^{\tau} M_i(t)\{\bar{\mathbf{Z}}_b(t)-\mathbf{e}_{bz}(t)\}dH(t) \\
&+\sum_{i=1}^{n}\int_0^{\tau} M_i(t)\{\bar{\mathbf{Z}}_b(t)-\mathbf{e}_{bz}(t)\}\left(\frac{G(X_i)}{\widehat{G}(X_i)}-1\right)dH(t)
\end{aligned}
$$

is of order $o_p(n^{1/2})$. By a similar argument, we have

$$
\sum_{i=1}^{n}\int_0^{\tau} M_i(t)\frac{G(X_i)}{\widehat{G}(X_i)}\{\mathbf{Z}_i-\mathbf{e}_{bz}(t)\}d(H-\tilde{H})(t) = o_p(n^{1/2}). \tag{S3.3}
$$

Then,

$$
\begin{aligned}
\mathbf{U}_{b1}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0) &= \sum_{i=1}^{n}\int_0^{\tau} M_i(t)\frac{G(X_i)}{\widehat{G}(X_i)}\{\mathbf{Z}_i-\mathbf{e}_{bz}(t)\}d\tilde{H}(t) \\
&\quad+\sum_{i=1}^{n}\int_0^{\tau} M_i(t)\frac{G(X_i)}{\widehat{G}(X_i)}\{\mathbf{Z}_i-\mathbf{e}_{bz}(t)\}d(H-\tilde{H})(t) \\
&\quad-\sum_{i=1}^{n}\int_0^{\tau} M_i(t)\frac{G(X_i)}{\widehat{G}(X_i)}\{\bar{\mathbf{Z}}_b(t)-\mathbf{e}_{bz}(t)\}dH(t) \\
&= \sum_{i=1}^{n}\int_0^{\tau} M_i(t)\frac{G(X_i)}{\widehat{G}(X_i)}\{\mathbf{Z}_i-\mathbf{e}_{bz}(t)\}d\tilde{H}(t)+o_p(n^{1/2})
\end{aligned}
$$

$$= \sum_{i=1}^{n} \int_0^{\tau} M_i(t)\{\mathbf{Z}_i - \mathbf{e}_{bz}(t)\} d\tilde{H}(t) + \sum_{i=1}^{n} \int_0^{\tau} M_i(t)\{\mathbf{Z}_i - \mathbf{e}_{bz}(t)\} d\tilde{H}(t) \frac{G(X_i) - \widehat{G}(X_i)}{\widehat{G}(X_i)} + o_p(n^{1/2}). \tag{S3.4}$$

By the martingale representation of Kaplan-Meier estimate, the second term on the right hand side (RHS) of (S3.4) equals

$$\begin{aligned}
&\sum_{i=1}^{n} \int_0^{\tau} M_i(t)\{\mathbf{Z}_i - \mathbf{e}_{bz}(t)\} d\tilde{H}(t) \frac{G(X_i) - \widehat{G}(X_i)}{G(X_i)} + o_p(n^{1/2}) \\
&= \sum_{i=1}^{n} \int_0^{\tau} M_i(t)\{\mathbf{Z}_i - \mathbf{e}_{bz}(t)\} d\tilde{H}(t) \\
&\qquad \times \int_0^{X_i} \left[1 + \frac{\widehat{G}(u-) - G(u)}{G(u)} \right] \frac{\sum_{j=1}^{n} dM_j^c(u)}{n\widehat{\pi}(u)} + o_p(n^{1/2}) \\
&= \sum_{i=1}^{n} \int_0^{\tau} M_i(t)\{\mathbf{Z}_i - \mathbf{e}_{bz}(t)\} d\tilde{H}(t) \int_0^{\tau} I(X_i \geq u) \frac{\sum_{j=1}^{n} dM_j^c(u)}{n\widehat{\pi}(u)} + o_p(n^{1/2}) \\
&= \sum_{i=1}^{n} \int_0^{\tau} \frac{\mathbf{S}_1(t)}{\widehat{\pi}(t)} dM_i^c(t) + o_p(n^{1/2}),
\end{aligned} \tag{S3.5}$$

where

$$\mathbf{S}_1(t) = n^{-1} \sum_{i=1}^{n} I(X_i \geq t) \int_0^{\tau} M_i(u)\{\mathbf{Z}_i - \mathbf{e}_{bz}(u)\} d\tilde{H}(u).$$

It follows from the USSLN that $\mathbf{S}_1(t)$ and $\widehat{\pi}(t)$ converge almost surely to nonrandom functions $\mathbf{s}_1(t)$ and $\pi(t)$ uniformly in $t \in [0, \tau]$, respectively. Using Lemma 1 of Lin et al. (2000), (S3.5) becomes

$$\sum_{i=1}^{n} \int_0^{\tau} \frac{\mathbf{s}_1(t)}{\pi(t)} dM_i^c(t) + o_p(n^{1/2}). \tag{S3.6}$$

By (S1.2) of Lemma 1 in Supplementary Material S1, we obtain

$$\begin{aligned}\mathbf{U}_{b1}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}})-\mathbf{U}_{b1}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0) &= -n\widehat{\mathbf{Q}}_{b1}\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})-\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\}^T\boldsymbol{\gamma}_0\\ &= -\sum_{i=1}^n \mathbf{Q}_{b1}\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0+o_p(n^{1/2}). \quad \text{(S3.7)}\end{aligned}$$

