Supplementary Materials for Generalized scale-change models for recurrent event processes under informative censoring

Gongjun Xu\textsuperscript{1}, Sy Han Chiou\textsuperscript{2}, Jun Yan\textsuperscript{3}, Kieren Marr\textsuperscript{4}, and Chiung-Yu Huang\textsuperscript{5}. \\
\textsuperscript{1}University of Michigan, \textsuperscript{2}University of Texas at Dallas, \textsuperscript{3}University of Connecticut, \textsuperscript{4}Johns Hopkins University, and \textsuperscript{5}University of California, San Francisco

Abstract: The Supplementary materials contains the proof of Theorem 1 in Section S1, additional simulation results in Section S2, and identifiability assumption check for the infection data in Sections S3.

S1. Proof of Theorem 1

S1.1 Asymptotic linearity of $\hat{\Lambda}_n(t; a)$. We first show the asymptotic linearity of $\hat{\Lambda}_n(t; a)$, which will be used to establish the asymptotic normality of $n^{1/2}(\hat{\alpha}_n - \alpha)$ and $n^{1/2}(\hat{\beta}_n - \beta)$. For any fixed $a$, let $\mathcal{N}_n(t; a) = \frac{1}{n} \sum_{i=1}^{n} N_i^*(t, a) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m_i} I\{t_{ij}^*(a) \leq t \land Y_i^*(a)\}$ and recall that

$$\mathcal{R}^{(k)}_n(t; a) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m_i} X_{ij}^k I\{t_{ij}^*(a) \leq t \leq Y_i^*(a)\}$$
for $k = 0$ and 1. Define stochastic processes $N(u; a)$ and $R^{(k)}(u; a)$ as follows:

$$
N(u; a) := E\left\{N_i(u e^{-X_i^T a} \wedge Y_i)\right\}
$$

$$
= E\left[ E\left\{ Z_i \Lambda_0(u e^{X_i^T (\alpha-a)} \wedge Y_i e^{X_i^T (\beta-a)}) e^{X_i^T (\beta-a)} \mid Z_i, Y_i, X_i \right\} \right]
$$

$$
= E\left( E\left[ \int_0^u Z_i I\{Y_i^*(a) \geq v\} \, d\Lambda_0\{v e^{X_i^T (\alpha-a)} \} \times e^{X_i^T (\beta-a)} \mid Z_i, Y_i, X_i \right] \right)
$$

$$
= E\left[ \int_0^u Z_i I\{Y_i^*(a) \geq v\} \, d\Lambda_0\{v e^{X_i^T (\alpha-a)} \} \times e^{X_i^T (\beta-a)} \right],
$$

and

$$
R^{(k)}(u; a) := E\left[ X_i^k \sum_{j=1}^{m_i} I\{t_{ij}^*(a) \leq u \leq Y_i^*(a)\} \right]
$$

$$
= E\left( E[X_i^k I\{Y_i^*(a) \geq u\} N(u e^{-X_i^T a}) \mid Y_i, Z_i, X_i] \right)
$$

$$
= E\left[ I(Y_i^*(a) \geq u) Z_i \Lambda_0\{u e^{X_i^T (\alpha-a)} \} X_i^k e^{X_i^T (\beta-a)} \right].
$$

When $a$ equals the true parameter $\alpha$, we have $N(u; \alpha) = \int_0^u E[Z_i I\{Y_i^*(a) \geq v\} e^{X_i^T (\beta-a)}] \, d\Lambda_0(v)$ and $R^{(k)}(u; \alpha) = E[Z_i I\{Y_i^*(\alpha) \geq u\} X_i^k e^{X_i^T (\beta-a)}] \Lambda_0(u)$. We have

$$
H(t) = \log \Lambda_0(t) = - \int_t^\tau \Lambda_0^{-1}(u) \, d\Lambda_0(u) = \int_t^\tau (R^{(0)})^{-1}(u; \alpha) \, dN(u; \alpha).
$$

This also establishes the mean zero property of the process $M_i^*(t)$ in Equation (4) of the main text.

Let $\Lambda_0(t, a) = \exp\{- \int_t^\tau (R^{(0)})^{-1}(u; a) \, dN(u; a)\}$. When $a = \alpha$, we have $\Lambda_0(t) = \Lambda_0(t, \alpha)$. By the uniform strong law of large numbers (van der Vaart and Wellner, 1996), we have $N_n(u, a) \to N(u, a)$ and $R^{(k)}_n(u, a) \to R^{(k)}(u, a)$ a.s. uniformly in $a$ and $u \in [0, \tau]$; furthermore we have $\|N_n(u, a) - N(u, a)\| = O_p(n^{-1/2})$ and $\|R^{(k)}_n(u, a) - R^{(k)}(u, a)\| = O_p(n^{-1/2})$. 

