Supplementary Materials for "A Class of Functional Methods for Error-Contaminated Survival Data under Additive Hazards Models with Replicate Measurements"

BY

Ying Yan and Grace Y. Yi

Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West

Waterloo, Ont., Canada N2L 3G1

Appendix

In the following derivations, we introduce some notations. For an $m \times 1$ vector $a = (a_1, a_2, \cdots, a_m)^T$, let $||a|| = (\sum a_i^2)^{1/2}$ denote the Euclidean norm for $a$. For a matrix $A$, define $||A|| = \max_{i,j} |a_{ij}|$, where $a_{ij}$ is the $(i, j)$th element of $A$. For vector processes $A_n(t)$ and $A(t)$, $A_n(t)$ is said to converge almost surely to $A(t)$ uniformly in $t$ if

$$\sup_{0 \leq t \leq \tau} ||A_n(t) - A(t)|| \overset{\text{a.s.}}{\to} 0, \text{ as } n \to \infty.$$ 

We define that a random matrix $A_n = o_{a.s.}(1)$ in the sense that $\Pr\{\lim_{n \to \infty} ||A_n|| = 0\} = 1$.

Appendix A1: Regularity Conditions

R1. $\{N_i(\cdot), Y_i(\cdot), Z_i(\cdot)\}, i = 1, \cdots, n$ are independent and identically distributed.

R2. $\Pr\{Y_i(\tau) = 1\} > 0$ for $i = 1, \cdots, n$.

R3. $T_i$ and $C_i$ are conditionally independent given $Z_i(t)$, $i = 1, \cdots, n$.

R4. $\sup_{t \in [0, \tau]} ||E\{Z_i^\otimes 2(t)\}|| < \infty, i = 1, \cdots, n.$

R5. Bounded variation condition: for $i = 1, \cdots, n; j = 1, \cdots, p + q$,

$$|Z_{ij}(0)| + \int_0^\tau |dZ_{ij}(u)| \leq K$$
holds almost surely for all the sample path, where \( K \) is a constant.

R6. All the \( n_i(i = 1, \cdots, n) \) are bounded by a constant \( N_0 \), and \( \lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} I\{n_i = j\} \) exists, where \( j = 1, \cdots, N_0 \).

R7. \( \|E(\epsilon_{ir}^{\otimes 2})\| < \infty, i = 1, \cdots, n; r = 1, \cdots, n_i. \)

R8. \( \int_0^T E[Y_i(t)\{Z_i(t) - e(t)\}^{\otimes 2}] \) and \( \Sigma_c \) are positive definite, \( i = 1, \cdots, n. \)

These regularity conditions are imposed for the technical development. The conditions R1, R2, R3, R4, R5 and R8 are conventionally used for developing asymptotic theory in survival analysis, and they are analogous to those by, for example, Andersen and Gill (1982), Spiekerman and Lin (1998), Lin, Wei, Yang and Ying (2000), and Hu and Lin (2004). In particular, condition R1 assumes homogeneity among the subjects in the study. Condition R2 says that each subject in the study has a positive probability to be observed, and this condition ensures the denominator of \( e(t) = E[Y_i(t)Z_i(t)]/E[Y_i(t)] \) to be bounded away from zero. Condition R3 is a common assumption for censoringship. Conditions R4 and R5 control the variability of the covariates in which condition R5 is a key assumption for using empirical process theory (e.g., Strong Uniform Law of Large Numbers by Pollard 1990). Condition R6 guarantees the existence of \( \rho_0 \) and \( \rho_1 \); imposing the upper bound for the \( n_i \) is often plausible as in practice, the \( n_i \) are usually not large. Condition R7 controls the variability of the error terms \( \epsilon_{ir} \), and condition R8 is needed for developing asymptotic normality of the proposed estimators.

**Appendix A2: Proof of the relationship (7) and its consequence**

Note that

\[
\hat{\beta}_{nv} = \left[ n^{-1} \sum_{i=1}^{n} \int_0^T Y_i(t) \left\{ \hat{Z}_i(t) - \bar{Z}(t) \right\}^{\otimes 2} dt \right]^{-1} \left[ n^{-1} \sum_{i=1}^{n} \int_0^T \left\{ \hat{Z}_i(t) - \bar{Z}(t) \right\} dN_i(t) \right].
\]