By (S3.5)–(S3.7), we have

$$\begin{aligned}\mathbf{U}_{b1}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}) &= \mathbf{U}_{b1}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0)+\mathbf{U}_{b1}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}})-\mathbf{U}_{b1}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0)\\ &= \sum_{i=1}^n \mathbf{U}_{bi1}+o_p(n^{1/2}), \quad \text{(S3.8)}\end{aligned}$$

where

$$\mathbf{U}_{bi1}=\int_0^\tau M_i(t)\{\mathbf{Z}_i-\mathbf{e}_{bz}(t)\}d\tilde{H}(t)+\int_0^\tau \frac{\mathbf{s}_1(t)}{\pi(t)}dM_i^c(t)-\mathbf{Q}_{b1}\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0.$$

$\mathbf{U}_{b2}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0)$ can be written as

$$\begin{aligned}&\sum_{i=1}^n\int_0^\tau\left[M_i(t)\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i-\bar{\boldsymbol{\xi}}_b^*(t;\boldsymbol{\theta}_0)\}+\frac{\Delta_i I(X_i>t)}{G(X_i)}\mathbf{D}(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0\right]\frac{G(X_i)}{\widehat{G}(X_i)}dH(t)\\ &=\sum_{i=1}^n\int_0^\tau\left[M_i(t)\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i-\mathbf{e}_{b\xi}(t)\}+\frac{\Delta_i I(X_i>t)}{G(X_i)}\mathbf{D}(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0\right]\frac{G(X_i)}{\widehat{G}(X_i)}d\tilde{H}(t)\\ &+\sum_{i=1}^n\int_0^\tau\left[M_i(t)\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i-\mathbf{e}_{b\xi}(t)\}+\frac{\Delta_i I(X_i>t)}{G(X_i)}\mathbf{D}(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0\right]\frac{G(X_i)}{\widehat{G}(X_i)}d(H-\tilde{H})(t)\\ &-\sum_{i=1}^n M_i(t)\frac{G(X_i)}{\widehat{G}(X_i)}\{\bar{\boldsymbol{\xi}}_b^*(t;\boldsymbol{\theta}_0)-\mathbf{e}_{b\xi}(t)\}dH(t). \quad \text{(S3.9)}\end{aligned}$$

By the USLLN, we have $\sup_{0\le t\le\tau}|\bar{\boldsymbol{\xi}}_b^*(t;\boldsymbol{\theta}_0)-\mathbf{e}_{b\xi}(t)|=o_p(1)$, thus, by a similar argument to derive (S3.3) we can show the third term of (S3.9) is $o_p(n^{1/2})$, similarly, the second term of (S3.9) is also $o_p(n^{1/2})$. The first term of (S3.9) equals

$$\begin{aligned}
&\sum_{i=1}^{n}\int_{0}^{\tau}\left[M_i(t)\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i-\mathbf{e}_{b\xi}(t)\}+\frac{\Delta_i I(X_i>t)}{G(X_i)}\mathbf{D}(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0\right]d\tilde{H}(t)\\
&+\sum_{i=1}^{n}\int_{0}^{\tau}\left[M_i(t)\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i-\mathbf{e}_{b\xi}(t)\}+\frac{\Delta_i I(X_i>t)}{G(X_i)}\mathbf{D}(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0\right]d\tilde{H}(t)\\
&\qquad\times\frac{G(X_i)-\widehat{G}(X_i)}{\widehat{G}(X_i)}. \qquad (S3.10)
\end{aligned}$$

By a similar argument as in (S3.5), the second term of (S3.10) equals

$$\begin{aligned}
&\sum_{i=1}^{n}\int_{0}^{\tau}\left[M_i(t)\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i-\mathbf{e}_{b\xi}(t)\}+\frac{\Delta_i I(X_i>t)}{G(X_i)}\mathbf{D}(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0\right]d\tilde{H}(t)\\
&\times\int_{0}^{\tau}I(X_i\geq u)\frac{\sum_{j=1}^{n}dM_j^c(u)}{n\widehat{\pi}(u)}+o_p(n^{1/2})\\
&=\sum_{i=1}^{n}\int_{0}^{\tau}\frac{\mathbf{S}_2(t)}{\widehat{\pi}(t)}dM_i^c(t)+o_p(n^{1/2}), \qquad (S3.11)
\end{aligned}$$

where

$$\mathbf{S}_2(t)=n^{-1}\sum_{i=1}^{n}I(X_i\geq t)\int_{0}^{\tau}\left[M_i(u)\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i-\mathbf{e}_{b\xi}(u)\}+\frac{\Delta_i I(X_i>u)}{G(X_i)}\mathbf{D}(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0\right]d\tilde{H}(u).$$

By the USLLN, $\mathbf{S}_2(t)$ converges almost surely to nonrandom function $\mathbf{s}_2(t)$ uniformly in $t\in[0,\tau]$, again by Lemma 1 of Lin et al. (2000), the RHS of (S3.11) becomes

$$\sum_{i=1}^{n}\int_{0}^{\tau}\frac{\mathbf{s}_2(t)}{\pi(t)}dM_i^c(t)+o_p(n^{1/2}). \qquad (S3.12)$$

Some algebraic manipulation yields

$$\begin{aligned}
\mathbf{U}_{b2}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}})-\mathbf{U}_{b2}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0) &= n\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})-\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\}\widehat{\mathbf{Q}}_{b2}^{*}-n\widehat{\mathbf{Q}}_{b3}^{*}\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})-\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\}^{T}\boldsymbol{\gamma}_0\\
&\quad+n\widehat{Q}_{b4}\{\mathbf{D}(\widehat{\boldsymbol{\theta}})-\mathbf{D}(\boldsymbol{\theta}_0)\}\boldsymbol{\gamma}_0, \qquad (S3.13)
\end{aligned}$$

where

$$\begin{aligned}
\widehat{\mathbf{Q}}_{b2}^* &= n^{-1}\sum_{i=1}^n \int_0^\tau \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)}\{\mathbf{V}_i - \bar{\mathbf{V}}_b(t)\}[(X_i - t) - \boldsymbol{\beta}_0^T\mathbf{Z}_i - \boldsymbol{\gamma}_0^T\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})\mathbf{V}_i]dH(t),\\
\widehat{\mathbf{Q}}_{b3}^* &= n^{-1}\sum_{i=1}^n \int_0^\tau \frac{\Delta_i I(X_i > t)}{\widehat{G}(X_i)}\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i - \bar{\boldsymbol{\xi}}_b^*(t;\boldsymbol{\theta}_0)\}\mathbf{V}_i^T dH(t).
\end{aligned}$$

