Supplementary Materials: Semiparametric analysis of mixture regression models with competing risks data

Appendix

We first introduce some notations. Let \( B(t) = \{b_{jk}(t)\}_{2 \times 2} \) be a 2 \times 2 matrix with components \( b_{jk}(t) = \int_0^t \{s_{0j}^{(k)}(u)/s_{0j}^{(0)}(u)\} \, d\Lambda_{j0}(u) \), \( j, k = 1, 2 \). Define \( V(t, u) \equiv \{V_{jk}(t, u)\}_{2 \times 2} = [\exp\{B(t)\}]^{-1} \exp\{B(u)\} \), and for \( j, k = 1, 2, i = 1, \ldots, n, \)

\[
w_{jk}(t) = \{s_{ij}^{(k)}(t)s_{0j}^{(0)}(u) - s_{ij}^{(0)}(u)s_{0j}^{(k)}(t)\}/\{s_{0j}^{(0)}(u)\}^2,
\]

\[
a_{jk}(u) = E \left\{ \int_u^\infty w_{j1}(t)V_{1k}(t, u)dN_{11}(t) + \int_0^u w_{j2}(t)V_{2k}(t, u)dN_{21}(t) \right\},
\]

\[
r_{ji}(t) = \int_0^t \{V_{j1}(t, u)/s_{01}^{(0)}(u)\}dM_{11}(u) + \int_0^t \{V_{j2}(t, u)/s_{02}^{(0)}(u)\}dM_{21}(u),
\]

\[
\zeta_{kl} = E\{W_1Y_1(u)I(\eta_1 = 0)g_{11}^{(k)}(1)X_k(X_1, u)\}.
\]

Furthermore, define \( g_{ji}^{(1)}(t) = \partial g_{ji}(t; \theta_0, \Lambda_0)/\partial \beta_1 \), \( g_{ji}^{(2)}(t) = \partial g_{ji}(t; \theta_0, \Lambda_0)/\partial \beta_2 \), \( g_{ji}^{(3)}(t) = \partial g_{ji}(t; \theta_0, \Lambda_0)/\partial \gamma \), and \( e_{jk}(t) = \{s_{ij}^{(0)}(u)\}^{I(j=k)} + E\{Y_1(t)\exp(\beta_{j0}^T Z_1)g_{j1}^{(k)}(t)\}, \)

\( j = 1, 2, k = 1, 2, 3 \). Let \( \tilde{B}(t) = \text{diag}\{B(t) \otimes I_p, B(t) \otimes I_p, B(t) \otimes I_q\} \), where \( I_p \) and \( I_q \) are the \( p \times p \) and \( q \times q \) identity matrices, and \( \otimes \) stands for Kronecker product so that \( \tilde{B}(t) \) is a \((4p + 2q) \times (4p + 2q)\) matrix. Let \( \tilde{V}(t, u) = \exp\{\tilde{B}(t)\}^{-1} \exp\{\tilde{B}(u)\} = \text{diag}\{\tilde{V}^{(1)}(t, u), \tilde{V}^{(2)}(t, u)\}, \)

where \( \tilde{V}^{(k)}(t, u), k = 1, 2, \) can be written as

\[
\begin{pmatrix}
\tilde{V}_{11}^{(k)}(t, u) & \tilde{V}_{12}^{(k)}(t, u) \\
\tilde{V}_{21}^{(k)}(t, u) & \tilde{V}_{22}^{(k)}(t, u)
\end{pmatrix}.
\]

Here \( \tilde{V}_{ij}^{(1)}(t, u) \) are \( p \times p \) matrices and \( \tilde{V}_{ij}^{(2)}(t, u) \) are \( q \times q \) matrices, \( i, j = 1, 2 \). For \( j = 1, 2 \) and \( k = 1, 2, 3 \), define

