Supplementary Materials

"Empirical Likelihood MLE for Joint Modeling

Right Censored Survival Data with Longitudinal Covariates"

By Jian-Jian Ren and Yuyin Shi

Appendix

Proof of (14): For fixed $t > 0$, consider partition points of interval $[0, t]$: $0 = s_0 < s_1 < \cdots < s_N = t$, which have equal width $\Delta = \frac{t}{N}$ for each subinterval $[s_{j-1}, s_j]$, where N is any large positive integer. Then, from (6.8) on Page 196 in Kalbfleisch and Prentice (2002) we have the following:

$$
\begin{aligned}
\int_0^t f(s \,|\, X^s)\, ds &= o(1) + \sum_{k=0}^{N-1} \Delta\, f(s_k \,|\, X^{s_k}) \\
&\overset{(13)}{=\!=} o(1) + \sum_{k=0}^{N-1} \big[P\{s_k \le T < s_k + \Delta \,|\, X^{s_k}\} + \Delta\, r_k \big] \\
&\overset{\text{(6.8) of K-P}}{=\!=\!=\!=} o(1) + \sum_{k=0}^{N-1} \big[P\{s_k \le T < s_k + \Delta \,|\, X^{t}\} + \Delta\, r_k \big] \\
&= o(1) + P\{0 \le T < t \,|\, X^t\} + \frac{t}{N} \sum_{k=0}^{N-1} r_k \\
&= o(1) + P\{0 \le T < t \,|\, X^t\} + O\big(\max_k |\, r_k \,| \big),
\end{aligned}
\tag{47}
$$

where r_k is the remainder term from (13) at each s_k. Letting $N \to \infty$, we know that (14) follows from (47) and the uniform convergence of (13). $\square$

Proof of Lemma 1: For $\delta = 1$, we know

$$
\begin{aligned}
g(t, 1 \,|\, X^t) &= \lim_{\Delta \to 0} \Delta^{-1} P\{t \le V < t + \Delta,\, \delta = 1 \,|\, X^t\} \\
&= \lim_{\Delta \to 0} \Delta^{-1} P\{t \le T < t + \Delta,\, T \le C \,|\, X^t\}.
\end{aligned}
$$

Thus, the 1st equation of (16) follows from (13) and the following inequalities:

$$
\begin{aligned}
&P\{t \le T < t+\Delta,\, T \le C \,|\, X^t\} \le P\{t \le T < t+\Delta,\, t \le C \,|\, X^t\} \\
&\quad = P\{t \le T < t+\Delta \,|\, X^t\} P\{t \le C \,|\, X^t\} = P\{t \le T < t+\Delta \,|\, X^t\}\, \bar{F}_C(t); \\
&P\{t \le T < t+\Delta,\, T \le C \,|\, X^t\} \ge P\{t \le T < t+\Delta,\, t+\Delta \le C \,|\, X^t\} \\
&\quad = P\{t \le T < t+\Delta \,|\, X^t\} P\{t+\Delta \le C \,|\, X^t\} = P\{t \le T < t+\Delta \,|\, X^t\}\, \bar{F}_C(t+\Delta).
\end{aligned}
$$

For $\delta = 0$, we know

$$
\begin{aligned}
g(t, 0 \,|\, X^t) &= \lim_{\Delta \to 0} \Delta^{-1} P\{t \le V < t+\Delta, \delta = 0 \,|\, X^t\} \\
&= \lim_{\Delta \to 0} \Delta^{-1} P\{t \le C < t+\Delta, T > C \,|\, X^t\}.
\end{aligned}
$$

Thus, the 2nd equation of (16) follows from (12)-(13) and the following inequalities:

$$
\begin{aligned}
&P\{t \le C < t+\Delta,\, T > C \,|\, X^t\} \le P\{t \le C < t+\Delta,\, T \ge t \,|\, X^t\} \\
&\quad = P\{t \le C < t+\Delta \,|\, X^t\} P\{T \ge t \,|\, X^t\} = P\{t \le C < t+\Delta\}\, \bar{F}(t \,|\, X^t); \\
&P\{t \le C < t+\Delta,\, T > C \,|\, X^t\} \ge P\{t \le C < t+\Delta,\, t+\Delta \le T \,|\, X^t\} \\
&\quad = P\{t \le C < t+\Delta \,|\, X^t\} P\{t+\Delta \le T \,|\, X^t\} = P\{t \le C < t+\Delta\}\, \bar{F}(t+\Delta \,|\, X^t). \ \square
\end{aligned}
$$