By the consistency of $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})$ and $\widehat{G}(\cdot)$, and the law of large numbers, $\widehat{\mathbf{Q}}_{b2}^*$, $\widehat{\mathbf{Q}}_{b3}^*$ and $\widehat{Q}_{b4}$ converge to some nonrandom $\mathbf{Q}_{b2}$, $\mathbf{Q}_{b3}$ and Q_{b4} respectively, then by (S1.2) and (S1.3) of Lemma 1 in Supplementary Material S1, (S3.13) becomes

$$\sum_{i=1}^n \mathbf{R}_i(\boldsymbol{\theta}_0)\mathbf{Q}_{b2} - \sum_{i=1}^n \mathbf{Q}_{b3}\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0 + \sum_{i=1}^n Q_{b4}\mathbf{P}_i(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0. \tag{S3.14}$$

By (S3.10), (S3.12) and (S3.14) we have

$$\begin{aligned}
\mathbf{U}_{b2}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}) &= \mathbf{U}_{b2}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0) + \mathbf{U}_{b2}(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}) - \mathbf{U}_{b2}(\boldsymbol{\alpha}_0;\boldsymbol{\theta}_0)\\
&= \sum_{i=1}^n \mathbf{U}_{bi2} + o_p(n^{1/2}), \qquad\qquad (\text{S3.15})
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{U}_{bi2} &= \int_0^\tau \left[M_i(t)\{\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\mathbf{V}_i - \mathbf{e}_{b\xi}(t)\} + \frac{\Delta_i I(X_i > t)}{G(X_i)}\mathbf{D}(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0\right] d\tilde{H}(t)\\
&\quad + \int_0^\tau \frac{\mathbf{s}_2(t)}{\pi(t)} dM_i^c(t) + \mathbf{R}_i(\boldsymbol{\theta}_0)\mathbf{Q}_{b2} - \mathbf{Q}_{b3}\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0 + Q_{b4}\mathbf{P}_i(\boldsymbol{\theta}_0)\boldsymbol{\gamma}_0.
\end{aligned}$$

Let $\mathbf{U}_{bi} = (\mathbf{U}_{bi1}^T, \mathbf{U}_{bi2}^T)^T$, then it follows from (S3.8) and (S3.15) that

$$\mathbf{U}_b(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}) = \sum_{i=1}^n \mathbf{U}_{bi} + o_p(n^{1/2}). \tag{S3.16}$$

Likewise, $n^{-1}\mathbf{U}_b(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}) \to 0$ in probability and $n^{-1/2}\mathbf{U}_b(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}})$ converges in distribution to a normal random vector with mean zero and covariance matrix $\boldsymbol{\Sigma}_b = E\{\mathbf{U}_{bi}^{\otimes 2}\}$. Note that

$$\widehat{\boldsymbol{\alpha}}_b - \boldsymbol{\alpha}_0 = n^{-1}\widehat{\mathbf{A}}_b^{-1}\ \mathbf{U}_b(\boldsymbol{\alpha}_0;\widehat{\boldsymbol{\theta}}), \tag{S3.17}$$

$\widehat{\mathbf{A}}_b$ converges to some $\mathbf{A}_b$ in probability by the law of large numbers and the consistency of $\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})$ and $\mathbf{D}(\widehat{\boldsymbol{\theta}})$. Then, based on (S3.17), $\widehat{\boldsymbol{\alpha}}_b$ converges in probability to $\boldsymbol{\alpha}_0$, and $n^{1/2}(\widehat{\boldsymbol{\alpha}}_b - \boldsymbol{\alpha}_0)$ is asymptotically normal with mean zero and covariance matrix $\mathbf{A}_b^{-1}\boldsymbol{\Sigma}_b\mathbf{A}_b^{-1}$, which can be consistently estimated by $\widehat{\mathbf{A}}_b^{-1}\widehat{\boldsymbol{\Sigma}}_b\widehat{\mathbf{A}}_b^{-1}$.

Proof of Theorem 4.

By the uniform law of large numbers, $\widehat{\mathbf{L}}_b(t)$ and $\bar{\mathbf{V}}_b(t)$ converge in probability to nonrandom $\mathbf{e}_{bl}(t)$ and $\mathbf{e}_{bv}(t)$ uniformly in $t \in [0,\tau]$, respectively. With the use of (S3.16) and (S3.17), we have

$$\begin{aligned}
& \widehat{m}_{b0}(t;\widehat{\boldsymbol{\alpha}}_b,\widehat{\boldsymbol{\theta}}) - \widehat{m}_{b0}(t;\boldsymbol{\alpha}_0,\widehat{\boldsymbol{\theta}}) \\
= \;& -\widehat{\mathbf{L}}_b(t)^T(\widehat{\boldsymbol{\alpha}}_b - \boldsymbol{\alpha}_0) \\
= \;& -n^{-1}\sum_{i=1}^n \mathbf{e}_{bl}(t)^T\mathbf{A}_b^{-1}\mathbf{U}_{bi} + o_p(n^{-1/2}),
\end{aligned} \tag{S3.18}$$

and using (S1.2) of Lemma 1 in Supplementary Material S1,

$$\widehat{m}_{b0}(t;\boldsymbol{\alpha}_0,\widehat{\boldsymbol{\theta}}) - \widehat{m}_{b0}(t;\boldsymbol{\alpha}_0,\boldsymbol{\theta}_0)$$

$$\begin{aligned} &= -\bar{\mathbf{V}}_b(t)^T\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}}) - \boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\}^T\boldsymbol{\gamma}_0 \\ &= -n^{-1}\sum_{i=1}^n \mathbf{e}_{bv}(t)^T\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0 + o_p(n^{-1/2}). \end{aligned} \tag{S3.19}$$