Let

\[
D_{nv} = n^{-1} \sum_{i=1}^{n} \int_0^T Y_i(t) \left\{ \hat{Z}_i(t) - \bar{Z}(t) \right\}^{\otimes 2} dt
\]
denote the denominator of $\hat{\beta}_{nv}$. By the Strong Uniform Law of Large Numbers (USLLN) (Pollard 1990), it follows that $D_{nv} \to D_{nv}$ almost surely as $n \to \infty$, where

$$D_{nv} = E \left[ \int_0^T Y_i(t) \{Z_i(t) - e(t)\}^2 dt \right] + \int_0^T \rho_0 E \{Y_i(t)\} \Sigma_1 dt \triangleq B_1 + B_2.$$ 

Indeed, we have

$$n^{-1} \sum_{i=1}^n \int_0^T Y_i(t) \left\{ \hat{Z}_i(t) - \tilde{Z}(t) \right\}^2 dt$$

$$= n^{-1} \sum_{i=1}^n \int_0^T Y_i(t) \left\{ \hat{Z}_i(t) - e(t) \right\}^2 dt + n^{-1} \sum_{i=1}^n \int_0^T Y_i(t) \left\{ e(t) - \tilde{Z}(t) \right\}^2 dt$$

$$+ n^{-1} \sum_{i=1}^n \int_0^T 2Y_i(t) \left\{ \hat{Z}_i(t) - e(t) \right\} \left\{ e(t) - \tilde{Z}(t) \right\}^T dt$$

$$= n^{-1} \sum_{i=1}^n \int_0^T Y_i(t) \left\{ \hat{Z}_i(t) - e(t) \right\}^2 dt$$

$$+ \int_0^T \left( -n^{-1} \sum_{i=1}^n Y_i(t)e^2(t) - n^{-1} \frac{\left\{ \sum_{j=1}^n Y_j(t)\hat{Z}_j(t) \right\}^2}{\sum_{j=1}^n Y_j(t)} + 2n^{-1} \sum_{i=1}^n Y_i(t)\hat{Z}_i(t)e^T(t) \right) dt.$$ 

By Condition R5, each component of the vector $n^{-1} \sum_{i=1}^n Y_i(t)Z_i(t)$ is of bounded variation, so $n^{-1} \sum_{i=1}^n Y_i(t)Z_i(t)$ is the difference of two nondecreasing functions. By Lemma A.1 and A.2 of Bilias, Gu and Ying (1997), $n^{-1} \sum_{i=1}^n Y_i(t)Z_i(t)$ is manageable (Pollard 1990, p.38), and $n^{-1} \sum_{i=1}^n Y_i(t)\hat{Z}_i(t)$ is manageable. Furthermore, Conditions R4, R6 and R7 justify that the envelope (Pollard 1990 p.19) $F_i$ of $\{ \hat{Z}_i(t), t \in [0, \tau] \}$ is finite and could only take at maximum $N_0$ possible values, and thus $\max_{i=1} E(||F_i^{\otimes 2}||) < \infty$. It follows that $\sum_{i=1}^\infty E(||F_i^{\otimes 2}||) < \infty$. The two conditions of the Strong Uniform Law of Large Numbers (SULLN) (Pollard 1990, p.41) are thus verified. Consequently, we obtain that $n^{-1} \sum_{i=1}^n Y_i(t)\hat{Z}_i(t) \xrightarrow{a.s.} E\{Y_i(t)Z_i(t)\}$ uniformly in $t$. Similarly, $n^{-1} \sum_{i=1}^n Y_i(t) \xrightarrow{a.s.} E\{Y_i(t)\}$ uniformly in $t$. By SULLN together with Condition R2, $\sum_{i=1}^n Y_i(t)\hat{Z}_i(t)/\sum_{i=1}^n Y_i(t) \xrightarrow{a.s.} e(t)$ un-
formly in $t$. Thus, we obtain that

$$
\int_0^T \left( -n^{-1} \sum_{i=1}^n Y_i(t) e^{\otimes 2}(t) - n^{-1} \left\{ \sum_{j=1}^n Y_j(t) \hat{Z}_j(t) \right\}^{\otimes 2} + 2n^{-1} \sum_{i=1}^n Y_i(t) \hat{Z}_i(t) e^T(t) \right) dt
$$

$$
= \int_0^T \left( -E\{Y_i(t)\} e^{\otimes 2}(t) - \left[ \frac{E\{Y_j(t) \hat{Z}_j(t)\}}{E\{Y_j(t)\}} \right]^{\otimes 2} + 2E\{Y_i(t) \hat{Z}_i(t)\} e^T(t) \right) dt + o_a.s.(1)
$$