\( R^{(k)}(u, a) \| = O_p(n^{-1/2}) \), where \( \| \cdot \| \) denotes the supremum norm. Arguing as in the proof of Theorem 1 in Wang et al. (2001) and Xu et al. (2017), we have

\[
\int_t^\tau \frac{dN_n(u; a)}{R_n^{(0)}(u; a)} = \int_t^\tau \frac{dN(u; a)}{R^{(0)}(u; a)} - \int_t^\tau \frac{\{R_n^{(0)}(u; a) - R^{(0)}(u; a)\} dN(u; a)}{R^{(0)}(u; a)^2} \\
+ \int_t^\tau \frac{d[N_n(u; a) - N(u; a)]}{R^{(0)}(u; a)} + o_p(n^{-1/2}) \\
= \int_t^\tau \frac{dN(u; a)}{R^{(0)}(u; a)} - \frac{1}{n} \sum_{i=1}^n \eta_i(t; a) + o_p(n^{-1/2}),
\]

where

\[
\eta_i(t; a) = \sum_{j=1}^{m_n} \int_t^\tau I\{t_{ij}^*(a) \leq u \leq Y_j^*(a)\} \frac{dN(u; a)}{R^{(0)}(u; a)^2} - \sum_{j=1}^{m_n} \int_t^\tau I\{t_{ij}^*(a) \leq u\} \frac{dI\{t_{ij}^*(a) \leq u\}}{R^{(0)}(u; a)}.
\]

Note that \( E(\eta_i) = 0 \). This gives an asymptotic i.i.d. representation of \( n^{1/2}\{\hat{\Lambda}_n(a, t) - \Lambda_0(a, t)\} \):

\[
n^{1/2}\{\hat{\Lambda}_n(t, a) - \Lambda_0(t, a)\} = n^{-1/2}\Lambda_0(t, a) \sum_{i=1}^n \eta_i(t, a) + o_p(1). \tag{S1.1}
\]

We next show the asymptotic linearity results for \( n^{1/2}\{\hat{\Lambda}_n(t; a) - \hat{\Lambda}_n(t; \alpha)\} \). Applying the techniques in the proof of Theorem 1 in Ying (1993), for some positive sequence \( d_n \to 0 \) and \( \|a - \alpha\| \leq d_n \), we have uniformly in \( a \) and \( t \in [0, \tau] \),

\[
n^{1/2}\left\{ \int_t^\tau \frac{dN_n(u; a)}{R_n^{(0)}(u; a)} - \int_t^\tau \frac{dN_n(u; a)}{R_n^{(0)}(u; a)} \right\} = n^{1/2}\left\{ \int_t^\tau \frac{dN(u; a)}{R^{(0)}(u; a)} - \int_t^\tau \frac{dN(u; a)}{R^{(0)}(u; a)} \right\} + o_p(n^{1/2}\|a - \alpha\| + 1).
\]

Furthermore, we have asymptotic approximation

\[
n^{1/2}\left\{ \int_t^\tau \frac{dN(u; a)}{R^{(0)}(u; a)} - \int_t^\tau \frac{dN(u; a)}{R^{(0)}(u; a)} \right\} = \kappa(t) n^{1/2}(a - \alpha) + o(n^{1/2}\|a - \alpha\| + 1),
\]
where $\kappa(t)$ is the corresponding derivative matrix given by

$$
\kappa(t) = \int_t^\tau E \frac{XZI \{ Y^*(\alpha) \geq u \}}{R^{(0)}(u; a)} \, d\lambda_0(u) + \int_t^\tau \frac{\partial E[ZI \{ Y^*(\alpha) \geq u \} \, R^{(0)}(u; a)^{-1}]}{\partial a} \Big|_{a=\alpha} \, d\Lambda_0(u).
$$

Therefore, we have uniformly for $\|a - \alpha\| \leq d_n \to 0$ and $t \in [0, \tau]$

$$
n^{1/2} \left\{ \int_t^\tau \frac{dN_n(u; a)}{R_n^{(0)}(u; a)} - \int_t^\tau \frac{dN_n(u; a)}{R_n^{(0)}(u; \alpha)} \right\} = \kappa(t)^\top n^{1/2}(a - \alpha) + o_p(n^{1/2}\|a - \alpha\| + 1).
$$

This further implies

$$
n^{1/2} \left\{ \hat{H}_n(t, a) - \hat{H}_n(t, \alpha) \right\} = \kappa(t)^\top n^{1/2}(a - \alpha) + o_p(n^{1/2}\|a - \alpha\| + 1)
$$

and $n^{1/2} \{ \hat{\Lambda}_n(t, a) - \hat{\Lambda}_n(t, \alpha) \} = \Lambda_0(t, \alpha) \kappa(t)^\top n^{1/2}(a - \alpha) + o_p(n^{1/2}\|a - \alpha\| + 1)$, uniformly for $\|a - \alpha\| \leq d_n \to 0$ and $t \in [0, \tau]$.

**S1.2 Asymptotic results of \(\hat{\alpha}_n\) and \(\hat{\Lambda}_n(t, \hat{\alpha}_n)\).** We show the asymptotic normality of $n^{1/2}(\hat{\alpha}_n - \alpha)$. First, a similar discussion as the previous section gives that

$$
n^{1/2} \left\{ \int_t^\tau \frac{R_n^{(1)}(u; a)}{R_n^{(0)}(u; a)} \, d\lambda_n(u) \right\} - \int_t^\tau \frac{R_n^{(1)}(u; a)}{R_n^{(0)}(u; a)} \, d\lambda_n(u; \alpha) \right\} = \{\kappa(1)(t)\}^\top n^{1/2}(a - \alpha) + o(n^{1/2}\|a - \alpha\| + 1),
$$

where $\kappa(1)(t)$ is the corresponding derivative matrix given by

$$
\kappa(1)(t) = \int_t^\tau E \frac{R^{(1)}(u; a)XZI \{ Y^*(\alpha) \geq u \}}{R^{(0)}(u; a)} \, d\lambda_0(u) + \int_t^\tau \frac{\partial E[ZI \{ Y^*(\alpha) \geq u \} \, R^{(1)}(u; a)R^{(0)}(u; a)^{-1}]}{\partial a} \Big|_{a=\alpha} \, d\Lambda_0(u).
$$