By the USLLN, $S^{(0)}(t)$ converges almost surely to $s^{(0)}(t)$ uniformly in $t \in [0,\tau]$. Again using the martingale representation of $\widehat{G}(t)$, we have

$$\begin{aligned} &\widehat{m}_{b0}(t;\boldsymbol{\alpha}_0,\boldsymbol{\theta}_0) - m_0(t) \\ &= S^{(0)}(t)^{-1}n^{-1}\sum_{i=1}^n\left[M_i(t) + M_i(t)\frac{G(X_i)-\widehat{G}(X_i)}{\widehat{G}(X_i)}\right] \\ &= S^{(0)}(t)^{-1}n^{-1}\sum_{i=1}^n\left[M_i(t) + M_i(t)\int_0^{X_i}\frac{\widehat{G}(u-)}{G(u)}\frac{\sum_{j=1}^n dM_j^c(u)}{n\widehat{\pi}(u)}\right] + o_p(n^{-1/2}) \\ &= S^{(0)}(t)^{-1}n^{-1}\sum_{i=1}^n\left[M_i(t) + \int_0^\tau\frac{S_3(t,u)}{\widehat{\pi}(u)}dM_i^c(u)\right] + o_p(n^{-1/2}) \\ &= s^{(0)}(t)^{-1}n^{-1}\sum_{i=1}^n\left[M_i(t) + \int_0^\tau\frac{s_3(t,u)}{\pi(u)}dM_i^c(u)\right] + o_p(n^{-1/2}), \end{aligned} \tag{S3.20}$$

where

$$S_3(t,u) = n^{-1}\sum_{i=1}^n M_i(t)I(X_i \geq u),$$

and it converges almost surely to nonrandom $s_3(t,u)$ by the USLLN. Then by (S3.18)–(S3.20), we have

$$n^{1/2}\{\widehat{m}_{b0}(t) - m_0(t)\} = n^{-1/2}\sum_{i=1}^n O_{bi}(t) + o_p(1),$$

where

$$O_{bi}(t) = s^{(0)}(t)^{-1}\left[M_i(t) + \int_0^\tau\frac{s_3(t,u)}{\pi(u)}dM_i^c(u)\right] - \mathbf{e}_{bl}(t)^T\mathbf{A}_b^{-1}\mathbf{U}_{bi} - \mathbf{e}_{bv}(t)^T\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0.$$

Upon checking the finite dimensional convergence and tightness, $n^{1/2}\{\widehat{m}_{b0}(t)-m_0(t)\}$ converges weakly to a zero-mean Gaussian process with covariance function at (t,s) given by $\Upsilon_b(t,s) = E\{O_{bi}(t)O_{bi}(s)\}$, which can be consistently estimated by $\widehat{\Upsilon}_b(t,s)$.

S3.3 Simulation

This part presents the estimation results of the factor analysis model in Simulation 1. A comparison between the empirical performances of CEE1 and CEE2 is also provided.

Parameter estimates of the factor analysis model in Simulation 1

The parameter estimates of the factor analysis model in Simulation 1 are presented in Tables S1 and S2. Table S1 shows that the values of ψ's are slightly underestimated and the CP's are in general less than the nominal level when $n = 500$. The performance of CP's gets improved with an increase in the sample size. For instance, the number of CP's above 95% for Case (II) increases from 1 to 4 when n increases from 500 to 1500 (Tables S1 and S2), and the CP's get improved further when the sample size becomes larger (not reported). Likewise, the biases associated with the variances of the residual errors in the factor analysis get closer to zero when n increases. Thus, a sufficiently large sample size is required to achieve high estimation accuracy because the latent variables introduce additional uncertainty to the estimation procedure. The phenomenon of underestimating the variances of residual errors in a factor analysis with finite sample sizes is also found in the existing works, such as Lee et al. (2003) and Pan et al. (2015). An explicit reason why latent variables cause such underestimation requires further investigation.

Comparison of CEE1 and CEE2 under covariate-independent censoring

Under the setting of Simulation 1 in Section 6.1 of the main text, given that the effect of integrator H is difficult to assess analytically (e.g., Lin et al., 2001; Sun and Zhang, 2009), we conduct simulation based on two different H functions: $H_1(t) = \sum_{i=1}^{n} \Delta_i I(X_i \leq t)$ and $H_2(t) = 0$ if $t < 0$ and 1 otherwise. The results are presented in Tables S3 and S4.

In comparison (for the simulation results of CEE1 please refer to the main text), we find that the performance of estimation is slightly better for CEE2-$H_1(t)$ (CEE2 with weight function $H_1(t)$) than for CEE2-$H_2(t)$ (CEE2 with weight function $H_2(t)$) and CEE1, whereas the performance of CEE2-$H_2(t)$ aligns with that of CEE1. As the sample size increases, the estimation performs better in all the cases.

As for the choice of $H(t)$ for CEE2, a common practice in the literature is to compare it through simulation on a case-by-case basis. For the two types of $H(t)$ considered in this study, $H_1(t) = \sum_{i=1}^{n} \Delta_i I(X_i \leq t)$ takes into account all uncensored observations and assigns them an equal weight, whereas $H_2(t)$ takes 0 if $t < 0$ and 1 otherwise. $H_1(t)$ apparently incorporates more information about the data and is thus expected to achieve a higher estimation efficiency in general. The simulation results shown in Tables S3 and S4 demonstrate that $H_1(t)$ does outperform $H_2(t)$ under the cases considered. However, $H_1(t)$ is relatively computationally demanding.