\[
c_{jk}(t) = \int_0^t \left\{ \frac{\tilde{V}_{j1}^{(I(k<3)+2I(k=3))}(t, u)e_{1k}(u)}{s_{01}^{(0)}(u)} \, d\Lambda_{10}(u) + \frac{\tilde{V}_{j2}^{(I(k<3)+2I(k=3))}(t, u)e_{2k}(u)}{s_{02}^{(0)}(u)} \, d\Lambda_{20}(u) \right\}.
\]
\[
\omega_{ij}^{(k)}(t) = E \left[ \{Z_1 Y_1(t) \exp(\beta_{j0}^T Z_1) g_{j1}^{(0)}(t)\}^{I(j=k)} + Y_1(t) \exp(\beta_{j0}^T Z_1) \times \left\{ g_{j1}^{(1)}(t) + g_{j1}^{(2)}(t) b_{1k}(t) \right\} \right],
\]

\[
\omega_{ij}^{(k)}(t) = E \left[ \{Z_1^\otimes 2 Y_1(t) \exp(\beta_{j0}^T Z_1) g_{j1}^{(0)}(t)\}^{I(j=k)} + Z_1 Y_1(t) \exp(\beta_{j0}^T Z_1) \times \left\{ g_{j1}^{(1)}(t) + g_{j1}^{(2)}(t) b_{1k}(t) \right\} \right].
\]

Regularity conditions (i), (ii) and (iii) are assumed here to ensure the uniform law of large numbers and the functional central limit theorem (Pollard, 1990). To prove the results established in Theorems 1 and 2, we need to show the consistency of \( \hat{\theta} \) to \( \theta_0 \) and the uniform consistency of \( \hat{\Lambda}(t) \) to \( \Lambda_0(t) \) on \([0, \tau]\). The proof of consistency can be similarly derived as in Lu and Ying (2004), which is omitted here for brevity of presentation. The remainder of the proof of Theorems 1 and 2 are given below.

**Proof of Theorem 1**

Let \( U(\theta) = \{U_1(\theta)^T, U_2(\theta)^T, U_3(\theta)^T\}^T \). By Taylor expansion and some empirical process approximation techniques, we have

\[-U(\theta_0) = U(\hat{\theta}) - U(\theta_0) = \{n^{-1/2} \partial U(\theta) / \partial \theta^T|_{\theta = \theta_0}\} n^{-1/2}(\hat{\theta} - \theta_0) + o_p(1).\]

We shall show in step (A) that \( U(\theta_0) \) has a strong iid representation, i.e. \( U(\theta_0) = n^{-1/2} \sum_{i=1}^n \psi_i + o_p(1) \), and in step (B) that \(-n^{-1/2} \partial U(\theta) / \partial \theta^T|_{\theta = \theta_0}\) converges to a deterministic matrix \( \Phi \). The definitions of \( \psi_i \) and \( \Phi \) are given in Section 3.2. By the multivariate central limit theorem and Slutsky’s theorem, Theorem 1 then follows.

**Step A:** By some simple algebra, we can show that

\[
U_j(\theta_0) = n^{-1/2} \sum_{i=1}^n \{Z_i - \mu_{Z,j}(t)\} dM_{ji}(t) - \tilde{U}_j(\hat{\Lambda}(\theta_0)) + o_p(1), \quad j = 1, 2,
\]

where \( \tilde{U}_j(\Lambda) \) is given by

\[
n^{-1/2} \sum_{i=1}^n \int_0^\infty \left[ \sum_{k=1}^n Z_k Y_k(t) \exp(\beta_{j0}^T Z_k) g_{jk}(t; \theta_0, \Lambda) - \sum_{k=1}^n Z_k Y_k(t) g_{jk}^{(0)}(t) \right] dN_{1i}(t).
\]

By applying Taylor expansion of \( \tilde{U}_j(\Lambda) \) around \( \Lambda_0 \) and the uniform law of large numbers, we have

\[
U_j(\theta_0) = n^{-1/2} \sum_{i=1}^n \{Z_i - \mu_{Z,j}(t)\} dM_{ji}(t) - n^{-1} \sum_{i=1}^n \int_0^\infty w_{ji}(t) n^{1/2}\{\hat{\Lambda}_1(t; \theta_0) - \Lambda_{10}(t)\} dN_{1i}(t)
\]
\[-n^{-1} \sum_{i=1}^{n} \int_{0}^{\infty} w_{j2}(t) n^{1/2} \{ \hat{\Lambda}_2(t; \theta_0) - \Lambda_{20}(t) \} dN_{2i}(t) + o_p(1). \]  

(A.1)

