Supplementary Materials for

Semiparametric Transformation Models With Random Effects for Joint Analysis of Recurrent and Terminal Events

Donglin Zeng* **and D. Y. Lin**

Web Appendix A

EM Algorithm

For simplicity, we consider the common setting of time-independent covariates. In the E-step, we calculate the conditional expectation of $g(b_i)$ given the observed data for some function $g(\cdot)$. We denote such conditional expectation by $\widehat{E}[g(b_i)]$. In view of the expression of the likelihood function, $\widehat{E}[g(b_i)]$ is equal to $\int_b g(b)\Gamma_i(b)f(b;\gamma)db/\int_b \Gamma_i(b)f(b;\gamma)db$, where

$$\Gamma_i(b) = \exp\left\{b^T(\sum_t R_i(t)\Delta N_i^*(t)\widetilde{Z}_i + \Delta_i(\phi\circ\widetilde{Z}_i))\right\}$$
$$\times \exp\left\{\int \log H'(e^{\alpha^T Z_i + b^T\widetilde{Z}_i}A(t))R_i(t)dN_i^*(t) + \Delta_i \log G'(e^{\beta^T Z_i + b^T(\phi\circ\widetilde{Z}_i)}\Lambda(Y_i))\right\}$$
$$\times \exp\left\{-H(e^{\alpha^T Z_i + b^T\widetilde{Z}_i}A(Y_i)) - G(e^{\beta^T Z_i + b^T(\phi\circ\widetilde{Z}_i)}\Lambda(Y_i))\right\}.$$

The integrations in both the numerator and the denominator are evaluated through numerical approximations, such as the Gaussian-quadrature approximation when b_i is normal.

In the M-step, we maximize the following function

$$\sum_{i=1}^n \left[\int R_i(t)\left\{\log A\{t\} + \alpha^T Z_i + \widehat{E}[b_i]^T\widetilde{Z}_i + \widehat{E}[\log H'(e^{\alpha^T Z_i + b_i^T\widetilde{Z}_i}A(Y_i))]\right\}dN_i^*(t)\right.$$
$$\left. -\widehat{E}[H(e^{\alpha^T Z_i + b_i^T Z_i}A(Y_i))]\right]$$
$$+\sum_{i=1}^n \left[\Delta_i\left\{\log\Lambda\{Y_i\} + \beta^T Z_i + \widehat{E}[b_i]^T(\phi\circ\widetilde{Z}_i) + \widehat{E}[\log G'(e^{\beta^T Z_i + b_i^T(\phi\circ\widetilde{Z}_i)}\Lambda(Y_i))]\right\}\right.$$
$$\left. -\widehat{E}[G(e^{\beta^T Z_i + b_i^T(\phi\circ\widetilde{Z}_i)}\Lambda(Y_i))]\right] + \sum_{i=1}^n \widehat{E}[\log f(b_i;\gamma)]. \qquad (A1)$$

Let $t_{11} < t_{21} < t_{31} < \ldots < t_{n_1,1}$ be the ordered time points where recurrent events are observed, and let $a_1, a_2, a_3, \ldots, a_{n_1}$ be the jump sizes of A at those time points. Similarly, let

$t_{12} < t_{22} < t_{32} < \ldots < t_{n_2,2}$ be the ordered terminal event times, and let $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_{n_2}$ be the jump sizes of Λ at those time points. By differentiating (A1) with respect to a_k and setting the derivative to be zero, we obtain

$$\frac{1}{a_k} + \sum_{i=1}^{n} \widehat{E}\left\{\int R_i(t) I(t \geq t_{k1}) \frac{e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} H''(e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} A(t))}{H'(e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} A(t))} dN_i^*(t) \right.$$
$$\left. - I(Y_i \geq t_{k1}) e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} H'(e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} A(Y_i))\right\} = 0,$$

where $H''(x) = dH^2(x)/dx^2$. Thus,

$$\frac{1}{a_{k+1}} = \frac{1}{a_k} + \sum_{i=1}^{n} \left\{\int R_i(t) I(t_{k1} \leq t < t_{k+1,1}) \frac{e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} H''(e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} A(t))}{H'(e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} A(t))} dN_i^*(t) \right.$$
$$\left. - I(t_{k1} \leq Y_i < t_{k+1,1}) e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} H'(e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} A(Y_i))\right\}. \quad (A2)$$