Proof of Lemma 2: From (18), for any $t \in [t_{iJ_i},\, V_i]$ we have for $t_{i0} = 0$:

$$
\begin{aligned}
-\int_0^t \lambda_0(s)\, e^{\beta_0 \hat{Z}_i(s)}\, ds &= -\int_{t_{iJ_i}}^t \lambda_0(s)\, e^{\beta_0 \hat{Z}_i(s)}\, ds - \sum_{j=1}^{J_i} \int_{t_{i,\,j-1}}^{t_{ij}} \lambda_0(s)\, e^{\beta_0 \hat{Z}_i(s)}\, ds \\
&= -\int_{t_{iJ_i}}^t \lambda_0(s)\, e^{\beta_0 \tilde{Z}_i(t_{iJ_i})}\, ds - \sum_{j=1}^{J_i} \int_{t_{i,\,j-1}}^{t_{ij}} \lambda_0(s)\, e^{\beta_0 \tilde{Z}_i(t_{i,\,j-1})}\, ds \\
&= -e^{\beta_0 \tilde{Z}_i(t_{iJ_i})} \int_{t_{iJ_i}}^t \frac{f_0(s)}{\bar{F}_0(s)}\, ds - \sum_{j=1}^{J_i} e^{\beta_0 \tilde{Z}_i(t_{i,\,j-1})} \int_{t_{i,\,j-1}}^{t_{ij}} \frac{f_0(s)}{\bar{F}_0(s)}\, ds \\
&= e^{\beta_0 \tilde{Z}_i(t_{iJ_i})} \ln\Big(\frac{\bar{F}_0(t)}{\bar{F}_0(t_{iJ_i})}\Big) + \sum_{j=1}^{J_i} e^{\beta_0 \tilde{Z}_i(t_{i,\,j-1})} \ln\Big(\frac{\bar{F}_0(t_{ij})}{\bar{F}_0(t_{i,\,j-1})}\Big).
\end{aligned}
$$

Thus, for $\hat{Z}_i(0) = 0$ and $c_{ij} = e^{\beta_0 \tilde{Z}_i(t_{ij})}$ with $c_{i0} = 1$ and $c_i = c_{iJ_i}$, (19) can be written as

$$\bar{F}(t \,|\, \mathcal{X}_i^t) = \exp\Big\{ c_{iJ_i} \, \ln\Big(\frac{\bar{F}_0(t)}{\bar{F}_0(t_{iJ_i})}\Big) + \sum_{j=1}^{J_i} c_{i,\,j-1} \, \ln\Big(\frac{\bar{F}_0(t_{ij})}{\bar{F}_0(t_{i,\,j-1})}\Big)\Big\}$$

$$= \Big(\frac{\bar{F}_0(t)}{\bar{F}_0(t_{iJ_i})}\Big)^{c_{iJ_i}} \prod_{j=1}^{J_i} \Big(\frac{\bar{F}_0(t_{ij})}{\bar{F}_0(t_{i,\,j-1})}\Big)^{c_{i,\,j-1}}$$

$$= \Big(\frac{\bar{F}_0(t)}{\bar{F}_0(t_{iJ_i})}\Big)^{c_{iJ_i}} \Big(\frac{\bar{F}_0(t_{i1})}{\bar{F}_0(t_{i0})}\Big)^{c_{i0}} \times \Big(\frac{\bar{F}_0(t_{i2})}{\bar{F}_0(t_{i1})}\Big)^{c_{i1}} \times \cdots \times \Big(\frac{\bar{F}_0(t_{iJ_i})}{\bar{F}_0(t_{i,\,J_i-1})}\Big)^{c_{i,\,J_i-1}}$$

$$= \big[\bar{F}_0(t)\big]^{c_{iJ_i}} \prod_{j=1}^{J_i} \big[\bar{F}_0(t_{ij})\big]^{c_{i,\,j-1}-c_{ij}} = \big[\bar{F}_0(t)\big]^{c_i} \prod_{j=1}^{J_i} \big[\bar{F}_0(t_{ij})\big]^{c_{i,\,j-1}-c_{ij}}.$$