By the definition of \( \hat{\Lambda}(\theta_0) \), we have

\[ n^{-1/2} \sum_{i=1}^{n} [dN_{ji}(t) - Y_i(t) \exp(\beta_{0j}^T Z_i) g_{ji} \{ t; \theta_0, \hat{\Lambda}(\theta_0) \}] = 0, \quad j = 1, 2. \]

It then follows from the definitions of \( M_{ji}(t) (j = 1, 2) \) and the uniform consistency of \( \hat{\Lambda}(t; \theta_0) \) that

\[
\begin{align*}
n^{-1/2} & \sum_{i=1}^{n} M_{ji}(t) \\
& = n^{-1/2} \sum_{i=1}^{n} \int_{0}^{t} Y_i(s) \exp(\beta_{0j}^T Z_i) [g_{ji} \{ s; \theta_0, \hat{\Lambda}(\theta_0) \} d\hat{\Lambda}_j(s; \theta_0) - g_{ji}^{(0)}(s) d\Lambda_{j0}(s)] \\
& = n^{-1/2} \sum_{i=1}^{n} \int_{0}^{t} Y_i(s) \exp(\beta_{0j}^T Z_i) g_{ji}^{(0)}(s) d\{ \hat{\Lambda}_j(s; \theta_0) - \Lambda_{10}(s) \} \\
& + n^{-1/2} \sum_{i=1}^{n} \int_{0}^{t} Y_i(s) \exp(\beta_{0j}^T Z_i) g_{ji}^{(1)}(s) \{ \hat{\Lambda}_1(s; \theta_0) - \Lambda_{10}(s) \} d\Lambda_{10}(s; \theta_0) \\
& + n^{-1/2} \sum_{i=1}^{n} \int_{0}^{t} Y_i(s) \exp(\beta_{0j}^T Z_i) g_{ji}^{(2)}(s) \{ \hat{\Lambda}_2(s; \theta_0) - \Lambda_{20}(s) \} d\Lambda_{20}(s; \theta_0) + o_p(1).
\end{align*}
\]

Taking the derivative with respect to \( t \) on both sides and applying the uniform law of large numbers, we get

\[
\begin{pmatrix}
-n^{-1/2} \sum_{i=1}^{n} \{ s_{01}^{(0)}(t) \}^{-1} dM_{1i}(t) \\
n^{-1/2} \sum_{i=1}^{n} \{ s_{02}^{(0)}(t) \}^{-1} dM_{2i}(t)
\end{pmatrix}
= d \begin{pmatrix}
\hat{\Lambda}_1(t; \theta_0) - \Lambda_{10}(t) \\
\hat{\Lambda}_2(t; \theta_0) - \Lambda_{20}(t)
\end{pmatrix}
+ \{ dB(t) \}
\begin{pmatrix}
\hat{\Lambda}_1(t; \theta_0) - \Lambda_{10}(t) \\
\hat{\Lambda}_2(t; \theta_0) - \Lambda_{20}(t)
\end{pmatrix}
+ o_p(1).
\]

Therefore,

\[
\begin{align*}
n^{1/2} \begin{pmatrix}
\hat{\Lambda}_1(t; \theta_0) - \Lambda_{10}(t) \\
\hat{\Lambda}_2(t; \theta_0) - \Lambda_{20}(t)
\end{pmatrix}
& = n^{-1/2} \sum_{i=1}^{n} \int_{0}^{t} \{ \exp\{B(t)\} \}^{-1} \exp\{B(u)\} \begin{pmatrix}
\{ s_{01}^{(0)}(u) \}^{-1} dM_{1i}(u) \\
\{ s_{02}^{(0)}(u) \}^{-1} dM_{2i}(u)
\end{pmatrix}
\end{align*}
\]

\[ + o_p(1) = n^{-1/2} \sum_{i=1}^{n} \begin{pmatrix}
r_{1i}(t) \\
r_{2i}(t)
\end{pmatrix}
+ o_p(1). \quad \text{(A.2)}
\]