Since $A(t) = a_1 + \ldots + a_k$ for t between t_{k1} and $t_{k+1,1}$, equation (A2) provides a forward recursive formula for calculating a_{k+1} from $a_1, \ldots, a_k$. To obtain a backward recursive formula, we let $\theta_1 = A(t_{n_1,1})$ and redefine $f_{k1} = a_k/\theta_1$. Then $\sum_k f_{k1} = 1$, and the above equation becomes

$$\frac{1}{f_{k1}} = \frac{1}{f_{k+1,1}} - \frac{1}{\alpha_1} \sum_{i=1}^{n} \widehat{E}\left\{\int R_i(t) I(t_{k1} \leq t < t_{k+1,1}) \frac{e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} H''(e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} A(t))}{H'(e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} A(t))} dN_i^*(t) \right.$$
$$\left. - I(t_{k1} \leq Y_i < t_{k+1,1}) e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} H'(e^{\alpha^T Z_i + b_i^T \widetilde{Z}_i} A(Y_i))\right\}. \quad (A3)$$

Since $A(t) = \theta_1(1 - \sum_{j=k+1}^{n_1} f_{j1})$ for t between t_{k1} and $t_{k+1,1}$, we can calculate a_k from $a_{k+1}, \ldots, a_{n_1}$ and θ_1. Likewise, we can obtain both forward and backward recursive formulae for estimating $\lambda_1, \lambda_2, \ldots$. Specifically,

$$\frac{1}{\lambda_{k+1}} = \frac{1}{\lambda_k} + \sum_{i=1}^{n} I(t_{k2} \leq Y_i < t_{k+1,2}) \widehat{E}\left\{\Delta_i \frac{e^{\beta^T Z_i + b_i^T(\phi \circ \widetilde{Z}_i)} G''(e^{\beta^T Z_i + b_i^T(\phi \circ \widetilde{Z}_i)} \Lambda(Y_i))}{G'(e^{\beta^T Z_i + b_i^T(\phi \circ \widetilde{Z}_i)} \Lambda(Y_i))} \right.$$
$$\left. - e^{\beta^T Z_i + b_i^T(\phi \circ \widetilde{Z}_i)} G'(e^{\beta^T Z_i + b_i^T(\phi \circ \widetilde{Z}_i)} \Lambda(Y_i))\right\}, \quad (A4)$$

and

$$\frac{1}{f_{k2}} = \frac{1}{f_{k+1,2}} - \frac{1}{\alpha_2} \sum_{i=1}^{n} I(t_{k2} \leq Y_i < t_{k+1,2}) \widehat{E}\left\{\Delta_i \frac{e^{\beta^T Z_i + b_i^T(\phi \circ \widetilde{Z}_i)} G''(e^{\beta^T Z_i + b_i^T(\phi \circ \widetilde{Z}_i)} \Lambda(Y_i))}{G'(e^{\beta^T Z_i + b_i^T(\phi \circ \widetilde{Z}_i)} \Lambda(Y_i))} \right.$$
$$\left. - e^{\beta^T Z_i + b_i^T(\phi \circ \widetilde{Z}_i)} G'(e^{\beta^T Z_i + b_i^T(\phi \circ \widetilde{Z}_i)} \Lambda(Y_i))\right\} \quad (A5)$$

where $f_{k2} = \lambda_k/\theta_2$, $\theta_2 = \Lambda(t_{n_2,2})$, and $G''(x) = d^2G(x)/dx^2$.