Therefore from the estimating equation, we have

$$
n^{1/2}(\hat{\alpha}_n - \alpha) = J^{-1} n^{1/2} S_n(\alpha) + o_p(1),
$$
where \( J \) is defined as \( J = \kappa^{(1)}(0) \) and

\[
S_n(\alpha) = n^{-1} \sum_{i=1}^{n} e_i(\alpha) + o_p(n^{-1/2}),
\]

where \( e_i(\alpha) = \sum_{j} \int_{0}^{\tau} X_i dM_{ij}^*(\tau) \). By the central limit theorem, this implies that \( n^{1/2}(\hat{\alpha}_n - \alpha) \) converges weakly to a multivariate normal distribution with mean zero and variance \( \Sigma_{\alpha} = J^{-1}E[e_i(\alpha)e_i(\alpha)^\top](J^{-1})^\top \).

The consistency of \( \hat{H}_n(t, \hat{\alpha}_n) \) follows from that of \( \hat{\alpha}_n \). From the asymptotic linearity results for \( n^{1/2}\{\hat{H}_n(t, \hat{\alpha}_n) - \hat{H}_n(t, \alpha)\} \) and \( n^{1/2}\{\hat{H}_n(t, \alpha) - H(t)\} \) in the preceding section, we have uniformly for \( t \in [0, \tau] \)

\[
n^{1/2}\{\hat{H}_n(t, \hat{\alpha}_n) - H(t)\} = n^{1/2}\{\hat{H}_n(t, \hat{\alpha}_n) - \hat{H}_n(t, \alpha)\} + n^{1/2}\{\hat{H}_n(t, \alpha) - H(t)\}
\]

\[
= n^{-1/2}\sum_{i=1}^{n} \{\kappa^{(1)}(t)^\top J^{-1} e_i(\alpha) + \eta_i(t; \alpha)\} + o_p(1).
\]

Applying the functional central limit theorem, we have the weak convergence of \( n^{1/2}\{\hat{H}_n(t, \hat{\alpha}_n) - H(t)\} \) to a mean-zero Gaussian process for \( t \in [0, \tau] \). This also gives the asymptotic normality of the process \( n^{1/2}\{\hat{\Lambda}_n(t, \hat{\alpha}_n) - \Lambda_0(t)\} \).

**S1.3 Asymptotic results of \( \hat{\theta}_n \)** From the asymptotic linearity property of \( \hat{\Lambda}_n \) developed in the last section, we have

\[
\hat{\Lambda}_n\{Y_i^*(a)\} - \hat{\Lambda}_n\{Y_i^*(\alpha)\} = \{1 + o_p(1)\} \Lambda_0\{Y_i^*(\alpha)\} \kappa\{Y_i^*(\alpha)\}^\top Y_i e^{X_i^\top \alpha} X_i^\top (a - \alpha) + o_p(n^{-1/2}).
\]

For any sequence \( d_n \to 0 \), consider \( r \) in a neighborhood of \( \theta \) such that \( \|a - \alpha\| \leq d_n \)
and \( \|r - \theta\| \leq d_n \). Thus, the following result holds uniformly for the considered \( r \)

\[
U_n(r, \hat{\alpha}_n) - U_n(\theta, \alpha) = \frac{1}{n} \sum_{i=1}^{n} \bar{X}_i^T m_i \hat{\Lambda}_n^{-1} \{Y_i^*(\alpha)\} \kappa \{Y_i^*(\alpha)\}^\top Y_i e^{X_i^\top \alpha} X_i^\top (\hat{\alpha}_n - \alpha) + \frac{1}{n} \sum_{j=1}^{n} \bar{X}_j^\top \exp(\bar{X}_j^\top \theta) \bar{X}_j^\top (r - \theta) + o_p(n^{-1/2} + \|r - \theta\|) \\
:= J_1(\hat{\alpha}_n - \alpha) + J_2(r - \theta) + o_p(n^{-1/2} + \|r - \theta\|),
\]

where \( J_1 \) and \( J_2 \) are as defined in the above displays. We can further derive the normality of \( n^{1/2}U_n(\theta, \alpha) \). In particular, \( U_n(\theta, \alpha) \) can be written as

\[
U_n(\theta, \alpha) = \frac{1}{n} \sum_{i=1}^{n} \bar{X}_i^\top m_i \hat{\Lambda}_n^{-1} \{Y_i^*(\alpha)\} - \frac{1}{n} \sum_{i=1}^{n} \bar{X}_i^\top \exp(\bar{X}_i^\top \theta) \\
= \frac{1}{n} \sum_{i=1}^{n} \bar{X}_i^\top m_i \hat{\Lambda}_n^{-1} \{Y_i^*(\alpha)\} - \Lambda_0^{-1} \{Y_i^*(\alpha)\} \\
+ \frac{1}{n} \sum_{i=1}^{n} \bar{X}_i^\top [m_i \Lambda_0^{-1} \{Y_i^*(\alpha)\} - \exp(\bar{X}_i^\top \theta)].
\]