3
By empirical process approximation techniques, it then follows from (A.1) and (A.2) that

\[
U_j(\theta_0) = n^{-1/2} \sum_{i=1}^{n} \int_{0}^{\infty} \{Z_i - \mu_{Z_j}(t)\}dM_{ji}(t) - n^{-3/2} \sum_{i=1}^{n} \sum_{k=1}^{n} \int_{0}^{\infty} w_{j1}(t)r_{1k}(t)dN_{1i}(t) \]

\[-n^{-3/2} \sum_{i=1}^{n} \sum_{k=1}^{n} \int_{0}^{\infty} w_{j2}(t)r_{2k}(t)dN_{2i}(t) + o_p(1) \]

\[= n^{-1/2} \sum_{i=1}^{n} \int_{0}^{\infty} \{Z_i - \mu_{Z_j}(t)\}dM_{ji}(t) - n^{-1/2} \int_{0}^{\infty} \frac{a_{j1}(u)}{s_{01}^{(0)}(u)}dM_{1i}(u) \]

\[-n^{-1/2} \int_{0}^{\infty} \frac{a_{j2}(u)}{s_{02}^{(0)}(u)}dM_{2i}(u) + o_p(1) = n^{-1/2} \sum_{i=1}^{n} \psi_{ji} + o_p(1), \quad j = 1, 2.\]

Applying similar techniques to \(U_3(\theta_0)\), we get

\[
U_3(\theta_0) = n^{-1/2} \sum_{i=1}^{n} W_i \{I(\eta_i = 1) + I(\eta_i = 0)g_{11}^{(0)}(X_i) - \pi(\gamma_0^TW_i)\} \]

\[+ n^{-3/2} \sum_{i=1}^{n} \sum_{k=1}^{n} W_i I(\eta_i = 0)g_{11}^{(1)}(X_i) \left\{ \int_{0}^{\infty} \sum_{l=1}^{2} Y^i_l(u)v_{1l}(X_i,u) \frac{dM_{lk}(u)}{s_{0l}^{(0)}(u)} \right\} \]

\[+ n^{-3/2} \sum_{i=1}^{n} \sum_{k=1}^{n} W_i I(\eta_i = 0)g_{11}^{(2)}(X_i) \left\{ \int_{0}^{\infty} \sum_{l=1}^{2} Y^i_l(u)v_{2l}(X_i,u) \frac{dM_{lk}(u)}{s_{0l}^{(0)}(u)} \right\} + o_p(1) \]

\[= n^{-1/2} \sum_{i=1}^{n} W_i \{I(\eta_i = 1) + I(\eta_i = 0)g_{11}^{(0)}(X_i) - \pi(\gamma_0^TW_i)\} + \sum_{r=1}^{2} \sum_{l=1}^{2} \int_{0}^{\infty} \frac{\zeta_{rl}(u)}{s_{0l}^{(0)}(u)}dM_{lk}(u) \]

\[+ o_p(1) = n^{-1/2} \sum_{i=1}^{n} \psi_{3i} + o_p(1).\]

Therefore, \(U(\theta_0) = n^{-1/2} \sum_{i=1}^{n} \psi_i + o_p(1).\)

**Step B:** Let \(\varrho_{j1}(t) = \partial \hat{\Lambda}_j(t; \theta)/\partial \beta_1|_{\theta = \theta_0}, \varrho_{j2}(t) = \partial \hat{\Lambda}_j(t; \theta)/\partial \beta_2|_{\theta = \theta_0}\) and \(\varrho_{j3}(t) = \partial \hat{\Lambda}_j(t; \theta)/\partial \gamma|_{\theta = \theta_0}\). It follows from the definition of \(\hat{\Lambda}(t; \theta)\) that

\[n^{-1/2} \sum_{i=1}^{n} [dN_{ji} - Y^i(t) \exp(\beta^T_j Z_i)g_{ji}\{t; \hat{\Lambda}(\theta)\}d\hat{\Lambda}_j(t; \theta)] = 0, \quad j = 1, 2.\]

Taking partial derivative with respect to \(\theta\) on both sides of the above equations, straightforward manipulations give that for \(j = 1, 2\) and \(k = 1, 2, 3\),

\[\{s_{0j}^{(1)}(t)d\Lambda_{j0}(t)\} \varrho_{1k}(t) + \{s_{0j}^{(2)}(t)d\Lambda_{j0}(t)\} \varrho_{2k}(t) + s_{0j}^{(0)} \varrho_{jk}(t) = -\epsilon_{jk}(t)d\Lambda_{j0}(t) + o_p(1).\]
Theorem (Pollard, 1990, Theorem 10.6). The weak convergence of $n^{1/2}\{\hat{\Lambda}(t) - \Lambda_0(t)\}$ follows from the functional central limit theorem (Pollard, 1990, Theorem 10.6).