Using the forward recursive formulae (A2) and (A4), the maximization of (A1) over α, β, ϕ, γ and all the jump sizes of A and Λ becomes the maximization over α, β, ϕ, γ, a_1 and λ_1. Likewise, using the backward recursive formulae (A3) and (A5), the maximization is over α, β, ϕ, γ, θ_1, θ_2, f_{n_1} and f_{n_2}. In either way, we reduce the large number of parameters to only a small number in the maximization of (A1). The Newton-Raphson algorithm can be used to update the estimates in the M-step. We iterate between the E-step and the M-step until convergence to obtain the NPMLEs. The EM algorithm also works for time-dependent covariates, although the above recursive formulae are no longer applicable.

Web Appendix B

Asymptotic Properties

Let α_0, β_0, ϕ_0, γ_0, $A_0(t)$ and $\Lambda_0(t)$ denote the true parameter values. We wish to show that the NPMLEs $\widehat{\alpha}$, $\widehat{\beta}$, $\widehat{\phi}$, $\widehat{\gamma}$, $\widehat{A}(\cdot)$ and $\widehat{\Lambda}(\cdot)$ are consistent and

$$n^{1/2}\begin{bmatrix} \widehat{\alpha}-\alpha_0 \\ \widehat{\beta}-\beta_0 \\ \widehat{\phi}-\phi_0 \\ \widehat{\gamma}-\gamma_0 \\ \widehat{A}(\cdot)-A_0(\cdot) \\ \widehat{\Lambda}(\cdot)-\Lambda_0(\cdot) \end{bmatrix}$$

converges weakly to a zero-mean Gaussian process in $R^d \times BV[0,\tau] \times BV[0,\tau]$, where d is the dimension of $(\alpha^T, \beta^T, \phi^T, \gamma^T)$, and $BV[0,\tau]$ denotes the space of all functions with bounded variations in $[0,\tau]$. We also wish to show that $(\widehat{\alpha}^T, \widehat{\beta}^T, \widehat{\phi}^T, \widehat{\gamma}^T)$ is asymptotically efficient and that the inverse observed information matrix is a consistent estimator of the limiting covariance matrix.

We impose the following conditions.

(D1) The parameter value $(\alpha_0^T, \beta_0^T, \phi_0^T, \gamma_0^T)^T$ belongs to the interior of a compact set Θ in R^d, and $A_0'(t) > 0$ and $\Lambda_0'(t) > 0$ for all $t \in [0,\tau]$.

(D2) With probability one, $Z_i(\cdot)$ is left-continuous with uniformly bounded left- and right-derivatives in $[0,\tau]$.

(D3) With probability one, $P(C_i \geq \tau | Z_i) > \delta_0 > 0$ for some constant δ_0.

(D4) With probability one, $E[N_i^*(\tau)] < \infty$.

(D5) The transformation functions $H(\cdot)$ and $G(\cdot)$ are four-times differentiable with $H(0) = G(0) = 0$, $H'(0) > 0$ and $G'(0) > 0$. In addition, there exist constants $\mu_0 > 0$ and $\kappa_0 > 0$ such that for any integer $m \geq 0$ and any sequence $0 < x_1 < \ldots < x_m \leq y$,

$$\prod_{l=1}^{m} \{(1 + x_l) H'(x_l)\} \exp\{-H(y)\} \leq \mu_0^m (1 + y)^{-\kappa_0}$$

and

$$(1 + x) G'(x) \exp\{-G(x)\} \leq \mu_0 (1 + x)^{-\kappa_0}.$$

There also exists a constant ρ_0 such that

$$\sup_x \left\{ \frac{|H''(x)| + |H^{(3)}(x)| + |H^{(4)}(x)|}{H'(x)(1 + x)^{\rho_0}} \right\} + \sup_x \left\{ \frac{|G''(x)| + |G^{(3)}(x)| + |G^{(4)}(x)|}{G'(x)(1 + x)^{\rho_0}} \right\} < \infty,$$

where $H^{(3)}$, $G^{(3)}$, $H^{(4)}$ and $G^{(4)}$ denote the third and fourth derivatives.