Equation (20) is obtained by $f(t \,|\, \mathcal{X}_i^t) = -\frac{d}{dt}\big(\bar{F}(t \,|\, \mathcal{X}_i^t)\big)$. □

Proof of (21): From Lemma 2, we have the following:

$$\prod_{i=1}^{n} \big[f(V_i \,|\, \mathcal{X}_i^{V_i})\big]^{\delta_i} \big[\bar{F}(V_i \,|\, \mathcal{X}_i^{V_i})\big]^{1-\delta_i}$$

$$= \prod_{i=1}^{n} \Big(c_i f_0(V_i) \big[\bar{F}_0(V_i)\big]^{c_i - 1} \prod_{j=1}^{J_i} \big[\bar{F}_0(t_{ij})\big]^{c_{i,\,j-1}-c_{ij}} \Big)^{\delta_i} \Big(\big[\bar{F}_0(V_i)\big]^{c_i} \prod_{j=1}^{J_i} \big[\bar{F}_0(t_{ij})\big]^{c_{i,\,j-1}-c_{ij}} \Big)^{1-\delta_i}$$

$$= \prod_{i=1}^{n} \big[c_i f_0(V_i)\big]^{\delta_i} \big[\bar{F}_0(V_i)\big]^{c_i - \delta_i} \Big(\prod_{j=1}^{J_i} \big[\bar{F}_0(t_{ij})\big]^{c_{i,\,j-1}-c_{ij}} \Big). \qquad \square$$

Proof of Lemma 3: Let $A_{kij} = I\{V_k \leq t_{ij} < V_{k+1}\}$ with $V_{n+1} = \infty$ and let $d_{ij} = c_{i,\,j-1} - c_{ij}$ with c_{ij}'s given by (23). Since for F given by (25), we have $F(t_{ij}) = 0$ for $0 \leq t_{ij} < V_1$, then the last component of (22) can be simplified as:

$$\prod_{i=1}^{n} \prod_{j=1}^{J_i} \big[\bar{F}(t_{ij})\big]^{d_{ij}} = \prod_{i=1}^{n} \prod_{j=1}^{J_i} \prod_{k=1}^{n} \big[\bar{F}(t_{ij})\big]^{A_{kij} d_{ij}} = \prod_{i=1}^{n} \prod_{j=1}^{J_i} \prod_{k=1}^{n} \big[\bar{F}(V_k)\big]^{A_{kij} d_{ij}}$$

$$= \prod_{k=1}^{n} \prod_{i=1}^{n} \prod_{j=1}^{J_i} \big[\bar{F}(V_k)\big]^{A_{kij} d_{ij}} = \prod_{k=1}^{n} \big[\bar{F}(V_k)\big]^{A_k} = \prod_{i=1}^{n} \big[\bar{F}(V_i)\big]^{A_i},$$

where $A_k = \sum_{i=1}^{n} \sum_{j=1}^{J_i} A_{kij} d_{ij}$. Then, (22) can be written as:

$$L(\beta, F) = \prod_{i=1}^{n} \big[c_i \, dF(V_i)\big]^{\delta_i} \big[\bar{F}(V_i)\big]^{d_i - \delta_i} \tag{48}$$

where for $A_i = \sum_{k=1}^{n} \sum_{j=1}^{J_k} A_{ikj} d_{kj}$, we have for c_{ij}'s given in (23),

$$d_i = c_i + A_i = c_i + \sum_{k=1}^{n} \sum_{j=1}^{J_k} A_{ikj} d_{kj} = c_{iJ_i} + \sum_{k=1}^{n} \sum_{j=1}^{J_k} (c_{k,\,j-1} - c_{kj}) I\{V_i \le t_{kj} < V_{i+1}\}.$$