Comparison of CEE1 and CEE2 under covariate-dependent censoring

Under the setting of censoring rate 73% and case (II): $\boldsymbol{\xi}_i = (\xi_{i1}, \xi_{i2})^T \sim \{\Gamma(4,2) - 2\}\mathbf{I}$ in Section 6.2 of the main text, we again conduct the analysis using CEE2-$H_1(t)$, and CEE2-$H_2(t)$, respectively. The results are presented in Table S5.

Although CEE2-$H_1(t)$ performs better than CEE2-$H_2(t)$, they both produce biased estimation in the presence of covariate-dependent censoring. Specifically, the parameters are consistently underestimated and the coverage rates are significantly lower than the nominal level (95%) under CEE2-$H_1(t)$ and CEE2-$H_2(t)$. Moreover, as the sample size increases, the biases remain the same levels and the coverage rates are even further from the nominal level, implying that CEE2-based estimators, their standard error estimates, and their asymptotic normality are questionable. In contrast, the biases of parameter estimates are close to zero and the coverage rates are close to the nominal level under CEE1. Furthermore, the performance of CEE1 is significantly improved when the sample size increases.

In conclusion, CEE1 performs slightly worse than CEE2 in the presence of covariate-independent censoring but much better than CEE2 when censoring is covariate-dependent. Given that CEE1 does not require specifying the censoring mechanism, such robust feature of CEE1 to censoring mechanisms is expected.

S3.4 Real data comparison between CEE1 and CEE2

For comparison, we re-conducted the analysis using the CEE2-based approach for the data in Section 7 of the main text. Regardless of the use of CEE2-$H_1(t)$ or CEE2-$H_2(t)$, the estimated effects of obesity, blood pressure, and lipid on the MRL function of CKD became nonsignificant. These results contradict the previous findings and those obtained in the medical literature (Song et al., 2008; Pan et al., 2015). A possible reason is that the assumption of covariate-independent censoring in the CEE2 approach is erroneous in the CKD study. Therefore, CEE1 is recommend when the censoring mechanism is unknown.

S4 Weak convergence of $\varphi(t;\mathbf{z})$

Definition of $\tilde{\varphi}_i(t;\mathbf{z})$.

$$
\begin{aligned}
\tilde{\varphi}_i(t;\mathbf{z}) &= I(\widehat{\mathbf{Z}}_i^* \le \mathbf{z})\widehat{\Omega}_i(t) + \int_0^t \widehat{O}_{ai}(s)d\widehat{J}_1(s;\mathbf{z}) - \widehat{\mathbf{J}}_2(t;\mathbf{z})^T\widehat{\mathbf{A}}_a^{-1}\widehat{\mathbf{U}}_{ai} \\
&\quad + \widehat{\mathbf{J}}_3(t;\mathbf{z})^T\mathbf{R}_i^\dagger(\widehat{\boldsymbol{\theta}})^T\widehat{\boldsymbol{\gamma}}_a - \int_0^t \widehat{J}_4(s;\mathbf{z})d\widehat{O}_{ai}(s),
\end{aligned}
$$

where

$$
\begin{aligned}
\widehat{J}_1(t;\mathbf{z}) &= n^{-1}\sum_{i=1}^n I(\widehat{\mathbf{Z}}_i^* \le \mathbf{z})N_i(t), \\
\widehat{\mathbf{J}}_2(t;\mathbf{z}) &= n^{-1}\sum_{i=1}^n I(\widehat{\mathbf{Z}}_i^* \le \mathbf{z})N_i(t)\widehat{\mathbf{Z}}_i^*, \\
\widehat{\mathbf{J}}_3(t;\mathbf{z}) &= n^{-1}\sum_{i=1}^n I(\widehat{\mathbf{Z}}_i^* \le \mathbf{z})N_i(t)\mathbf{V}_i, \\
\widehat{J}_4(t;\mathbf{z}) &= n^{-1}\sum_{i=1}^n I(\widehat{\mathbf{Z}}_i^* \le \mathbf{z})Y_i(t).
\end{aligned}
$$

Proof of the asymptotic equivalence of $\varphi(t;\mathbf{z})$ and $\tilde{\varphi}(t;\mathbf{z})$.

Let $\mathbf{Z}_i^* = (\mathbf{Z}_i^T, \widehat{\boldsymbol{\xi}}_i(\boldsymbol{\theta}_0)^T)^T$. Some algebraic manipulations yield

$$
\begin{aligned}
\varphi(t;\mathbf{z}) &= n^{-1/2}\sum_{i=1}^n I(\mathbf{Z}_i^* \le \mathbf{z})\Omega_i(t;\boldsymbol{\alpha}_0,\boldsymbol{\theta}_0) \\
&\quad + n^{1/2}\int_0^t \{\widehat{m}_{a0}(s) - m_0(s)\}d\widehat{J}_1(s;\mathbf{z}) \\
&\quad + n^{1/2}\widehat{\mathbf{J}}_2(t;\mathbf{z})^T(\widehat{\boldsymbol{\alpha}}_a - \boldsymbol{\alpha}_0)
\end{aligned}
$$

$$
\begin{aligned}
&+n^{1/2}\widehat{\mathbf{J}}_3(t;\mathbf{z})^T\{\boldsymbol{\Gamma}(\widehat{\boldsymbol{\theta}})-\boldsymbol{\Gamma}(\boldsymbol{\theta}_0)\}^T\boldsymbol{\gamma}_0 \\
&-n^{1/2}\int_0^t \widehat{J}_4(s;\mathbf{z})d\{\widehat{m}_{a0}(s)-m_0(s)\} \\
&+n^{-1/2}\sum_{i=1}^n\{I(\widehat{\mathbf{Z}}_i^*\le\mathbf{z})-I(\mathbf{Z}_i^*\le\mathbf{z})\}\Omega_i(t;\boldsymbol{\alpha}_0,\boldsymbol{\theta}_0).
\end{aligned}
$$