$$
= \int_0^T \left( -E\{Y_i(t)\} e^{\otimes 2}(t) - \left[ \frac{E\{Y_j(t) Z_j(t)\}}{E\{Y_j(t)\}} \right]^{\otimes 2} + 2E\{Y_i(t) Z_i(t)\} e^T(t) \right) dt + o_a.s.(1)
$$

$$
= \int_0^T \left[ -E\{Y_i(t)\} e^{\otimes 2}(t) - E\{Y_i(t)\} e^{\otimes 2}(t) + 2E\{Y_i(t)\} e^{\otimes 2}(t) \right] dt + o_a.s.(1)
$$

$$
= o_a.s.(1),
$$

where the second last identity follows from the definition of $e(t)$. Let $\hat{\epsilon}_i = (n_i^{-1} \sum_{k=1}^{n_i} \epsilon_{ik}^T, 0^T)^T$, and $\rho_0 = \lim_{n \to \infty} n^{-1} \sum_{i=1}^n n_i^{-1}$, where the existence of the limit is ensured by the Regularity
Conditions. By SULLN and observing that $E\{\hat{Z}_i^{\otimes 2}(t)\} = E\{Z_i^{\otimes 2}(t)\} + \Sigma_1/n_i$, we obtain that

$$D_{nv} = n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ \hat{Z}_i(t) - e(t) \right\}^{\otimes 2} dt + o_{a.s.}(1)$$

$$= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^{\otimes 2} dt + n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \epsilon_i^{\otimes 2} dt$$

$$+ n^{-1} \sum_{i=1}^{n} 2Y_i(t) \left\{ Z_i(t) - e(t) \right\} \epsilon_i^{T} dt + o_{a.s.}(1)$$

$$= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^{\otimes 2} dt + n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \epsilon_i^{\otimes 2} dt$$

$$+ \int_{0}^{\tau} 2E[Y_i(t) \left\{ Z_i(t) - e(t) \right\} \epsilon_i^{T}] dt + o_{a.s.}(1)$$

$$= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^{\otimes 2} dt + n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \epsilon_i^{\otimes 2} dt + o_{a.s.}(1)$$

$$= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^{\otimes 2} dt + \int_{0}^{\tau} \lim_{n \to \infty} \left[ n^{-1} \sum_{i=1}^{n} E(Y_i(t) \epsilon_i^{\otimes 2}) \right] dt + o_{a.s.}(1)$$

$$= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^{\otimes 2} dt + \int_{0}^{\tau} E(Y_i(t)) \lim_{n \to \infty} \left\{ n^{-1} \sum_{i=1}^{n} E(\epsilon_i^{\otimes 2}) \right\} dt + o_{a.s.}(1)$$

$$= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^{\otimes 2} dt + \int_{0}^{\tau} E(Y_i(t)) \lim_{n \to \infty} \left\{ n^{-1} \sum_{i=1}^{n} \Sigma_1/n_i \right\} dt + o_{a.s.}(1)$$

$$= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^{\otimes 2} dt + \int_{0}^{\tau} \rho_0 E(Y_i(t)) \Sigma_1 dt + o_{a.s.}(1),$$

where the last two steps follow from the definition of $\Sigma_1$ and $\rho_0$. Thus, by the definition of $D_{nv}$, we obtain that

$$D_{nv} = D_{nv} + o_{a.s.}(1),$$

i.e., $D_{nv} \to D_{nv}$ almost surely as $n \to \infty$.

Similarly, by SULLN we have $n^{-1} \sum_{i=1}^{n} N_i(t) \overset{a.s.}{\to} E\{N_i(t)\}$ uniformly in $t$. By Lemma 1 of Lin, Wei, Yang and Ying (2000), $n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} \hat{Z}(t) dN_i(t) \overset{a.s.}{\to} \int_{0}^{\tau} e(t) E\{dN_i(t)\}$. Similarly,
\[ n^{-1} \sum_{i=1}^{n} \int_0^\tau \hat{Z}_i(t) dN_i(t) \overset{a.s.}{\to} \int_0^\tau E\{Z_i(t) dN_i(t)\}. \]
Thus, we obtain that
\[
\begin{align*}
& n^{-1} \sum_{i=1}^{n} \int_0^\tau \{\hat{Z}_i(t) - \bar{Z}(t)\} dN_i(t) \\
& = n^{-1} \sum_{i=1}^{n} \int_0^\tau \{\hat{Z}_i(t) - e(t)\} dN_i(t) + o_{a.s.}(1) \\
& = E \left[ \int_0^\tau \{\hat{Z}_i(t) - e(t)\} dN_i(t) \right] + o_{a.s.}(1) \\
& = E \left[ \int_0^\tau \{Z_i(t) - e(t)\} dN_i(t) \right] + o_{a.s.}(1).
\end{align*}
\]