**Proof of Theorem 3**
Let $\Lambda^*(t; \theta) = \{\Lambda_1^*(t; \theta), \Lambda_2^*(t; \theta)\}$ denote the solution of equation (15) and (17) for fixed $\theta$. Define $U^*(\theta) = \{U_1^*(\theta)\}^T, U_2^*(\theta), U_3^*(\theta)\}^T$, where

$$U_j^*(\theta) = n^{-1/2} \sum_{i=1}^{n} \int_0^\infty \xi_i \left[ Z_i - \frac{\sum_{k=1}^{n} \xi_k Z_k Y_k(t) \exp(\beta_j^T Z_k) g_{jk}\{t; \theta, \Lambda^*(t; \theta)\}}{\sum_{k=1}^{n} \xi_k Z_k Y_k(t) \exp(\beta_j^T Z_k) g_{jk}\{t; \theta, \Lambda^*(t; \theta)\}} \right] dN_{ji}(t), j = 1, 2,$$

$$U_3^*(\theta) = n^{-1/2} \sum_{i=1}^{n} \xi_i W_i \left[ I(\eta_i = 1) + I(\eta_i = 0) g_{i0} \{X_i; \theta, \Lambda^*(t; \theta)\} - \pi(\gamma^T W_i) \right].$$

By the consistency of $\hat{\theta}$ and $\hat{\Lambda}$, and $E(\xi_1) = 1$, following the step B of the proof of Theorem 1, it is easy to show that $-n^{-1/2} \partial U^*(\theta)/\partial \theta^T|_{\theta = \hat{\theta}} \rightarrow P \Phi$. Thus, it suffices to show that for every realization of $\{\xi_1, \ldots, \xi_n\}$ the conditional distribution of $U^*(\hat{\theta})$ given the observed data converges to a normal random vector with mean 0 and covariance matrix $E(\psi_1 \psi_1^T)$. Note that $U(\hat{\theta}) = 0$ and for $j = 1, 2$

$$U_j(\hat{\theta}) = n^{-1/2} \sum_{i=1}^{n} \int_0^\infty \left[ Z_i - \frac{\sum_{i=1}^{n} Z_i Y_i(t) \exp(\hat{\beta}_j^T Z_i) g_{ji}\{t; \hat{\theta}, \hat{\Lambda}(t)\}}{\sum_{i=1}^{n} Y_i(t) \exp(\hat{\beta}_j^T Z_i) g_{ji}\{t; \hat{\theta}, \hat{\Lambda}(t)\}} \right] d\hat{M}_{ji}(t),$$

where $\hat{M}_{ji}(t) = N_{ji}(t) - \int_0^t Y_i(u) \exp(\hat{\beta}_j^T Z_i) g_{ji}\{u; \hat{\theta}, \hat{\Lambda}(t)\} d\hat{\Lambda}_j(u)$. By simple algebra, we have for $j = 1, 2$

$$U_j^*(\hat{\theta}) = U_j^*(\hat{\theta}) - U_j(\hat{\theta})$$

$$= n^{-1/2} \sum_{i=1}^{n} \int_0^\infty \xi_i \left[ Z_i - \frac{\sum_{k=1}^{n} \xi_k Z_k Y_k(t) \exp(\hat{\beta}_j^T Z_k) g_{jk}\{t; \hat{\theta}, \hat{\Lambda}^*(t; \hat{\theta})\}}{\sum_{k=1}^{n} \xi_k Y_k(t) \exp(\hat{\beta}_j^T Z_k) g_{jk}\{t; \hat{\theta}, \hat{\Lambda}^*(t; \hat{\theta})\}} \right] dN_{ji}(t)$$