Thus, (26) follows from (25) and (48). $\square$

Proof of Theorem 1: Let $a_i = \frac{p_i}{b_i}$ with $b_i = \sum_{j=i}^{n+1} p_j$, then we have $b_{i+1} = (b_i - p_i)$, $b_1 = 1$, $b_{n+1} = p_{n+1}$ and $(1 - a_i) = \frac{b_{i+1}}{b_i}$. From $\prod_{i=1}^{n} (1 - a_i) = b_{n+1}$ and

$$\prod_{i=1}^{n} a_i^{d_i} (1 - a_i)^{n - h_i} = \Big(\prod_{i=1}^{n} a_i^{d_i}\Big) \prod_{i=1}^{n} \Big(\frac{b_{i+1}}{b_i}\Big)^{n - h_i} = \Big(\prod_{i=1}^{n} a_i^{d_i}\Big) b_{n+1}^{n - h_n} \Big(\prod_{i=1}^{n} b_i^{d_i}\Big) = \Big(\prod_{i=1}^{n} p_i^{d_i}\Big) b_{n+1}^{n - h_n},$$

where $h_i = d_1 + \cdots + d_i$, we obtain the following for (26):

$$\begin{aligned} L(\beta, \boldsymbol{p}) &= \prod_{i=1}^{n} \big[c_i \, p_i\big]^{\delta_i} \big(b_{i+1}\big)^{d_i - \delta_i} = \prod_{i=1}^{n} \big[c_i \, p_i\big]^{\delta_i} \big(b_i - p_i\big)^{d_i - \delta_i} \\ &= \prod_{i=1}^{n} c_i^{\delta_i} p_i^{d_i} \Big(\frac{1 - a_i}{a_i}\Big)^{d_i - \delta_i} = \Big(\prod_{i=1}^{n} c_i^{\delta_i} a_i^{d_i} (1 - a_i)^{h_n - h_i}\Big) \prod_{i=1}^{n} \Big(\frac{1 - a_i}{a_i}\Big)^{d_i - \delta_i} \\ &= \prod_{i=1}^{n} \big(c_i \, a_i\big)^{\delta_i} \big(1 - a_i\big)^{e_i - \delta_i} \equiv L_1(\boldsymbol{a};\, \beta) \end{aligned} \tag{49}$$

where $e_i = d_i + \cdots + d_n$. For fixed β satisfying (AS1), from the 1st and 2nd partial derivatives of $\log L_1(\boldsymbol{a};\, \beta)$ with respect to a_i's, we know that the solution of equations $\frac{\partial \log L_1}{\partial a_i} = 0$, $1 \le i \le n$, is given by $\hat{a}_i = \frac{\delta_i}{e_i}$, $i = 1, \cdots, n$, which maximizes $L_1(\boldsymbol{a};\, \beta)$ with all $0 \le \hat{a}_i \le 1$. Thus, (27) follows from the fact that the d.f. F corresponding to $\hat{a}_i$'s is given by $\bar{\hat{F}}_n(t; \beta) = \prod_{V_i \le t} (1 - \hat{a}_i)$. $\square$

Proof of (28) and (30)-(31): Notice that we obtain (28) by plugging $\hat{a}_i$'s into (49). From notations $c_i' = \frac{dc_i}{d\beta}$ and $e_i' = \frac{de_i}{d\beta}$, we have $\frac{c_i'}{c_i} = \tilde{Z}_i(t_{iJ_i})$ and we obtain equation (30) from the following:

$$\begin{aligned}
\Psi_n(\beta) &= n^{-1}\tfrac{d}{d\beta}\big(\log \ell(\beta)\big) = n^{-1}\tfrac{d}{d\beta}\Big(\sum_{i=1}^{n}\big[\delta_i(\ln c_i - \ln e_i) + (e_i - \delta_i)\big(\ln(e_i - \delta_i) - \ln e_i\big)\big]\Big) \\
&= n^{-1}\sum_{i=1}^{n}\Big[\delta_i\Big(\frac{c_i'}{c_i} - \frac{e_i'}{e_i}\Big) + e_i' \log\Big(1 - \tfrac{\delta_i}{e_i}\Big) + (e_i - \delta_i)\Big(\tfrac{e_i'}{e_i - \delta_i} - \tfrac{e_i'}{e_i}\Big)\Big] \\
&= n^{-1}\sum_{i=1}^{n}\Big[\delta_i\tilde{Z}_i(t_{iJ_i}) + e_i' \log\Big(1 - \tfrac{\delta_i}{e_i}\Big)\Big] = \sum_{i=1}^{n}\delta_i\Big(\tilde{Z}_i(t_{iJ_i}) + e_i' \log\big(1 - \tfrac{1}{e_i}\big)\Big).
\end{aligned}$$