As a result, we obtain that \( \hat{\beta}_{nv}^* \), the asymptotic limit of \( \hat{\beta}_{nv} \), is given by
\[
\begin{align*}
\beta_{nv}^* &= D_{nv}^{-1} E \left[ \int_0^\tau \{Z_i(t) - e(t)\} dN_i(t) \right] = (B_1 + B_2)^{-1} E \left[ \int_0^\tau \{Z_i(t) - e(t)\} dN_i(t) \right]. \quad (A1)
\end{align*}
\]

On the other hand, Lin and Ying (1994) showed that, in the absence of measurement error, the estimator
\[
\hat{\beta} = \left[ \sum_{i=1}^{n} \int_0^\tau Y_i(t) (Z_i(t) - \bar{Z}(t))^{\otimes 2} dt \right]^{-1} \left[ \sum_{i=1}^{n} \int_0^\tau \{Z_i(t) - \bar{Z}(t)\} dN_i(t) \right]
\]
converges in probability to \( \beta \). By analogy to the preceding arguments, we can show that
\[
\beta = B_1^{-1} E \left[ \int_0^\tau \{Z_i(t) - e(t)\} dN_i(t) \right]. \quad (A2)
\]

Comparison between (A1) and (A2) leads to the expression of \( \beta_{nv}^* \) and \( \beta \), we obtain that \( \beta_{nv}^* = (B_1 + B_2)^{-1} B_1 \beta \). The proof is then completed.

It is straightforward that if \( \|\beta\| = 0 \), then \( \|\beta_{nv}^*\| = 0 \). When \( Z_i(t) \) is univariate, then \( |(B_1 + B_2)^{-1} B_1| < 1 \), and it follows that \( |\beta_{nv}^*| < |\beta| \). If \( X_i \) and \( V_i(t) \) are univariates and are independent, and either \( V_i(t) \) or \( X_i \) are independent of the followup process, then \( B_1 \) is a \( 2 \times 2 \) diagonal matrix, i.e.,
\[
B_1 = \begin{pmatrix} \int_0^\tau E [Y_i(t)\{X_i - e_1(t)\}^2] dt & 0 \\ 0 & \int_0^\tau E [Y_i(t)\{V_i(t) - e_2(t)\}^2] dt \end{pmatrix},
\]

\[ 6 \]
where $e_1(t)$ and $e_2(t)$ are the two components of the vector $e(t)$. Note that

$$B_2 = \left( \begin{array}{cc} \rho_0 \Sigma \int_0^T E\{Y_i(t)\} dt & 0 \\ 0 & 0 \end{array} \right),$$

It follows that

$$(B_1 + B_2)^{-1} B_1 = \left( \begin{array}{cc} \int_0^T E\{Y_i(t)\} \{X_i - e_1(t)\}^2 dt & 0 \\ \int_0^T E\{Y_i(t)\} \{X_i - e_1(t)\}^2 dt + \rho_0 \Sigma \int_0^T E\{Y_i(t)\} dt & 0 \end{array} \right).$$

Therefore, the relationship (7) implies that $|\beta^{*}_{nv,x}| < |\beta_x|$ and $\beta^{*}_{nv,v} = \beta_v$. 
Appendix A3: Derivation of $E\{U_{nv}(\beta)|\mathcal{F}_\tau\}$