(D6) For any constant $c_0 > 0$,

$$\sup_\gamma E\left[\int_b \exp\{c_0 (N_i^*(\tau) + 1)|b|\} f(b; \gamma) db \right] < \infty,$$

and

$$\left| \frac{\partial f(b;\gamma)/\partial\gamma}{f(b;\gamma)} \right| + \left| \frac{\partial^2 f(b;\gamma)/\partial\gamma^2}{f(b;\gamma)} \right| + \left| \frac{\partial^3 f(b;\gamma)/\partial\gamma^3}{f(b;\gamma)} \right| \leq \exp\{c_0(1 + |b|)\},$$

(D7) If there exist $c(t)$ and v such that $c(t) + v^T Z(t) = 0$ with probability 1, then $c(t) = 0$ and $v = 0$. In addition, there exists some $t \in [0, \tau]$ such that $\{\widetilde{Z}(t)\}$ spans the whole space of b.

(D8) $f(b;\gamma) = f(b;\gamma_0)$ almost surely if and only if $\gamma = \gamma_0$. In addition, if $v^T f'(b;\gamma_0) = 0$ almost surely, then $v = 0$.

It can be easily verified that condition (D5) holds for all commonly used transformations, including the classes of Box-Cox and logarithmic transformations and that (D6) and (D8) hold for normal and many other distributions. Condition (D7) entails the linear independence of covariates, which is a natural requirement in regression analysis.

The desired asymptotic properties will hold if we can verify that conditions (C1)-(C8) in Appendix B of Zeng and Lin (2007) hold under our conditions (D1)-(D8). It follows from the arguments in Appendix B of Zeng and Lin (2007) that our conditions (D1)-(D8) imply (C1)-(C4), (C6) and (C8) of Zeng and Lin (2007). It remains to verify the two identifiability conditions (C5) and (C7) of Zeng and Lin (2007).

We first verify (C5). Suppose that $(\alpha, \beta, \phi, \gamma, A, \Lambda)$ yields the same likelihood as $(\alpha_0, \beta_0, \phi_0, \gamma_0, A_0, \Lambda_0)$. That is,

$$\int_b \left[\prod_t \left\{ a(t) e^{\alpha^T Z_i(t) + b^T \widetilde{Z}_i(t)} H'(q_1(t)) \right\}^{R_i(t)\Delta N_i^*(t)} e^{-H(q_1(Y_i))} \right]$$
$$\times \left[\left\{ \lambda(Y_i) e^{\beta^T Z_i(Y_i) + b^T(\phi \circ \widetilde{Z}_i(Y_i))} G'(q_2(Y_i)) \right\}^{\Delta_i} e^{-G(q_2(Y_i))} \right] f(b;\gamma) db$$
$$= \int_b \left[\prod_t \left\{ a_0(t) e^{\alpha_0^T Z_i(t) + b^T \widetilde{Z}_i(t)} H'(q_{01}(t)) \right\}^{R_i(t)\Delta N_i^*(t)} e^{-H(q_{01}(Y_i))} \right]$$
$$\times \left[\left\{ \lambda_0(Y_i) e^{\beta_0^T Z_i(Y_i) + b^T(\phi_0 \circ \widetilde{Z}_i(Y_i))} G'(q_{02}(Y_i)) \right\}^{\Delta_i} e^{-G(q_{02}(Y_i))} \right] f(b;\gamma_0) db, \qquad (A6)$$

where $q_1(t) = \int_0^t e^{\alpha^T Z(s) + b^T \widetilde{Z}(s)} dA(s)$, $q_2(t) = \int_0^t e^{\beta^T Z(s) + b^T(\phi \circ \widetilde{Z}(s))} d\Lambda(s)$, and q_{01} and q_{02} are q_1 and q_2 evaluated at the true parameter values. We take the following actions on both sides of (A6):

(i) For the terminal event, we obtain an equation from (A6) by setting $\Delta = 0$ and $Y = \tau$; we obtain a second equation by integrating t from t_2 to τ on both sides under $\Delta = 1$ and $Y = t$. We then take the difference between these two equations.