Using Taylor's expansion $\log(1-x) \approx -x$, term $\log\big(1 - \frac{1}{e_i}\big)$ in (30) is approximated by $-\frac{1}{e_i}$, which gives (31). $\square$

Proof of Lemma 4 (a): Using the notations of A_{kij}'s and d_{ij}'s in the proof of Lemma 3, we have the following for e_i in (27) based on d_i's given in (26):

$$\begin{aligned}
e_i = \sum_{q=i}^{n} d_q &= \sum_{q=i}^{n}\Big(c_{qJ_q} + \sum_{k=1}^{n}\sum_{j=1}^{J_k} d_{kj} I\{V_q \le t_{kj} < V_{q+1}\}\Big) \\
&= \sum_{k=i}^{n} c_{kJ_k} + \sum_{q=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{J_k} d_{kj} A_{qkj} I\{q \ge i\} \\
&= \sum_{k=i}^{n} c_{kJ_k} + \sum_{k=1}^{n}\sum_{j=1}^{J_k}\sum_{q=1}^{n} d_{kj} A_{qkj} I\{q \ge i\} \\
&= \sum_{k=i}^{n} c_{kJ_k} + \sum_{k=1}^{n}\sum_{j=1}^{J_k}\sum_{q=i}^{n} d_{kj} I\{V_q \le t_{kj} < V_{q+1}\} \\
&= \sum_{k=i}^{n} c_{kJ_k} + \sum_{k=1}^{n}\sum_{j=1}^{J_k} d_{kj} I\{V_i \le t_{kj}\}.
\end{aligned}$$

Notice that from data (11) and assumption (24), we know that for each $k < i$, we always have $t_{kj} \le V_k < V_i$ for any $j = 1, \cdots, J_k$. Thus, we can simplify and express above equation as below:

$$e_i = \sum_{k=i}^{n} \Big(c_{kJ_k} + \sum_{j=1}^{J_k} d_{kj} I\{V_i \le t_{kj}\} \Big) = \sum_{k=i}^{n} B_{ik}, \tag{50}$$

where $B_{ik} = c_{kJ_k} + \sum_{j=1}^{J_k} d_{kj} I\{V_i \le t_{kj}\}$. Obviously, if $\mathcal{E}_{ik} = \emptyset$, we have $B_{ik} = c_{kJ_k}$. If $\mathcal{E}_{ik} \ne \emptyset$, for J_{ik} given in the statement of Lemma 4 we have

$$B_{ik} = c_{kJ_k} + \sum_{j=J_{ik}+1}^{J_k} d_{kj} = c_{kJ_k} + \sum_{j=J_{ik}+1}^{J_k} \big(c_{k,j-1} - c_{kj}\big) = c_{kJ_{ik}}$$

which gives

$$B_{ik} \overset{(23)}{=\!=} e^{\beta \tilde{Z}_k(t_{kJ_k})} I\{\mathcal{E}_{ik} = \emptyset\} + e^{\beta \tilde{Z}_k(t_{kJ_{ik}})} I\{\mathcal{E}_{ik} \ne \emptyset\}. \qquad \square \tag{51}$$

Proof of Lemma 4 (b): For $e''_i = \frac{de'_i}{d\beta}$, we obtain derivative of $\psi_n(\beta)$ in (31):

$$\psi'_n(\beta) = n^{-1} \sum_{i=1}^{n} \delta_i \Big(\frac{(e'_i)^2 - e_i\, e''_i}{e_i^2} \Big). \tag{52}$$