Let $J_1(t;\mathbf{z})$,$\mathbf{J}_2(t;\mathbf{z})$, $\mathbf{J}_3(t;\mathbf{z})$ and $J_4(t;\mathbf{z})$ denote the limits of $\widehat{J}_1(t;\mathbf{z})$, $\widehat{\mathbf{J}}_2(t;\mathbf{z})$, $\widehat{\mathbf{J}}_3(t;\mathbf{z})$ and $\widehat{J}_4(t;\mathbf{z})$, respectively. By (S1.2) of Supplementary Material S1, (S2.14) and (S2.18) of Supplementary Material S2, and noting that the last term of the right hand side is asymptotically negligible, we have, uniformly in $t\in[0,\tau]$,

$$
\begin{aligned}
\varphi(t;\mathbf{z}) \quad = \quad & n^{-1/2}\sum_{i=1}^n I(\mathbf{Z}_i^*\le\mathbf{z})\Omega_i(t;\boldsymbol{\alpha}_0,\boldsymbol{\theta}_0) \\
&+n^{-1/2}\sum_{i=1}^n\int_0^t O_{ai}(s)dJ_1(s;\mathbf{z}) \\
&-n^{-1/2}\mathbf{J}_2(t;\mathbf{z})^T\sum_{i=1}^n\mathbf{A}_a^{-1}\mathbf{U}_{ai} \\
&+n^{-1/2}\mathbf{J}_3(t;\mathbf{z})^T\sum_{i=1}^n\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0 \\
&-n^{-1/2}\sum_{i=1}^n\int_0^t J_4(s;\mathbf{z})dO_{ai}(s)+o_p(1),
\end{aligned}
$$

which is a sum of i.i.d. zero-mean random variables for fixed t and $\mathbf{z}$. By the multivariate central limit theorem, $\varphi(t;\mathbf{z})$ converges in finite dimensional distributions to a zero-mean Gaussian process. It follows from the functional central limit theorem that the first, the second and the fifth terms of $\varphi(t;\mathbf{z})$ are tight. The third and fourth terms are tight because

$\mathbf{J}_2(t;\mathbf{z})$ and $\mathbf{J}_3(t;\mathbf{z})$ are deterministic functions and $n^{-1/2}\sum_{i=1}^{n}\mathbf{A}_a^{-1}\mathbf{U}_{ai}$ and $n^{-1/2}\sum_{i=1}^{n}\mathbf{R}_i(\boldsymbol{\theta}_0)^T\boldsymbol{\gamma}_0$ converge in distribution. Thus, $\varphi(t;\mathbf{z})$ converges weakly to a zero-mean Gaussian process, which can be approximated by the zero-mean Gaussian process $\tilde{\varphi}(t;\mathbf{z})$.

S5 Exploratory factor analysis in the CKD study

The structure of factor loading matrix (7.9) in the CKD study can be cross validated by the result of an exploratory factor analysis through well-known software, such as LISREL (Jöreskog and Sörbom, 1996) and Mplus (Muthén and Muthén, 1998-2007). Table S6 shows that the estimated number of factors is four, Factor 1 is clearly interpreted as "obesity" because large loadings are associated with WAIST and BMI, and small loadings are associated with others, and Factors 2, 3, and 4 can be likewise interpreted as "blood pressure," "glycemia," and "lipid," respectively. Consequently, a non-overlapping structure of $\mathbf{B}$ in (7.9) is determined by fixing those with small loadings at zero.

Bibliography

[1] Bickel, P. J., Klaassen, C. A. J., Ritov, Y., and Wellner, J. A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models.* Baltimore: Johns Hopkins University Press.

[2] Browne, M. W. (1974). Generalized least square estimators in the analysis of covariance structures. *South African Statistical Journal*, **8**, 1-24.

[3] Browne, M. W. (1984). Asymptotic distribution free methods in analysis of covariance structure. *British Journal of Mathematical and Statistical Psychology*, **37**, 62-83.

[4] Jöreskog, K. G., and Sörbom, D. (1996). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language.* Scientific Software International.

[5] Liang, K. Y., and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models, *Biometrika*, **73**, 13-22.

[6] Lee, S. Y., Song, X. Y., and Lee, J. C. K. (2003). Maximum likelihood estimation of nonlinear structural equation models with ignorable mossing data. *Journal of Educational and Behavioral Statistics*, **28**, 111-134.

[7] Lin, D. Y., Wei, L. J., Yang, I., and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society, Ser. B*, **62**, 711-730.

[8] Lin, D. Y., Wei, L. J., and Ying, Z. (2001). Semiparametric transformation models for point processes. *Journal of the American Statistical Association*, **96**, 620-628.

[9] Muthén, L. K., and Muthén, B. O. (1998-2007). Mplus User's Guide. Fifth Edition. Los Angeles, CA, Muthén & Muthén.

[10] Pan, D., He, H. J., Song, X. Y., and Sun, L. Q. (2015). Regression analysis of additive hazards model with latent variables. *Journal of the American Statistical Association*, **110**, 1148-1159.

[11] Pollard, D. (1990). *Empirical Processes: Theory and Applications*. Hayward, CA: IMS.

[12] Robins, J. M., and Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. in *AIDS Epidemiology: Methodologic Issues*, eds. N. Jewell, K. Dietz, and V. Farewell, Boston: Birkhauser, pp. 297-331.

[13] Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, **89**, 846-866.