(ii) For the recurrent events, we let $R(t) = 1$ and let $N^*(t)$ have jumps at $s_1, s_2, \ldots, s_m$ and $s'_1, \ldots, s'_{m'}$ for any arbitrary $(m + m')$ times in $[0, \tau]$. We integrate l_1 of $s_1, \ldots, s_m$ from 0 to t_{11}, l_2 of them from 0 to t_{12}, …, l_K of them from 0 to t_{1K}, and integrate $s'_1, \ldots, s'_{m'}$ from 0 to t_2. We then multiply both sides by $\left\{ \prod_k (\mathrm{i}\omega_{1k})^{l_k} / \prod_k l_k! \right\} / m'!$ and sum over $l_1, \ldots, l_K, m' = 0, 1, \ldots$.

After these sequential operations, we obtain

$$\int_b \exp\{\sum_{k=1}^K \mathrm{i}\omega_k H(q_1(t_{1k}))\} \exp\{-G(q_2(t_2))\} f(b;\gamma) db$$

$$= \int_b \exp\{\sum_{k=1}^K \mathrm{i}\omega_k H(q_{01}(t_{1k}))\} \exp\{-G(q_{02}(t_2))\} f(b;\gamma_0) db. \quad (A7)$$

With $t_2 = 0$, this equation implies that $H(q_1(t_{1k}))$ $(k = 1, \ldots, K)$ as a function of $b_1 \sim f(b;\gamma)$ and $H(q_{01}(t_{1k}))$ $(k = 1, \ldots, K)$ as a function of $b_2 \sim f(b;\gamma_0)$ have the same distribution, and so is true for $q_1(t_{1k})$ $(k = 1, \ldots, K)$ and $q_{01}(t_{1k})$ $(k = 1, \ldots, K)$ because of the one-to-one mapping. Thus, $\log q_1'(t_{1k})$ $(k = 1, \ldots, K)$ and $\log q_{01}'(t_{1k})$ $(k = 1, \ldots, K)$ must also have the same distribution. Because $E[b_1] = E[b_2] = 0$, we obtain $\log a(t) + Z(t)^T\alpha = \log a_0(t) + Z(t)^T\alpha_0$. It follows from (D7) that $a(t) = a_0(t)$ and $\alpha = \alpha_0$. In addition, $b_1^T\widetilde{Z}(t_{1k})$ $(k = 1, \ldots, K)$ and $b_2^T\widetilde{Z}(t_{1k})$ $(k = 1, \ldots, K)$ have the same distribution. Since $\widetilde{Z}(t)$ spans the space for b, b_1 and b_2 must have the same distribution. It then follows from (D8) that $\gamma = \gamma_0$. The foregoing arguments also imply that $b_1 \sim \exp\{-G(q_2(t_2))\}f(b;\gamma)$ and $b_2 \sim \exp\{-G(q_{02}(t_2))\}f(b;\gamma_0)$ have the same distribution. Thus, $\exp\{-G(q_2(t_2))\}f(b;\gamma) = \exp\{-G(q_{02}(t_2))\}f(b;\gamma_0)$. That is, $\log\lambda(t) + \beta^T Z(t) + b^T(\phi \circ \widetilde{Z}(t)) = \log\lambda_0(t) + \beta_0^T Z(t) + b^T(\phi_0 \circ \widetilde{Z}(t))$. Taking the expectation with respect to b, we see that $\lambda = \lambda_0$ and $\beta = \beta_0$. Clearly, $\phi = \phi_0$.