Let $C_{ik} = \tilde{Z}_k(t_{kJ_k}) I\{\mathcal{E}_{ik} = \emptyset\} + \tilde{Z}_k(t_{kJ_{ik}}) I\{\mathcal{E}_{ik} \ne \emptyset\}$, then from (50)-(51) we have

$$e'_i = \sum_{k=i}^{n} C_{ik} B_{ik} \quad \text{and} \quad e''_i = \sum_{k=i}^{n} C_{ik}^2 B_{ik}.$$

From Cauchy-Schwarz inequality, we know

$$\big(e'_i\big)^2 = \Big(\sum_{k=i}^{n} C_{ik} B_{ik} \Big)^2 = \Big(\sum_{k=i}^{n} \sqrt{B_{ik}}\, \big(C_{ik} \sqrt{B_{ik}}\,\big) \Big)^2 \le \Big(\sum_{k=i}^{n} B_{ik} \Big) \Big(\sum_{k=i}^{n} C_{ik}^2 B_{ik} \Big) = e_i\, e''_i$$

which implies $\psi'_n(\beta) \le 0$ in (52). $\square$

Data-Based Choice of M: In our simulation studies, we only consider the cases with all nonnegative $\tilde{Z}_i(t_{ij})$'s, thus from the proof of Lemma 4 we know that $e_i' \geq 0$ and $E_i' \geq 0$ always. Applying Taylor's expansion to (30) and (37), we have the following:

$$\begin{cases} \Psi_n(\beta) = \psi_n(\beta) - n^{-1} \sum_{i=1}^{n} \dfrac{\delta_i \, e_i'}{2(1-\xi_i)^2 \, e_i^2} \\ \Psi_{M,n}(\beta) = \psi_n(\beta) - n^{-1} \sum_{i=1}^{n} \dfrac{\delta_i \, E_i'}{2(1-\zeta_i)^2 \, E_i^2} \end{cases}$$

where ξ_i is between 0 and $\frac{1}{e_i}$, and ζ_i is between 0 and $\frac{1}{E_i}$, which implies

$$\Psi_n(\beta) < \psi_n(\beta) \quad \text{and} \quad \Psi_{M,n}(\beta) < \psi_n(\beta), \qquad \text{for } \forall \, \beta. \tag{53}$$

For fixed β, we differentiate $\Psi_{M,n}(\beta)$ with respective to M:

$$\frac{\partial \Psi_{M,n}(\beta)}{\partial M} = n^{-1} \sum_{i=1}^{n} \delta_i E_i' \Big(\log\big(1 - \tfrac{1}{E_i}\big) + \frac{\frac{1}{E_i}}{1 - \frac{1}{E_i}} \Big) \equiv n^{-1} \sum_{i=1}^{n} \delta_i \, g\big(\tfrac{1}{E_i}\big) > 0$$

where $g(x) = \log(1-x) + \frac{x}{1-x}$, $0 < x < 1$ satisfies

$$g'(x) = \frac{-1}{1-x} + \frac{1}{(1-x)^2} = \frac{x}{(1-x)^2} > 0 \quad \Rightarrow \quad g(x) > g(0) = 0.$$

Thus, we know that for any fixed β, $\Psi_{M,n}(\beta)$ is increasing in M.

Based on above facts, we have the following steps for data-based choice of M, which starts after finding $\tilde{\beta}_n$ from equation $\psi_n(\beta) = 0$.

Steps to Choose M:

Step 1. Compute $e_1, \cdots, e_n$ at $\beta = (\tilde{\beta}_n - \rho)$, then order them into $e_{(1)} \leq e_{(2)} \leq \cdots \leq e_{(n)}$. If $e_{(1)} \geq 1$, no need to choose M and go to Step 2; if $e_{(1)} < 1$, go to Step 3.

Step 2. Search $\hat{\beta}_n$ from equation $\Psi_n(\beta) = 0$: increase $\rho > 0$ until $\Psi_n(\tilde{\beta}_n - \rho) > 0$, then search $\hat{\beta}_n$ in interval $(\tilde{\beta}_n - \rho,\, \tilde{\beta}_n)$, because (53) implies $\Psi_n(\tilde{\beta}_n) < \psi_n(\tilde{\beta}_n) = 0$.