From (S1.1), the first part of the above display can be further written as

\[
\frac{1}{n} \sum_{i=1}^{n} \bar{X}_i^\top m_i \hat{\Lambda}_n^{-1} \{Y_i^*(\alpha)\} - \Lambda_0^{-1} \{Y_i^*(\alpha)\} \\
= -\frac{1}{n} \sum_{i=1}^{n} \bar{X}_i^\top m_i \Lambda_0^{-1} \{Y_i^*(\alpha)\} \eta_i \{y^*(\alpha)\} + o_p(n^{-1/2}) \\
= -\frac{1}{n} \sum_{i=1}^{n} \int \bar{x}^\top m \Lambda_0^{-1} \{y_i^*(\alpha)\} \eta_i \{y^*(\alpha)\} dV(x, y, m) + o_p(n^{-1/2}),
\]

where \( V(x, y, m) \) denotes the joint distribution function of \( (X, Y, m) \) (e.g., Proof of (9) in Wang et al. [2001]). Therefore,

\[
n^{1/2}U_n(\theta, \alpha) = n^{-1/2} \sum_{i=1}^{n} d_i(\theta, \alpha) + o_p(1)
\]
where
\[ d_i(\theta, \alpha) = -\int x^\top m \Lambda_0^{-1} \{y_i^*(\alpha)\} \eta_i \{y^*(\alpha)\} \, dV(x, y, m) + \bar{X}_i \Lambda_0^{-1} Y^*_i - \exp(\bar{X}_i^\top \theta). \]

This gives the normality of \( n^{1/2}U_n(\theta, \alpha) \) with asymptotic mean 0 and covariance matrix \( E[d_i(\theta, \alpha)d_i(\theta, \alpha)^\top] \).

Together with the result that \( n^{1/2}(\hat{\alpha}_n - \alpha) = J^{-1}n^{-1/2} \sum_{i=1}^n e_i(\alpha) + o_p(1) \), we have
\[ n^{1/2}(\hat{\theta} - \theta) = n^{-1/2} J_2^{-1} \sum_{i=1}^n \{d_i(\theta, \alpha) - J_1J^{-1} e_i(\alpha)\} + o_p(1) \]
and it converges weakly to a multivariate normal distribution with mean zero and variance
\[ \Sigma_\theta = J_2^{-1} E[(d_i(\theta, \alpha) - J_1J^{-1} e_i(\alpha))(d_i(\theta, \alpha) - J_1J^{-1} e_i(\alpha))^\top](J_2^{-1})^\top. \]

In particular, \( \{n^{1/2}(\hat{\alpha}_n - \alpha), n^{1/2}(\hat{\theta} - \theta)\} \) jointly converges weakly to a multivariate normal distribution with mean zero and covariance matrix induced by the above linear asymptotic approximations that
\[ \Sigma(\alpha, \theta) = \begin{pmatrix} \Sigma_\alpha & \text{Cov}(\alpha, \theta) \\ \text{Cov}(\alpha, \theta) & \Sigma_\theta \end{pmatrix}, \]
where \( \text{Cov}(\alpha, \theta) = \text{Cov}[J^{-1}e_i(\alpha), J_2^{-1}\{d_i(\theta, \alpha) - J_1J^{-1} e_i(\alpha)\}] \). As a consequence, we have the asymptotic normality of \( \{n^{1/2}(\hat{\alpha}_n - \alpha), n^{1/2}(\hat{\beta} - \beta)\} \) with mean zero and covariance matrix
\[ \Sigma(\alpha, \beta) = \begin{pmatrix} I_p & 0 & 0_p \\ I_p & 0 & I_p \\ I_p & 0 & I_p \end{pmatrix} \Sigma(\alpha, \theta) \begin{pmatrix} I_p & 0 & 0_p \\ I_p & 0 & I_p \end{pmatrix}^\top. \]
S2. Additional simulation results

Table 1 presents the simulation results under an accelerated rate model. For this scenario, regardless of the choice of the initial value, the proposed estimator is fairly unbiased, with the proposed variance estimator in close agreement to the empirical counterpart. The corresponding empirical coverage probabilities are reasonably close to the nominal level of 95%. On the other hand, the estimator of Sun and Su (2008) requires the initial value to be specified at the true value, in which case it yields moderate bias when the noninformative censoring assumption is violated.

When the data generating model reduces to the accelerated means model ($\alpha = \beta$), in addition to the proposed estimator and the conventional estimator of Sun and Su (2008), we also report the result from the estimator of Xu et al. (2017). Table 2 shows the proposed estimator outperforms the conventional estimator of Sun and Su (2008) under both the noninformative and informative censoring setting. The estimator of Xu et al. (2017) also yields small biases. The standard errors of the estimator of Xu et al. (2017) were obtained from an efficient resampling variance estimator that shares the similar spirit as the proposed version. Comparing to the proposed estimator, the estimator of Xu et al. (2017) has a slightly larger bias but smaller standard errors. This biases-efficiency trade-off comes from the additional model assumption on $\alpha = \beta$ that is imposed in Xu et al. (2017). Nonetheless, both of these estimators accommodate informative censoring and yield comparable coverage probabilities. Fig-
Figure 1 presents the estimates and the empirical pointwise 95% confidence intervals for the baseline cumulative rate function with \( n = 200 \). The averages of the estimated baseline cumulative rate function are indistinguishable from the truth for the cases considered.