\[
E\{U_{nv}(\beta)|\mathcal{F}_\tau\} = \sum_{i=1}^{n} \int_{0}^{\tau} \left[ E\{\hat{Z}_i(t)|\mathcal{F}_\tau\} - E\{\bar{Z}(t)|\mathcal{F}_\tau\} \right] dN_i(t)
- \sum_{i=1}^{n} \int_{0}^{\tau} E\{\hat{Z}_i(t)^2|\mathcal{F}_\tau\} Y_i(t) \beta dt + \sum_{i=1}^{n} \int_{0}^{\tau} E\{\bar{Z}(t)\hat{Z}_i^T(t)|\mathcal{F}_\tau\} Y_i(t) \beta dt
= \sum_{i=1}^{n} \int_{0}^{\tau} \{Z_i(t) - \bar{Z}(t)\} dN_i(t) - \sum_{i=1}^{n} \int_{0}^{\tau} \{Z_i^2(t) + \Sigma_1/n_i\} Y_i(t) \beta dt
+ \sum_{i=1}^{n} \int_{0}^{\tau} \left\{ \frac{Y_i(t)Z_i(t)\Sigma_1/n_i + \sum_{j=1}^{n} Y_j(t)Z_j(t)Z_i^T(t)}{\sum_{j=1}^{n} Y_j(t)} \right\} Y_i(t) \beta dt
= \sum_{i=1}^{n} \int_{0}^{\tau} \{Z_i(t) - \bar{Z}(t)\} dN_i(t) - \sum_{i=1}^{n} \int_{0}^{\tau} \{Z_i^2(t) + \Sigma_1/n_i\} Y_i(t) \beta dt
+ \sum_{i=1}^{n} \int_{0}^{\tau} \left\{ \frac{Y_i(t)\Sigma_1/n_i}{\sum_{j=1}^{n} Y_j(t)} \right\} Y_i(t) \beta dt + \sum_{i=1}^{n} \int_{0}^{\tau} \bar{Z}(t)Z_i^T(t) Y_i(t) \beta dt
= U(\beta) - \int_{0}^{\tau} \left\{ 1 - \frac{1}{\sum_{j=1}^{n} Y_j(t)} \right\} \sum_{i=1}^{n} \{Y_i(t)\Sigma_1/n_i\} dt.
\]

Appendix A4: Proof of Theorem 1

Recall that

\[
\hat{\beta}_c = \left[ n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \{\hat{Z}_i(t) - \bar{Z}(t)\}^2 dt - n^{-1} \int_{0}^{\tau} \left\{ 1 - \frac{1}{\sum_{j=1}^{n} Y_j(t)} \right\} \sum_{i=1}^{n} \{Y_i(t)\Sigma_1/n_i\} dt \right]^{-1}
\times \left[ n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} \{\hat{Z}_i(t) - \bar{Z}(t)\} dN_i(t) \right].
\]
Let
\[
D_c = n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ \hat{Z}_i(t) - \tilde{Z}(t) \right\} \otimes^2 dt - n^{-1} \int_{0}^{\tau} \left\{ 1 - \frac{1}{\sum_{j=1}^{n} Y_j(t)} \right\} \sum_{i=1}^{n} \left\{ Y_i(t) \hat{\Sigma}_i / n \right\} dt
\]
denote the denominator of $\hat{\beta}_c$. Let
\[
D_c = E \left[ \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\} \otimes^2 dt \right].
\]

We first prove the following lemmas:

**Lemma A.1** Under Regularity Conditions R1-R8, $D_c$ converges to $D_c$ almost surely as $n \to \infty$.

**Proof:** The proof consists of two parts. In the first part, we examine the asymptotic behavior of the first term of $D_c$ while in the second part we look at the second term of $D_c$.

**Part 1:**
\[
\begin{align*}
n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ \hat{Z}_i(t) - \tilde{Z}(t) \right\} \otimes^2 dt &= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ \hat{Z}_i(t) - e(t) \right\} \otimes^2 dt + n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ e(t) - \tilde{Z}(t) \right\} \otimes^2 dt \\
&+ n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} 2Y_i(t) \left\{ \hat{Z}_i(t) - e(t) \right\} \left\{ e(t) - \tilde{Z}(t) \right\}^T dt \\
&= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ \hat{Z}_i(t) - e(t) \right\} \otimes^2 dt \\
&+ \int_{0}^{\tau} \left( -n^{-1} \sum_{i=1}^{n} Y_i(t) e^{\otimes^2} - n^{-1} \left\{ \sum_{j=1}^{n} Y_j(t) \hat{Z}_j(t) \right\} \otimes^2 - 2n^{-1} \sum_{i=1}^{n} Y_i(t) \hat{Z}_i(t) e^T \right) dt.
\end{align*}
\]

Now we examine the second term of the expression above by applying the Strong Uniform
Law of Large Numbers individually to each term and obtain