Step 3. Find $e^* = \min_{1\le i\le n}\{e_{(i)} \,|\, e_{(i)} \ge 10^{-7}\}$ and compute $M_0 = -\log e^*$, then search $\hat{\beta}_n$ from equation $\Psi_{M_0,\,n}(\beta) = 0$: increase $\rho > 0$ until $\Psi_{M_0,\,n}(\tilde{\beta}_n - \rho) > 0$, then search $\hat{\beta}_n$ in interval $(\tilde{\beta}_n - \rho,\, \tilde{\beta}_n)$, because (53) implies $\Psi_{M_0,\,n}(\tilde{\beta}_n) < \psi_n(\tilde{\beta}_n) = 0$.

Step 4. If encounter some $E_i \le 1$ in Step 3, let $\Psi_{M_0,\,n}(\hat{\beta}_n^0) = 0$ by dropping at most two terms involving $E_i \le 1$. Then, use $M = M_0 + \gamma$ to search $\hat{\beta}_n$ from equation $\Psi_{M,\,n}(\beta) = 0$ in interval $(\tilde{\beta}_n^0,\, \tilde{\beta}_n)$, where $\gamma > 0$ and $\rho > 0$ need to be appropriately chosen based on important facts aforementioned regarding $\Psi_{M,\,n}(\beta) = 0$.

Note: Above choice of M_0 is to achieve $E_i \ge 1$ in order to avoid computing floating errors, which is difficult to control in intensive Monte Carlo simulation studies. But such choice of M_0 is conservative and can still have occasional $E_i \le 1$, thus adjustment to use $M = M_0 + \gamma$ is for the purpose of not too big M which can eliminate cases $E_i \le 1$. □

Proof of (44)-(46): For $\tilde{\delta}_k$ and $\tilde{c}_k$ given in (45), we have

$$\begin{aligned}
&\prod_{k=1}^{N}\prod_{V_i=U_k}\big[c_i\, dF(V_i)\big]^{\delta_i}\big[\bar{F}(V_i)\big]^{c_i-\delta_i} \\
&\quad= \prod_{k=1}^{N}\Big(\prod_{V_i=U_k} e^{\beta\tilde{Z}_i(t_{iJ_i})\delta_i}\Big)\big[dF(V_k)\big]^{\tilde{\delta}_k}\big[\bar{F}(V_k)\big]^{\sum_{i=1}^{n}c_iI\{V_i=U_k\}-\tilde{\delta}_k} \\
&\quad= \prod_{k=1}^{N}\big(e^{\beta\sum_{i=1}^{n}\tilde{Z}_i(t_{iJ_i})\delta_iI\{V_i=U_k\}}\big)\big[dF(V_k)\big]^{\tilde{\delta}_k}\big[\bar{F}(V_k)\big]^{\sum_{i=1}^{n}c_iI\{V_i=U_k\}-\tilde{\delta}_k} \\
&\quad= \prod_{k=1}^{N}\big[\tilde{c}_k\, dF(V_k)\big]^{\tilde{\delta}_k}\big[\bar{F}(V_k)\big]^{\sum_{i=1}^{n}c_iI\{V_i=U_k\}-\tilde{\delta}_k}.
\end{aligned} \tag{54}$$

Using the notations of A_{kij}'s and d_{ij}'s in the proof of Lemma 3, for F given by $F(t) = \sum_{k=1}^{N} p_k I\{U_k \le t\}$, we have the following for the last component in the 1st equation of (44):