We extended the simulation studies to explore scenarios with different variance configuration for the subject-specific latent variable. Specifically, we considered scenarios with \( Z \) generated from gamma distributions with mean 1 and variance 0.5 or 1. Since the proposed estimator performs well regardless of the choice of the initial value, we only present the results that use the zero vector as the initial in Table 3. The standard errors were estimated through the proposed resampling approach with 200 bootstrap size. For all scenarios considered, our estimator remains virtually unbiased, with estimated standard errors reasonably close to the empirical standard errors. The magnitude of the estimated standard errors seem to increase with the variance of the frailty variable, but the empirical coverage rates remain close to the nominal level in all scenarios.

[Table 1 about here.]
[Table 2 about here.]
[Table 3 about here.]
[Figure 1 about here.]
A small-scale simulation study was conducted to examine the performance of the proposed method under scenarios where a smaller number of events are observed. We considered the simulation settings described in the manuscript, but set \( \tau = 10 \) and generated the censoring time from an exponential distribution with mean \( 7e^{-X_1}/Z \). Simulation results with \( \alpha = \beta = (-1, 1)^\top \) are presented in Table 4. Under this setting, the average number of observed recurrent event was about 0.6 per subject, which is smaller than what we observed in the two transplant cohorts. The proposed estimators are virtually unbiased with only a few exceptions with \( n = 200 \) and the bias the exception cases quickly diminishes as the sample size increases. The estimated standard errors are reasonably close to the empirical standard error, with better agreement with \( n = 400 \). The coverage rates are closer to the nominal level of 95% when \( n = 400 \). These results suggest that our model gives valid inference even when the average number of observed recurrent event is small.

[Table 4 about here.]

Lastly, we also considered situations when the recurrent events were generated from a non-Poisson process given the frailty by altering the interarrival time distribution from exponential and adding a second latent variable to the rate function. In particular, we generated the recurrent events from the inversion algorithm for nonstationary renewal processes (e.g., Gerhardt and Nelson, 2009). Let \( G \) be the distribution of the scaled interarrival time such that the resulting renewal process has rate 1. For
example, $G$ can be unit exponential (for Poisson process) or uniform over $(0, 2)$ (for a non-Poisson process). Let $\Lambda_{ij}(t) = \int_0^t \lambda_{ij}(u) \, du$, where

$$\lambda_{ij}(t) = Z_i \zeta_{ij} \lambda_0(t e^{\alpha_1 X_1 + \alpha_2 X_2}) e^{\beta_1 X_1 + \beta_2 X_2},$$

$\lambda_0(t) = 1/(5+t)$, and $\zeta_{ij}$'s are additional latent variables, independent of $Z_i$, following a gamma distribution with mean 1 and variance 0.5. If $\zeta_{ij}$'s are dependent within each $i$, the process allows for association of the arrival time of the recurrent events within the $i$th subject. The data generation algorithm is outlined as follows:

Step 1. Initialize $\epsilon_0 = 0$.

Step 2. Generate a random number $\eta$ from $G$.

Step 3. Set $t_{ij} = \Lambda_{ij}^{-1}(\epsilon_j)$, where $\epsilon_j = \epsilon_{j-1} + \eta$, for $j \geq 1$, and $\Lambda_{ij}^{-1}(\cdot)$ is the inverse of $\Lambda_{ij}(\cdot)$.

Step 4. Repeat Steps 2 and 3 until $t_{i,m_i+1} \equiv t_{i,j+1} > Y_i$.

The sequence $\{t_{ij}, j = 1, \ldots, m_i\}$ is then the observed recurrent event times for the $i$th subject. In this algorithm, when $\zeta_{ij} \equiv 1$ and $G$ is unit exponential, the process reduces to the conditional Poisson process given $Z_i$; otherwise, the process is not Poisson conditional on $Z_i$. In the simulation, we set $G$ to be uniform over $(0, 2)$. We used the same configurations used in the manuscript to generate $X_i$, $Z_i$, and the censoring time. The average number of the observed recurrent events in this simulation ranges from 2.6 to 5.1.
Table 5 summarizes the results based on 1000 replications. For all scenarios, our estimator yields small bias, with estimated standard errors reasonably close to the empirical standard errors. The empirical coverage probabilities are somewhat lower than the nominal level under the accelerated rate model, with $\alpha = (-1, -1)^\top, \beta = (0, 0)^\top$ when $n = 200$. As $n$ increases to 400, however, the empirical coverage probabilities become closer to the anticipated level of 95%. Overall, our estimator remains satisfactory under the non-Poisson settings.

S3. Assessment of the identifiability assumption for the infection datasets

Under the Weibull model, $\log\{\Lambda_0(t)\}$ is linear in $\log(t)$. This motivates a check for model identifiability by assessing whether $\log\{\hat{\Lambda}_0(t)\}$ and $\log(t)$ form a linear pattern. Under the Weibull model, Model (1) reduces to the special case of the Cox-type model considered by Wang et al. (2001), and $\hat{\Lambda}_0(t)$ can be obtained using their conditional likelihood method; Figure 2 shows $\log\{\hat{\Lambda}_0(t)\}$ versus $\log(t)$ for the kidney transplant cohort and the HSCT cohort. In both cases, the linearity does not seem to be appropriate. A formal test would be of future interest.
References


Table 1: Simulation results with $\alpha = (-1, -1)^\top$ and $\beta = (0, 0)^\top$. Columns without an asterisk (*) present results using zero vector as initial value; Columns with an asterisk present results using true value as initial value; Bias is the empirical bias; ESE is the empirical standard error; ASE is the average standard error; CP is the empirical coverage probability (%) of 95% confidence intervals.