\[
\int_0^\tau \left( -n^{-1} \sum_{i=1}^n Y_i(t)e^{\otimes 2}(t) - n^{-1} \frac{\left( \sum_{j=1}^n Y_j(t) \hat{Z}_j(t) \right)^{\otimes 2}}{\sum_{j=1}^n Y_j(t)} + 2n^{-1} \sum_{i=1}^n Y_i(t) \hat{Z}_i(t)e^T(t) \right) dt \\
= \int_0^\tau \left( -E\{Y_i(t)\}e^{\otimes 2}(t) - \frac{\left[ E\{Y_j(t)\} \hat{Z}_j(t) \right]^{\otimes 2}}{E\{Y_j(t)\}} + 2E\{Y_i(t)\} \hat{Z}_i(t)e^T(t) \right) dt + o_{a.s.}(1) \\
= \int_0^\tau \left( -E\{Y_i(t)\}e^{\otimes 2}(t) - \frac{\left[ E\{Y_j(t)Z_j(t)\} \right]^{\otimes 2}}{E\{Y_j(t)\}} + 2E\{Y_i(t)\}Z_i(t)e^T(t) \right) dt + o_{a.s.}(1) \\
= \int_0^\tau \left[ -E\{Y_i(t)\}e^{\otimes 2}(t) - E\{Y_i(t)\}e^{\otimes 2}(t) + 2E\{Y_i(t)\}e^{\otimes 2}(t) \right] dt + o_{a.s.}(1) \\
= o_{a.s.}(1),
\]

where the second last identity comes from the definition of \( e(t) \).
Then it follows by the Strong Uniform Law of Large Numbers that

\[
 n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ \tilde{Z}_i(t) - \tilde{Z}(t) \right\}^2 dt \\
= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ \tilde{Z}_i(t) - e(t) \right\}^2 dt + o_{a.s.}(1) \\
= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^2 dt + n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \epsilon_i^2 dt \\
+ n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} 2Y_i(t) \left\{ Z_i(t) - e(t) \right\} \epsilon_i^2 dt + o_{a.s.}(1) \\
= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^2 dt + n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \epsilon_i^2 dt \\
+ \int_{0}^{\tau} 2E[Y_i(t) \left\{ Z_i(t) - e(t) \right\} \epsilon_i^2] dt + o_{a.s.}(1) \\
= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^2 dt + n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \epsilon_i^2 dt + o_{a.s.}(1) \\
= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^2 dt + \int_{0}^{\tau} \lim_{n \to \infty} \left[ n^{-1} \sum_{i=1}^{n} E\{Y_i(t) \epsilon_i^2 \} \right] dt + o_{a.s.}(1) \\
= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^2 dt + \int_{0}^{\tau} E\{Y_i(t)\} \lim_{n \to \infty} \left\{ n^{-1} \sum_{i=1}^{n} E(\epsilon_i^2) \right\} dt + o_{a.s.}(1) \\
= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^2 dt + \int_{0}^{\tau} E\{Y_i(t)\} \lim_{n \to \infty} \left\{ n^{-1} \sum_{i=1}^{n} \Sigma_i/n_i \right\} dt + o_{a.s.}(1) \\
= n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Y_i(t) \left\{ Z_i(t) - e(t) \right\}^2 dt + \int_{0}^{\tau} \rho_0 E\{Y_i(t)\} \Sigma_1 dt + o_{a.s.}(1),
\]

where the last two steps follow from the definition of \( \Sigma_1 \) and \( \rho_0 \).
Part 2: Now we examine the second term of $D_c$:

$$n^{-1} \int_0^\tau \left\{ 1 - \frac{1}{\sum_{j=1}^n Y_j(t)} \right\} \sum_{i=1}^n \left\{ Y_i(t) \hat{\Sigma}_1/n_i \right\} dt$$

$$= n^{-1} \int_0^\tau \sum_{i=1}^n \left\{ Y_i(t) \hat{\Sigma}_1/n_i \right\} dt - n^{-1} \hat{\Sigma}_1 \int_0^\tau \frac{n^{-1} \sum_{i=1}^n Y_i(t)/n_i dt}{\sum_{i=1}^n Y_i(t)}$$

$$= n^{-1} \int_0^\tau \sum_{i=1}^n \left\{ Y_i(t) \hat{\Sigma}_1/n_i \right\} dt - n^{-1} \hat{\Sigma}_1 \int_0^\tau \lim_{n \to \infty} \frac{n^{-1} \sum_{i=1}^n E(Y_i(t))/n_i}{E\{Y_i(t)\}} dt + o_{a.s.}(1)$$

$$= n^{-1} \int_0^\tau \sum_{i=1}^n \left\{ Y_i(t) \hat{\Sigma}_1/n_i \right\} dt - n^{-1} \hat{\Sigma}_1 \rho_0 \tau + o_{a.s.}(1)$$

$$= n^{-1} \int_0^\tau \sum_{i=1}^n \left\{ Y_i(t) \hat{\Sigma}_1/n_i \right\} dt + o_{a.s.}(1)$$

$$= \int_0^\tau \rho_0 E\{Y_i(t)\} \Sigma_1 dt + o_{a.s.}(1).$$

As a result, combining parts 1 and 2 gives

$$D_c = n^{-1} \sum_{i=1}^n \int_0^\tau Y_i(t) \{ Z_i(t) - e(t) \} \otimes^2 dt + \int_0^\tau \rho_0 E\{Y_i(t)\} \Sigma_1 dt - \int_0^\tau \rho_0 E\{Y_i(t)\} \Sigma_1 dt + o_{a.s.}(1)$$

$$= D_c + o_{a.s.}(1).$$

The proof of Lemma A.1 is now completed.