$$- n^{-1/2} \sum_{i=1}^{n} \int_0^\infty \left[ Z_i - \frac{\sum_{i=1}^{n} Z_i Y_i(t) \exp(\hat{\beta}_j^T Z_i) g_{ji}\{t; \hat{\theta}, \hat{\Lambda}(t)\}}{\sum_{i=1}^{n} Y_i(t) \exp(\hat{\beta}_j^T Z_i) g_{ji}\{t; \hat{\theta}, \hat{\Lambda}(t)\}} \right] d\hat{M}_{ji}(t)$$

$$= n^{-1/2} \sum_{i=1}^{n} \int_0^\infty \xi_i \left[ Z_i - \frac{\sum_{k=1}^{n} \xi_k Z_k Y_k(t) \exp(\hat{\beta}_j^T Z_k) g_{jk}\{t; \hat{\theta}, \hat{\Lambda}(t)\}}{\sum_{k=1}^{n} \xi_k Z_k Y_k(t) \exp(\hat{\beta}_j^T Z_k) g_{jk}\{t; \hat{\theta}, \hat{\Lambda}(t)\}} \right] d\hat{M}_{ji}(t)$$

$$- n^{-1/2} \sum_{i=1}^{n} \int_0^\infty \left[ Z_i - \frac{\sum_{i=1}^{n} Z_i Y_i(t) \exp(\hat{\beta}_j^T Z_i) g_{ji}\{t; \hat{\theta}, \hat{\Lambda}(t)\}}{\sum_{i=1}^{n} Y_i(t) \exp(\hat{\beta}_j^T Z_i) g_{ji}\{t; \hat{\theta}, \hat{\Lambda}(t)\}} \right] d\hat{M}_{ji}(t)$$

$$+ n^{-1/2} \sum_{i=1}^{n} \int_0^\infty \xi_i \left[ \frac{\sum_{k=1}^{n} \xi_k Z_k Y_k(t) \exp(\hat{\beta}_j^T Z_k) g_{jk}\{t; \hat{\theta}, \hat{\Lambda}^*(t; \hat{\theta})\}}{\sum_{k=1}^{n} \xi_k Y_k(t) \exp(\hat{\beta}_j^T Z_k) g_{jk}\{t; \hat{\theta}, \hat{\Lambda}^*(t; \hat{\theta})\}} \right] dN_{ji}(t).$$
Then it follows from some empirical process approximation techniques that

\[
U_j^*(\hat{\theta}) = n^{-1/2} \sum_{i=1}^{n} \left( \xi_i - 1 \right) \left[ Z_i - \frac{\sum_{i=1}^{n} Z_i Y_i(t) \exp(\hat{\beta}_j^T Z_i) g_{ij}(t; \hat{\theta}, \hat{\Lambda}(t))}{\sum_{i=1}^{n} Y_i(t) \exp(\hat{\beta}_j^T Z_i) g_{ij}(t; \hat{\theta}, \hat{\Lambda}(t))} \right] d\hat{M}_{ji}(t)
\]

\[
- n^{-1} \sum_{k=1}^{2} \sum_{i=1}^{n} \int_{0}^{\infty} \hat{w}_{jk}(t) n^{1/2} \left\{ \Lambda_k^*(t; \hat{\theta}) - \hat{\Lambda}_k(t) \right\} \xi_i dN_{ki}(t) + o_p(1),
\]

where \( \hat{w}_{jk}(t) \) is obtained by substituting \((\hat{\theta}, \hat{\Lambda})\) for \((\theta_0, \Lambda_0)\) and replacing the expectation \( E \) by its empirical counterpart. In the sequel, an estimator \( \hat{h} \) of a deterministic quantity \( h \) is similarly defined. Following the arguments for (A.2) and repeatedly using \( E(\xi) = 1 \) and the consistency of \( \hat{\theta} \) and \( \hat{\Lambda}(t) \), it can be shown that \( n^{1/2} \left\{ \Lambda^*(t; \hat{\theta}) - \hat{\Lambda}(t) \right\} = n^{-1/2} \sum_{i=1}^{n} (\xi_i - 1) \tilde{r}_i(t) + o_p(1) \). This gives that for \( j = 1, 2 \),