$$
\begin{aligned}
\prod_{k=1}^{N}\prod_{V_i=U_k}\prod_{j=1}^{J_i}\big[\bar F(t_{ij})\big]^{d_{ij}} &= \prod_{k=1}^{N}\prod_{i=1}^{n}\prod_{j=1}^{J_i}\big[\bar F(t_{ij})\big]^{d_{ij}I\{V_i=U_k\}} \\
&= \prod_{k=1}^{N}\prod_{i=1}^{n}\prod_{j=1}^{J_i}\prod_{q=1}^{N}\big[\bar F(t_{ij})\big]^{A_{qij}d_{ij}I\{V_i=U_k\}} = \prod_{k=1}^{N}\prod_{i=1}^{n}\prod_{j=1}^{J_i}\prod_{q=1}^{N}\big[\bar F(U_q)\big]^{A_{qij}d_{ij}I\{V_i=U_k\}} \\
&= \prod_{q=1}^{N}\prod_{k=1}^{N}\prod_{i=1}^{n}\prod_{j=1}^{J_i}\big[\bar F(U_q)\big]^{A_{qij}d_{ij}I\{V_i=U_k\}} \\
&= \prod_{q=1}^{N}\big[\bar F(U_q)\big]^{\sum_{k=1}^{N}\sum_{i=1}^{n}\sum_{j=1}^{J_i}A_{qij}d_{ij}I\{V_i=U_k\}} = \prod_{k=1}^{N}\big[\bar F(U_k)\big]^{\tilde A_k}
\end{aligned}
\tag{55}
$$

where

$$
\begin{aligned}
\tilde A_k &= \sum_{q=1}^{N}\sum_{i=1}^{n}\sum_{j=1}^{J_i}A_{kij}d_{ij}I\{V_i=U_q\} = \sum_{i=1}^{n}\sum_{j=1}^{J_i}\Big(\sum_{q=1}^{N}I\{V_i=U_q\}\Big)A_{kij}d_{ij} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{J_i}A_{kij}d_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{J_i}(c_{i,j-1}-c_{ij})I\{U_k\le t_{ij}<U_{k+1}\}.
\end{aligned}
$$

Thus, from (54) and for $\tilde d_k$ given in (45) we have

$$
\sum_{i=1}^{n}c_iI\{V_i=U_k\}+\tilde A_k = \sum_{i=1}^{n}\Big(c_{iJ_i}I\{V_i=U_k\}+\sum_{j=1}^{J_i}(c_{i,j-1}-c_{ij})I\{U_k\le t_{ij}<U_{k+1}\}\Big) = \tilde d_k,
$$

which gives (44)-(45) by putting (54) and (55) together.

From the derivation of (49) in the proof of Theorem 1, we know that for any fixed β satisfying $\tilde d_k \ge \tilde\delta_k$, $k=1,\cdots,N$, likelihood function in (44) is maximized by:

$$
1-\hat F_n(t;\,\beta) = \prod_{U_k\le t}\Big(1-\frac{\tilde\delta_k}{\tilde e_k}\Big), \tag{56}
$$

where $\tilde e_k = \tilde d_k+\cdots+\tilde d_N,\ k=1,\cdots,N$; in turn, by similar derivation to that of (28) the profile likelihood function for β_0 is given by:

$$\ell(\beta) = L\big(\beta, \hat{F}_n(\cdot\,;\,\beta)\big) = \prod_{k=1}^{N} \left(\frac{\tilde{c}_k \tilde{\delta}_k}{\tilde{e}_k}\right)^{\tilde{\delta}_k} \left(1 - \frac{\tilde{\delta}_k}{\tilde{e}_k}\right)^{\tilde{e}_k - \tilde{\delta}_k}.$$

From differentiation, we obtain the *profile estimating function*: $\Psi_n(\beta) = n^{-1}\frac{d}{d\beta}\big(\log \ell(\beta)\big)$, which, after computation and algebraic simplification, is given by the 1st equation in (46) with the given $\tilde{Z}_k$.

From Taylor's expansion, we use linear approximation $\log(1-x) \approx -x$ on term $\log\big(1 - \frac{\tilde{\delta}_k}{\tilde{e}_k}\big)$ in $\Psi_n(\beta)$ given by the 1st equation in (46), then we obtain *approximated profile estimating function* $\psi_n(\beta)$, which is given by the 2nd equation in (46).

Finally, using the "shifting" method on the likelihood function given in (44) which is described in Section 3.2, we follow the derivation of equation (37) step by step, then with some algebraic work, we obtain *generalized profile estimating function* $\Psi_{M,n}(\beta)$, which is given by the 3rd equation in (46). $\square$