<table>
<thead>
<tr>
<th>n</th>
<th>Proposed</th>
<th>Sun and Su (2008)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias ESE ASE CP</td>
<td>Bias ESE ASE CP Bias* ESE*</td>
</tr>
<tr>
<td>200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-0.024 0.369 0.350 92.5</td>
<td>-0.018 0.375 0.754 0.388 0.127 11.3</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>-0.025 0.336 0.310 93.4</td>
<td>0.030 0.335 0.736 0.364 0.136 12.3</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.013 0.236 0.229 94.1</td>
<td>0.013 0.244 0.470 0.260 0.110 19.5</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-0.017 0.217 0.208 93.7</td>
<td>0.023 0.215 0.468 0.242 0.109 17.3</td>
</tr>
<tr>
<td>400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-0.007 0.257 0.249 93.6</td>
<td>0.012 0.232 0.688 0.460 0.172 24.4</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.001 0.222 0.222 94.3</td>
<td>-0.003 0.216 0.619 0.417 0.160 23.4</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-0.007 0.170 0.158 94.5</td>
<td>0.012 0.156 0.429 0.293 0.123 28.3</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-0.004 0.148 0.147 94.1</td>
<td>-0.002 0.148 0.395 0.273 0.118 26.5</td>
</tr>
</tbody>
</table>

$Z = 1$ $\sim$ Gamma(4,4)
Table 2: Simulation results with $\alpha = \beta = (-1, 1)^\top$. Columns without an asterisk (*) present results using zero vector as initial value; Columns with an asterisk present results using true value as initial value; Bias is the empirical bias; ESE is the empirical standard error; ASE is the average standard error; CP is the empirical coverage probability (%) of 95% confidence intervals.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\text{Bias}$</td>
<td>$\text{ESE}$</td>
<td>$\text{ASE}$</td>
</tr>
<tr>
<td>200</td>
<td>$\alpha_1$</td>
<td>-0.016</td>
<td>0.333</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>-0.008</td>
<td>0.316</td>
</tr>
<tr>
<td></td>
<td>$\beta_1$</td>
<td>-0.001</td>
<td>0.230</td>
</tr>
<tr>
<td></td>
<td>$\beta_2$</td>
<td>-0.004</td>
<td>0.215</td>
</tr>
<tr>
<td>400</td>
<td>$\alpha_1$</td>
<td>0.003</td>
<td>0.244</td>
</tr>
<tr>
<td></td>
<td>$\alpha_2$</td>
<td>-0.006</td>
<td>0.220</td>
</tr>
<tr>
<td></td>
<td>$\beta_1$</td>
<td>-0.002</td>
<td>0.165</td>
</tr>
<tr>
<td></td>
<td>$\beta_2$</td>
<td>0.001</td>
<td>0.156</td>
</tr>
</tbody>
</table>

$Z \sim \text{Gamma}(4, 4)$
Table 3: Additional simulation results with different variance for the subject-specified frailty variable. Bias is the empirical bias; ESE is the empirical standard error; ASE is the average standard error; CP is the empirical coverage probability (%) of 95% confidence intervals.

<table>
<thead>
<tr>
<th>n</th>
<th>$Z \sim \text{Gamma}(1, 1)$</th>
<th></th>
<th></th>
<th>$Z \sim \text{Gamma}(2, 2)$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha = (0, 0)^T, \beta = (-1, 1)^T$</td>
<td></td>
<td></td>
<td>$\alpha = (0, 0)^T, \beta = (-1, 1)^T$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$ &amp; -0.003 &amp; 0.177 &amp; 0.162 &amp; 91.4 &amp; -0.019 &amp; 0.173 &amp; 0.160 &amp; 92.0 &lt;br&gt; $\alpha_2$ &amp; 0.003 &amp; 0.163 &amp; 0.158 &amp; 93.0 &amp; 0.013 &amp; 0.158 &amp; 0.143 &amp; 91.6 &lt;br&gt; $\beta_1$ &amp; 0.020 &amp; 0.214 &amp; 0.198 &amp; 91.8 &amp; -0.005 &amp; 0.166 &amp; 0.159 &amp; 92.8 &lt;br&gt; $\beta_2$ &amp; 0.027 &amp; 0.201 &amp; 0.189 &amp; 92.6 &amp; 0.016 &amp; 0.161 &amp; 0.156 &amp; 93.6</td>
<td>$\alpha_1$ &amp; -0.014 &amp; 0.111 &amp; 0.107 &amp; 92.8 &amp; -0.014 &amp; 0.114 &amp; 0.103 &amp; 92.6 &lt;br&gt; $\alpha_2$ &amp; 0.007 &amp; 0.106 &amp; 0.102 &amp; 94.6 &amp; 0.001 &amp; 0.102 &amp; 0.099 &amp; 93.6 &lt;br&gt; $\beta_1$ &amp; 0.014 &amp; 0.155 &amp; 0.147 &amp; 92.2 &amp; -0.003 &amp; 0.126 &amp; 0.119 &amp; 94.7 &lt;br&gt; $\beta_2$ &amp; 0.022 &amp; 0.144 &amp; 0.136 &amp; 93.4 &amp; 0.006 &amp; 0.115 &amp; 0.111 &amp; 94.6</td>
<td>$\alpha_1$ &amp; -0.021 &amp; 0.397 &amp; 0.375 &amp; 91.1 &amp; -0.028 &amp; 0.374 &amp; 0.355 &amp; 91.4 &lt;br&gt; $\alpha_2$ &amp; 0.011 &amp; 0.355 &amp; 0.334 &amp; 91.8 &amp; 0.008 &amp; 0.322 &amp; 0.316 &amp; 95.0 &lt;br&gt; $\beta_1$ &amp; -0.005 &amp; 0.283 &amp; 0.267 &amp; 92.0 &amp; 0.007 &amp; 0.257 &amp; 0.247 &amp; 94.8 &lt;br&gt; $\beta_2$ &amp; -0.003 &amp; 0.258 &amp; 0.252 &amp; 92.8 &amp; 0.005 &amp; 0.236 &amp; 0.229 &amp; 93.8</td>
<td>$\alpha_1$ &amp; -0.003 &amp; 0.265 &amp; 0.252 &amp; 95.4 &amp; -0.016 &amp; 0.245 &amp; 0.245 &amp; 94.4 &lt;br&gt; $\alpha_2$ &amp; 0.011 &amp; 0.236 &amp; 0.228 &amp; 94.0 &amp; 0.010 &amp; 0.222 &amp; 0.225 &amp; 94.6 &lt;br&gt; $\beta_1$ &amp; -0.013 &amp; 0.202 &amp; 0.191 &amp; 93.0 &amp; -0.006 &amp; 0.181 &amp; 0.177 &amp; 94.0 &lt;br&gt; $\beta_2$ &amp; 0.010 &amp; 0.182 &amp; 0.177 &amp; 93.6 &amp; 0.007 &amp; 0.164 &amp; 0.161 &amp; 95.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Additional simulation results with small average number of observed recurrent event. Bias is the empirical bias; ESE is the empirical standard error; ASE is the average standard error; CP is the empirical coverage probability (%) of 95% confidence intervals.