From the proof of Lemma A.1, we obtain that the inverse matrix in $\hat{\beta}_c$, i.e., $D_c$, converges almost surely to a positive definite matrix under mild regularity conditions. Thus, the estimator $\hat{\beta}_c$ does not have the singularity and unstability issues.

Lemma A.2 Under Regularity Conditions R1-R8, $\hat{\beta}_c$ converges to $\beta$ almost surely as $n \to \infty$.

Proof: Recall that we have proved in Appendix A2 that

$$n^{-1} \sum_{i=1}^n \int_0^\tau \left\{ \tilde{Z}_i(t) - \tilde{Z}(t) \right\} dN_i(t) = E \left[ \int_0^\tau \{ Z_i(t) - e(t) \} dN_i(t) \right] + o_{a.s.}(1).$$
Combined with Lemma A.1, we obtain that

\[
\hat{\beta}_c = D_c^{-1}n^{-1} \sum_{i=1}^{n} \int_0^t \{ \hat{Z}_i(t) - \bar{Z}(t) \} \, dN_i(t)
\]

\[
= D_c^{-1}E \left[ \int_0^t \{ Z_i(t) - e(t) \} \, dN_i(t) \right] + o_{a.s.}(1)
\]

\[
= D_c^{-1}E \left[ \int_0^t \{ Z_i(t) - e(t) \} \, dN_i(t) \right] + o_{a.s.}(1)
\]

\[
= \beta + o_{a.s.}(1).
\]

The proof of Lemma A.2 is now completed.

We now return to the proof of Theorem 1. Note that \( U_c(\beta) = U_1 - U_2 \), where

\[
U_1 = \sum_{i=1}^{n} \int_0^t \{ \hat{Z}_i(t) - \bar{Z}(t) \} \, d\tilde{M}_i(t; \beta, \Lambda_0),
\]

and \( U_2 = \int_0^t \left\{ 1 - \frac{1}{\sum_{j=1}^{n} Y_j(t)} \right\} \sum_{i=1}^{n} \{ Y_i(t)\hat{\Sigma}_1\beta/n_i \} \, dt. \)

We now separately study the asymptotic expansion of \( n^{-1/2}U_1 \) and \( n^{-1/2}U_2 \).

By analogy to the proof of Theorem 1 of Kulich and Lin (2000), we have

\[
n^{-1/2}U_1 = n^{-1/2} \sum_{i=1}^{n} \int_0^t \{ \hat{Z}_i(t) - e(t) \} \, d\tilde{M}_i(t; \beta, \Lambda_0) + o_p(1).
\]

Note that

\[
n^{-1/2}U_2 = n^{-1/2} \int_0^t \sum_{i=1}^{n} \{ Y_i(t)\hat{\Sigma}_1\beta/n_i \} \, dt - n^{-1/2} \int_0^t \frac{1}{\sum_{j=1}^{n} Y_j(t)} \sum_{i=1}^{n} \{ Y_i(t)\hat{\Sigma}_1\beta/n_i \} \, dt
\]

\[
= n^{-1/2} \int_0^t \sum_{i=1}^{n} \{ Y_i(t)\hat{\Sigma}_1\beta/n_i \} \, dt + o_p(1)
\]

\[
= n^{1/2}(\hat{\Sigma}_1 - \Sigma_1)\rho_0 E(S_i)\beta + n^{-1/2} \int_0^t \sum_{i=1}^{n} \{ Y_i(t)\Sigma_1\beta/n_i \} \, dt + o_p(1)
\]

\[
= n^{1/2} \frac{\sum_{i=1}^{n} \sum_{r=1}^{n_i} (W_{ir} - \bar{W}_i)^2}{\sum_{i=1}^{n} (n_i - 1)} \rho_0 E(S_1)\beta - n^{1/2} \sum_{i=1}^{n} S_i \Sigma_1\beta/n_i + o_p(1)
\]

\[
= n^{-1/2} \rho_0 \rho_1^{-1} \sum_{i=1}^{n} \sum_{r=1}^{n_i} (W_{ir} - \bar{W}_i)^2 E(S_i)\beta - n^{1/2} \sum_{i=1}^{n} S_i \Sigma_1\beta/n_i + o_p(1),
\]
where the third identity comes from that

$$n^{1/2} \left[ \int_0^T \sum_{i=1}^n \left\{ \frac{Y_i(t)}{n_i} \right\} dt - \rho_0 E(S_i) \right](\hat{\Sigma}_1 - \Sigma_1)\beta = o_p(1).$$