[14] Song, X. Y., Lee, S. Y., and Hser, Y. I. (2008). A two-level structural equation model approach for analyzing multivariate longitudinal responses. *Statistics in Medicine*, **27**, 3017-3041.

[15] Sun, L., and Zhang, Z. (2009). A class of transformed mean residual life models with censored survival data. *Journal of the American Statistical Association*, **104**, 803-815.

[16] van der Laan, M. J., and Robins, J. M. (2003). *Unified methods for Censored Longitudinal Data and Causality*. New York: Springer Verlag.

Table S1: Results of the factor analysis model in Simulation 1, $n = 500$

Case	Par	Bias	SE	SEE	CP	Par	Bias	SE	SEE	CP
(I)	b_{11}	-0.004	0.039	0.038	0.934	ψ_1	-0.003	0.031	0.029	0.933
	b_{21}	-0.004	0.040	0.037	0.940	ψ_2	-0.005	0.030	0.029	0.929
	b_{31}	-0.005	0.039	0.037	0.940	ψ_3	-0.005	0.031	0.029	0.920
	b_{42}	-0.004	0.039	0.038	0.934	ψ_4	-0.005	0.029	0.029	0.951
	b_{52}	-0.004	0.039	0.037	0.942	ψ_5	-0.004	0.030	0.029	0.930
	b_{62}	0.003	0.037	0.037	0.950	ψ_6	-0.005	0.029	0.029	0.942
	ϕ_{12}	0.000	0.051	0.049	0.944					
(II)	b_{11}	0.004	0.043	0.043	0.937	ψ_1	-0.005	0.030	0.031	0.935
	b_{21}	-0.006	0.045	0.043	0.915	ψ_2	-0.005	0.030	0.029	0.945
	b_{31}	-0.007	0.045	0.043	0.940	ψ_3	-0.003	0.029	0.031	0.953
	b_{42}	-0.008	0.044	0.044	0.939	ψ_4	-0.004	0.032	0.030	0.923
	b_{52}	-0.007	0.043	0.044	0.939	ψ_5	-0.005	0.031	0.031	0.936
	b_{62}	-0.008	0.044	0.044	0.922	ψ_6	-0.005	0.031	0.030	0.927
	ϕ_{12}	0.001	0.054	0.051	0.920					
(III)	b_{11}	-0.006	0.036	0.036	0.945	ψ_1	-0.003	0.031	0.029	0.929
	b_{21}	0.001	0.036	0.036	0.948	ψ_2	-0.007	0.031	0.029	0.913
	b_{31}	0.000	0.036	0.036	0.960	ψ_3	-0.006	0.031	0.029	0.922
	b_{42}	-0.002	0.037	0.036	0.934	ψ_4	-0.004	0.031	0.029	0.932
	b_{52}	-0.003	0.036	0.036	0.942	ψ_5	-0.005	0.029	0.029	0.945
	b_{62}	-0.002	0.036	0.036	0.951	ψ_6	-0.006	0.030	0.029	0.932
	ϕ_{12}	0.003	0.044	0.044	0.946					

Table S2: Results of the factor analysis model in Simulation 1, $n = 1500$

Case	Par	Bias	SE	SEE	CP	Par	Bias	SE	SEE	CP
(I)	b_{11}	-0.001	0.022	0.022	0.944	ψ_1	-0.002	0.017	0.017	0.943
	b_{21}	-0.001	0.023	0.022	0.942	ψ_2	-0.002	0.018	0.017	0.942
	b_{31}	-0.001	0.022	0.022	0.949	ψ_3	-0.001	0.017	0.017	0.958
	b_{42}	-0.001	0.022	0.022	0.956	ψ_4	-0.002	0.018	0.017	0.941
	b_{52}	-0.001	0.022	0.022	0.940	ψ_5	-0.002	0.017	0.017	0.942
	b_{62}	-0.001	0.021	0.022	0.953	ψ_6	-0.001	0.017	0.017	0.942
	ϕ_{12}	-0.001	0.029	0.029	0.945					
(II)	b_{11}	-0.001	0.026	0.025	0.941	ψ_1	-0.002	0.017	0.017	0.945
	b_{21}	-0.002	0.026	0.025	0.940	ψ_2	0.000	0.017	0.017	0.951
	b_{31}	-0.002	0.026	0.025	0.940	ψ_3	-0.001	0.017	0.017	0.944
	b_{42}	-0.001	0.025	0.025	0.960	ψ_4	-0.001	0.018	0.017	0.941
	b_{52}	-0.001	0.026	0.025	0.939	ψ_5	-0.002	0.017	0.017	0.958
	b_{62}	-0.001	0.026	0.025	0.950	ψ_6	-0.001	0.018	0.017	0.949
	ϕ_{12}	0.000	0.031	0.030	0.935					
(III)	b_{11}	-0.001	0.022	0.021	0.937	ψ_1	-0.001	0.017	0.017	0.954
	b_{21}	-0.001	0.022	0.021	0.943	ψ_2	-0.002	0.018	0.017	0.935
	b_{31}	-0.002	0.021	0.021	0.945	ψ_3	-0.002	0.018	0.017	0.927
	b_{42}	-0.001	0.021	0.021	0.947	ψ_4	-0.001	0.017	0.017	0.952
	b_{52}	-0.001	0.021	0.021	0.951	ψ_5	-0.002	0.017	0.017	0.947
	b_{62}	0.001	0.021	0.021	0.946	ψ_6	-0.002	0.017	0.017	0.947
	ϕ_{12}	0.003	0.026	0.026	0.945					