\[
U_j^*(\hat{\theta}) = n^{-1/2} \sum_{i=1}^{n} (\xi_i - 1) \left[ Z_i - \frac{\sum_{i=1}^{n} Z_i Y_i(t) \exp(\hat{\beta}_j^T Z_i) g_{ij}(t; \hat{\theta}, \hat{\Lambda}(t))}{\sum_{i=1}^{n} Y_i(t) \exp(\hat{\beta}_j^T Z_i) g_{ij}(t; \hat{\theta}, \hat{\Lambda}(t))} \right] d\hat{M}_{ji}(t)
\]

\[
- n^{-1/2} \sum_{i=1}^{n} (\xi_i - 1) \sum_{l=1}^{2} \int_{0}^{\infty} \frac{\hat{a}_{j1}(u) + \hat{a}_{j2}(u)}{\bar{s}_0(u)} d\hat{M}_{li}(u) + o_p(1)
\]

\[
= n^{-1/2} \sum_{i=1}^{n} (\xi_i - 1) \hat{\psi}_{ji} + o_p(1).
\]

Employing similar techniques, we get

\[
U_3^*(\hat{\theta}) = U_3^*(\hat{\theta}) - U_3(\hat{\theta}) = n^{-1/2} \sum_{i=1}^{n} (\xi_i - 1) W_i \left[ I(\eta_i = 1) + I(\eta_i = 0) g_{1i}(X_i; \hat{\theta}, \hat{\Lambda}(t)) - \pi(\hat{\gamma}^T W_i) \right]
\]

\[
+ n^{-1/2} \sum_{i=1}^{n} \xi W_i I(\eta_i = 0) \left[ g_{1i}(X_i; \hat{\theta}, \Lambda^*(t, \hat{\theta})) - g_{1i}(X_i; \hat{\theta}, \hat{\Lambda}(t)) \right]
\]

\[
= n^{-1/2} \sum_{i=1}^{n} (\xi_i - 1) W_i \left[ I(\eta_i = 1) + I(\eta_i = 0) g_{1i}(X_i; \hat{\theta}, \hat{\Lambda}(t)) - \pi(\hat{\gamma}^T W_i) \right]
\]

\[
+ n^{-1/2} \sum_{k=1}^{2} \sum_{l=1}^{2} \sum_{i=1}^{n} (\xi_i - 1) \int_{0}^{\infty} \frac{\hat{\zeta}_{kl}(u)}{\bar{s}_0(u)} d\hat{M}_{li}(u) + o_p(1)
\]

\[
= n^{-1/2} \sum_{i=1}^{n} (\xi_i - 1) \hat{\psi}_{3i} + o_p(1).
\]

Therefore, \( U^*(\hat{\theta}) = n^{-1/2} \sum_{i=1}^{n} (\xi_i - 1) \hat{\psi}_i + o_p(1) \). Note that \( E(\xi_i) = 1 \), \( E\{((\xi_i - 1)^2)\} = 1 \) and \( \lim_{n \to \infty} (1/n) \sum_{i=1}^{n} \hat{\psi}_i^{\otimes 2} = E(\psi_1 \psi_1^T) \). It follows from the multivariate central limit theorem that the limiting distribution of \( U^*(\hat{\theta}) \) given the observed data is the same as
that of $U(\theta_0)$. In addition, we can show that

$$n^{1/2}\{\Lambda^*(t) - \hat{\Lambda}(t)\} = n^{1/2}\{\Lambda^*(t; \theta^*) - \Lambda^*(t; \hat{\theta})\} + n^{1/2}\{\Lambda^*(t; \hat{\theta}) - \hat{\Lambda}(t)\}$$

$$= \left\{\frac{\partial \Lambda^*(t; \theta)}{\partial \theta^T}|_{\theta = \hat{\theta}}\right\}n^{1/2}(\theta^* - \hat{\theta}) + n^{-1/2}\sum_{i=1}^{n}(\xi_i - 1)\hat{r}_i(t) + o_p(1)$$

$$= n^{-1/2}\sum_{i=1}^{n}(\xi_i - 1)\{\hat{H}(t)\hat{\Phi}^{-1}\hat{\psi}_i + \hat{r}_i(t)\} + o_p(1).$$

Thus, by the functional central limit theorem, the limiting distribution of $n^{1/2}\{\Lambda^*(t) - \hat{\Lambda}(t)\}$ given the observed data is the same as that of $n^{1/2}\{\hat{\Lambda}(t) - \Lambda_0(t)\}$. 