To verify (C7) of Zeng and Lin (2007), we write out the score equation along the path $(\alpha_0 + \epsilon v_1, \beta_0 + \epsilon v_2, \gamma_0 + \epsilon v_\gamma, \phi_0 + \epsilon v_3, A_0 + \epsilon \int h_1 dA_0, \Lambda_0 + \epsilon \int h_2 d\Lambda_0)$. We perform operations (i) and (ii) on the score equation to obtain

$$\int_b \left[\sum_{k=1}^K \mathrm{i}\omega_k A_1(t_{1k}, b) - A_2(t_2, b) + \frac{f'(b;\gamma_0)^T v_\gamma}{f(b;\gamma_0)}\right] \exp\left\{\sum_{k=1}^K \mathrm{i}\omega_k H(q_{01}(t_{1k})) - G(q_{02}(t_2))\right\}$$
$$\times f(b;\gamma_0) db = 0, \quad (A8)$$

where $A_1(t,b) = \int_0^t (h_1(s) + v_1^T Z(s) + b^T(v_1 \circ \widetilde{Z}(s))) q_{01}'(s) ds H'(q_{01}(t)) H(q_{01}(t))$, and $A_2(t,b) = \int_0^t (h_2(s) + v_2^T Z(s) + b^T(v_2 \circ \widetilde{Z}(s))) q_{02}'(s) ds G'(q_{02}(t)) G(q_{02}(t))$. Applying the Fourier transformation to both sides yields that for $b \sim f(b;\gamma_0)$,

$$\sum_{k=1}^K \frac{\partial}{\partial g_k} E_b\left[A_1(t_{1k}, b) \middle| H(q_{01}(t_{11})) = g_1, \ldots, H(q_{01}(t_{1K})) = g_K\right] f(g_1, \ldots, g_K)$$
$$-E_b\left[A_2(t_2, b) \middle| H(q_{01}(t_{11})) = g_1, \ldots, H(q_{01}(t_{1K})) = g_K\right] f(g_1, \ldots, g_K)$$
$$+E_b\left[\frac{f'(b;\gamma_0)^T v_\gamma}{f(b;\gamma_0)} \middle| H(q_{01}(t_{11})) = g_1, \ldots, H(q_{01}(t_{1K})) = g_K\right] f(g_1, \ldots, g_K) = 0,$$

where $f(g_1, \ldots, g_K)$ is the joint density of $H(q_{01}(t_{11})), \ldots, G_K(q_{01}(t_{1K}))$. We integrate out $g_1, \ldots, g_K$ and set $t_2 = 0$. Note that the last term is zero and the remainder is a homogeneous equation for $v_1^T Z(t_{1k}) + h_1(t_{1k})$ $(k = 1, \ldots, K)$, which has a trivial solution. By condition (D7), $v_1 = 0$ and $h_1 = 0$. Then (A8) becomes

$$\int_b \left[-A_2(t_2, b) + \frac{f'(b; \gamma_0)^T v_\gamma}{f(b; \gamma_0)} \right] \exp\left\{ \sum_{k=1}^{K} \mathrm{i}\omega_k H(q_{01}(t_{1k})) - G(q_{02}(t_2)) \right\} f(b; \gamma_0) db = 0.$$

By the arguments used in proving the identifiability, we obtain

$$-A_2(t_2, b) + \frac{f'(b; \gamma_0)^T v_\gamma}{f(b; \gamma_0)} = 0.$$

Clearly, $f'(b; \gamma_0)^T v_\gamma = 0$, so it follows from condition (D8) that $v_\gamma = 0$. Thus, $v_2^T Z(t_2) + h_2(t) + b^T(v_3 \circ \widetilde{Z}(t)) = 0$, implying that $v_2 = 0$, $h_2 = 0$ and $v_3 = 0$.