Table S3: Results of the MRL model - CEE2 - $n = 500$

			$H_1(t)$				$H_2(t)$			
Case	CR	Par	Bias	SE	SEE	CP	Bias	SE	SEE	CP
(I)	10%	β	0.000	0.039	0.039	0.951	-0.004	0.077	0.078	0.954
		γ_1	-0.001	0.021	0.021	0.952	0.002	0.045	0.043	0.936
		γ_2	-0.005	0.034	0.028	0.936	-0.004	0.045	0.045	0.950
	30%	β	0.001	0.045	0.044	0.946	-0.003	0.083	0.083	0.946
		γ_1	-0.001	0.025	0.024	0.941	0.002	0.047	0.046	0.947
		γ_2	-0.005	0.038	0.032	0.923	-0.004	0.047	0.048	0.954
(II)	10%	β	0.000	0.042	0.039	0.952	-0.004	0.082	0.078	0.941
		γ_1	-0.001	0.027	0.025	0.940	0.004	0.048	0.044	0.934
		γ_2	-0.006	0.039	0.031	0.930	-0.008	0.058	0.050	0.950
	30%	β	0.001	0.048	0.045	0.951	-0.003	0.085	0.083	0.942
		γ_1	-0.004	0.031	0.039	0.946	0.004	0.051	0.047	0.930
		γ_2	-0.006	0.043	0.037	0.942	-0.009	0.061	0.053	0.942
(III)	10%	β	-0.001	0.039	0.039	0.951	-0.003	0.080	0.078	0.938
		γ_1	0.000	0.020	0.020	0.945	0.001	0.043	0.044	0.942
		γ_2	-0.004	0.032	0.026	0.940	-0.001	0.046	0.045	0.951
	30%	β	-0.003	0.045	0.045	0.944	-0.003	0.085	0.083	0.934
		γ_1	0.001	0.024	0.023	0.938	0.002	0.047	0.047	0.942
		γ_2	-0.003	0.037	0.030	0.952	-0.001	0.051	0.048	0.944

Table S4: Results of the MRL model - CEE2 - $n = 1500$

			$H_1(t)$				$H_2(t)$			
Case	CR	Par	Bias	SE	SEE	CP	Bias	SE	SEE	CP
(I)	10%	β	0.000	0.022	0.022	0.945	0.000	0.046	0.045	0.944
		γ_1	0.000	0.012	0.012	0.952	0.000	0.025	0.025	0.950
		γ_2	−0.002	0.017	0.016	0.930	−0.002	0.027	0.026	0.943
	30%	β	0.002	0.026	0.026	0.939	0.000	0.049	0.048	0.938
		γ_1	0.000	0.014	0.014	0.960	0.000	0.027	0.027	0.944
		γ_2	0.000	0.019	0.018	0.929	−0.002	0.029	0.028	0.944
(II)	10%	β	−0.001	0.023	0.023	0.950	0.001	0.044	0.045	0.952
		γ_1	0.000	0.012	0.012	0.953	−0.001	0.025	0.026	0.951
		γ_2	−0.002	0.018	0.018	0.947	−0.003	0.029	0.029	0.949
	30%	β	−0.001	0.026	0.026	0.959	0.001	0.047	0.048	0.948
		γ_1	0.000	0.014	0.014	0.936	−0.001	0.027	0.027	0.948
		γ_2	−0.002	0.021	0.020	0.934	−0.003	0.031	0.031	0.947
(III)	10%	β	0.000	0.023	0.023	0.944	0.000	0.046	0.045	0.943
		γ_1	0.000	0.011	0.012	0.943	0.001	0.027	0.026	0.939
		γ_2	0.000	0.015	0.015	0.956	−0.001	0.027	0.026	0.945
	30%	β	0.000	0.026	0.026	0.946	0.000	0.048	0.048	0.947
		γ_1	0.000	0.013	0.013	0.948	0.001	0.028	0.027	0.942
		γ_2	0.000	0.018	0.017	0.930	−0.001	0.028	0.028	0.947

Table S5: Results of the MRL model - CEE2 - $\boldsymbol{\xi}_i \sim \{\Gamma(4,2)-2\}\mathbf{I}$, CR = 73%

n	Method	Par	Bias	SE	SEE	CP
500	$H_1(t)$	β	-0.040	0.115	0.104	0.897
		γ_1	-0.017	0.049	0.052	0.927
		γ_2	-0.026	0.092	0.087	0.905
	$H_2(t)$	β	-0.275	0.191	0.175	0.686
		γ_1	-0.085	0.085	0.080	0.838
		γ_2	-0.177	0.129	0.123	0.686
1500	$H_1(t)$	β	-0.053	0.060	0.058	0.836
		γ_1	-0.017	0.028	0.027	0.882
		γ_2	-0.034	0.048	0.046	0.843
	$H_2(t)$	β	-0.265	0.098	0.101	0.212
		γ_1	-0.084	0.045	0.046	0.564
		γ_2	-0.176	0.071	0.071	0.280
3000	$H_1(t)$	β	-0.051	0.042	0.041	0.751
		γ_1	-0.018	0.019	0.019	0.831
		γ_2	-0.037	0.033	0.032	0.774
	$H_2(t)$	β	-0.268	0.073	0.071	0.017
		γ_1	-0.084	0.033	0.032	0.260
		γ_2	-0.180	0.049	0.050	0.037

Table S6: The result of the exploratory factor analysis in the CKD study

Variable	Factor1	Factor2	Factor3	Factor4
WAIST	**1.004**	0.014	-0.060	-0.016
BMI	**0.830**	0.009	0.027	-0.008
SBP	0.002	**0.946**	0.037	-0.038
DBP	0.045	**0.643**	-0.004	0.038
HbA1c	0.045	-0.035	**-0.941**	-0.043
FPG	0.003	-0.004	**-0.738**	0.025
TC	-0.109	0.120	-0.155	**0.358**
HCL-C	-0.177	0.091	-0.093	**-0.416**
TG	-0.024	-0.020	0.024	**1.014**