<table>
<thead>
<tr>
<th>n</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>-0.065</td>
<td>0.060</td>
<td>-0.008</td>
<td>-0.002</td>
<td>-0.080</td>
<td>0.060</td>
<td>-0.013</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>0.635</td>
<td>0.575</td>
<td>0.382</td>
<td>0.317</td>
<td>0.691</td>
<td>0.591</td>
<td>0.472</td>
<td>0.399</td>
</tr>
<tr>
<td></td>
<td>0.626</td>
<td>0.535</td>
<td>0.404</td>
<td>0.337</td>
<td>0.679</td>
<td>0.581</td>
<td>0.514</td>
<td>0.397</td>
</tr>
<tr>
<td></td>
<td>89.0</td>
<td>90.9</td>
<td>93.1</td>
<td>93.4</td>
<td>87.3</td>
<td>88.9</td>
<td>92.6</td>
<td>94.4</td>
</tr>
<tr>
<td>400</td>
<td>-0.050</td>
<td>0.060</td>
<td>-0.003</td>
<td>0.007</td>
<td>-0.043</td>
<td>0.034</td>
<td>0.021</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>0.460</td>
<td>0.398</td>
<td>0.257</td>
<td>0.207</td>
<td>0.484</td>
<td>0.402</td>
<td>0.334</td>
<td>0.242</td>
</tr>
<tr>
<td></td>
<td>0.440</td>
<td>0.388</td>
<td>0.266</td>
<td>0.218</td>
<td>0.441</td>
<td>0.392</td>
<td>0.308</td>
<td>0.251</td>
</tr>
<tr>
<td></td>
<td>91.7</td>
<td>92.0</td>
<td>95.1</td>
<td>95.2</td>
<td>90.0</td>
<td>92.3</td>
<td>94.4</td>
<td>96.0</td>
</tr>
</tbody>
</table>
Table 5: Additional simulation results with recurrent events generated from a non-Poisson process. Bias is the empirical bias; ESE is the empirical standard error; ASE is the average standard error; CP is the empirical coverage probability (%) of 95% confidence intervals.

\[
Z = 1, \quad Z \sim \text{Gamma}(4, 4)
\]

<table>
<thead>
<tr>
<th>n</th>
<th>(\alpha_1)</th>
<th>(\alpha_2)</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
<th>(\alpha_1)</th>
<th>(\alpha_2)</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>0.018</td>
<td>0.038</td>
<td>0.019</td>
<td>0.028</td>
<td>0.015</td>
<td>0.031</td>
<td>0.011</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>0.447</td>
<td>0.380</td>
<td>0.162</td>
<td>0.142</td>
<td>0.323</td>
<td>0.255</td>
<td>0.106</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td>0.442</td>
<td>0.369</td>
<td>0.176</td>
<td>0.156</td>
<td>0.321</td>
<td>0.245</td>
<td>0.121</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>90.8</td>
<td>91.0</td>
<td>95.1</td>
<td>95.3</td>
<td>91.6</td>
<td>92.5</td>
<td>95.8</td>
<td>95.1</td>
</tr>
<tr>
<td>400</td>
<td>-0.003</td>
<td>-0.009</td>
<td>-0.016</td>
<td>-0.020</td>
<td>-0.009</td>
<td>-0.031</td>
<td>-0.017</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td>0.251</td>
<td>0.240</td>
<td>0.130</td>
<td>0.122</td>
<td>0.179</td>
<td>0.164</td>
<td>0.091</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>0.253</td>
<td>0.239</td>
<td>0.125</td>
<td>0.120</td>
<td>0.177</td>
<td>0.165</td>
<td>0.088</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>95.1</td>
<td>95.2</td>
<td>93.8</td>
<td>93.5</td>
<td>95.5</td>
<td>95.4</td>
<td>94.6</td>
<td>94.9</td>
</tr>
</tbody>
</table>

\[
\alpha = (-1, -1)^T, \beta = (1, 1)^T
\]

<table>
<thead>
<tr>
<th>n</th>
<th>(\alpha_1)</th>
<th>(\alpha_2)</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
<th>(\alpha_1)</th>
<th>(\alpha_2)</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>0.067</td>
<td>0.061</td>
<td>-0.026</td>
<td>-0.001</td>
<td>0.050</td>
<td>0.038</td>
<td>-0.022</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>0.431</td>
<td>0.350</td>
<td>0.220</td>
<td>0.184</td>
<td>0.312</td>
<td>0.268</td>
<td>0.177</td>
<td>0.133</td>
</tr>
<tr>
<td></td>
<td>0.403</td>
<td>0.335</td>
<td>0.254</td>
<td>0.205</td>
<td>0.298</td>
<td>0.265</td>
<td>0.177</td>
<td>0.138</td>
</tr>
<tr>
<td></td>
<td>88.8</td>
<td>89.7</td>
<td>96.0</td>
<td>95.3</td>
<td>91.1</td>
<td>91.7</td>
<td>95.2</td>
<td>95.9</td>
</tr>
<tr>
<td>400</td>
<td>-0.008</td>
<td>-0.013</td>
<td>-0.026</td>
<td>-0.001</td>
<td>0.050</td>
<td>0.038</td>
<td>-0.022</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>0.318</td>
<td>0.285</td>
<td>0.163</td>
<td>0.136</td>
<td>0.312</td>
<td>0.268</td>
<td>0.177</td>
<td>0.133</td>
</tr>
<tr>
<td></td>
<td>0.320</td>
<td>0.283</td>
<td>0.162</td>
<td>0.138</td>
<td>0.298</td>
<td>0.265</td>
<td>0.177</td>
<td>0.138</td>
</tr>
<tr>
<td></td>
<td>92.8</td>
<td>92.5</td>
<td>93.1</td>
<td>95.1</td>
<td>91.1</td>
<td>91.7</td>
<td>95.2</td>
<td>95.9</td>
</tr>
</tbody>
</table>

\[
\alpha = (0, 0)^T, \beta = (-1, -1)^T
\]

<table>
<thead>
<tr>
<th>n</th>
<th>(\alpha_1)</th>
<th>(\alpha_2)</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
<th>(\alpha_1)</th>
<th>(\alpha_2)</th>
<th>(\beta_1)</th>
<th>(\beta_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>0.008</td>
<td>-0.010</td>
<td>-0.001</td>
<td>-0.012</td>
<td>0.005</td>
<td>-0.016</td>
<td>0.001</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>0.318</td>
<td>0.163</td>
<td>0.136</td>
<td>0.191</td>
<td>0.191</td>
<td>0.106</td>
<td>0.094</td>
<td>0.099</td>
</tr>
<tr>
<td></td>
<td>0.320</td>
<td>0.162</td>
<td>0.138</td>
<td>0.205</td>
<td>0.205</td>
<td>0.112</td>
<td>0.099</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>92.8</td>
<td>93.1</td>
<td>95.1</td>
<td>95.5</td>
<td>91.1</td>
<td>91.7</td>
<td>95.2</td>
<td>95.9</td>
</tr>
<tr>
<td>400</td>
<td>0.001</td>
<td>0.012</td>
<td>0.005</td>
<td>0.012</td>
<td>0.005</td>
<td>0.106</td>
<td>0.094</td>
<td>0.107</td>
</tr>
<tr>
<td></td>
<td>0.327</td>
<td>0.172</td>
<td>0.146</td>
<td>0.203</td>
<td>0.203</td>
<td>0.112</td>
<td>0.107</td>
<td>0.108</td>
</tr>
<tr>
<td></td>
<td>0.329</td>
<td>0.178</td>
<td>0.153</td>
<td>0.213</td>
<td>0.213</td>
<td>0.122</td>
<td>0.107</td>
<td>0.108</td>
</tr>
<tr>
<td></td>
<td>93.6</td>
<td>95.8</td>
<td>96.2</td>
<td>95.2</td>
<td>95.2</td>
<td>94.9</td>
<td>95.6</td>
<td>95.6</td>
</tr>
</tbody>
</table>
Figure 1: Plots of estimated $\hat{\Lambda}_0(t)$ with the empirical pointwise 95% confidence intervals for $n = 200$. Solid lines (—) present the empirical average of the estimated $\hat{\Lambda}_0(t)$; Dashed lines (- - -) present the true curve, $\Lambda_0(t) = 0.5 \log(1 + t)$; Dotted lines (.....) present the empirical pointwise confidence intervals obtained from the 2.5th and the 97.5th quantile of the estimated $\hat{\Lambda}_0(t)$. 

$Z = \sim \text{Gamma}(4, 4)$
Figure 2: Plot of the estimated log baseline rate function versus log event time.

(a) Kidney transplant cohort.  (b) HSCT cohort.