Therefore,

$$n^{-1/2}U_c(\beta) = n^{-1/2} \sum_{i=1}^n U_{c,i} + o_p(1),$$

where

$$U_{c,i} = \int_0^T \left\{ \hat{Z}_i(t) - e(t) \right\} d\tilde{M}_i(t; \beta, \Lambda_0) + \rho_0 \rho_1^{-1} \sum_{r=1}^{n_i} (W_{ir} - \bar{W}_i)^{\otimes 2} E(S_i)\beta - \rho_0 E(S_i) \Sigma_1 \beta + \frac{S_i \Sigma_1 \beta}{n_i}.$$

By the Taylor series expansion, $0 = n^{-1/2}U_c(\hat{\beta}_c) = n^{-1/2}U_c(\beta) + \left[ n^{-1} \frac{\partial U_c(\beta)}{\partial \beta} \right] n^{1/2}(\hat{\beta}_c - \beta)$, we obtain that

$$n^{1/2}(\hat{\beta}_c - \beta) = - \left[ n^{-1} \frac{\partial U_c(\beta)}{\partial \beta} \right]^{-1} n^{-1/2}U_c(\beta).$$

By the derivation in Appendix A2,

$$-n^{-1} \frac{\partial U_c(\beta)}{\partial \beta} = n^{-1} \sum_{i=1}^n \int_0^T \left\{ \frac{Y_i(t)}{n_i} \right\} dt - n^{-1} \sum_{i=1}^n \int_0^T \left\{ \frac{Y_i(t)\hat{\Sigma}_1/n_i}{n_i} \right\} dt$$

$$\overset{a.s.}{\longrightarrow} \lim_{n \to \infty} n^{-1} E \left[ \sum_{i=1}^n \int_0^T \left\{ \frac{Y_i(t)\hat{\Sigma}_1/n_i}{n_i} \right\} dt \right] - \rho_0 E(S_i) \Sigma_1$$

$$= D_c.$$

By Condition R6, $E(||n^{-1/2}U_{c,i}||^2)I(||n^{-1/2}U_{c,i}|| > \epsilon)$ can only take at most $N_0$ possible values for a given $\epsilon > 0$. Without loss of generality, suppose when $i = 1$, it achieves the maximum value. It follows from the Markov inequality that $Pr\{||n^{-1/2}U_{c,1}|| > \epsilon\} \leq n^{-1} E(||U_{c,1}||^2)/\epsilon^2 \to 0$ as $n \to \infty$, and thus

$$\sum_{i=1}^n E(||n^{-1/2}U_{c,i}||^2)I\{||n^{-1/2}U_{c,i}|| > \epsilon\} \leq E(||U_{c,1}||^2)I\{||n^{-1/2}U_{c,1}|| > \epsilon\} \to 0 \text{ as } n \to \infty,$$

suggesting that the Lindeberg condition (van der Vaart 1998, p.20) is satisfied.

By the multivariate Lindeberg-Feller Central Limit Theorem (van der Vaart 1998, p.20),
we obtain that \( n^{-1/2} U_c(\beta) \) is asymptotically normal with mean 0 and covariance matrix \( \Sigma_c = \lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} E(U_{c,i}) \otimes^2 \). It follows that that \( n^{1/2}(\hat{\beta}_c - \beta) \) is asymptotically normal with mean zero and covariance matrix \( D_c^{-1}\Sigma_c D_c^{-T} \).

**Appendix A5: Proof of Theorem 2**

It can be calculated that

\[
\frac{1}{\sqrt{n}} \left\{ \hat{\Lambda}_0(t; \hat{\beta}_c) - \Lambda_0(t) \right\} = \frac{1}{\sqrt{n}} \int_0^t \sum_{i=1}^{n} \frac{d\tilde{M}_i(u; \beta, \Lambda_0)}{\sum_{j=1}^{n} Y_j(u)} - \left\{ \int_0^t \sum_{i=1}^{n} \frac{Y_i(u)\tilde{Z}_i(u)}{\sum_{j=1}^{n} Y_j(u)} du \right\} \frac{1}{\sqrt{n}}(\hat{\beta}_c - \beta).
\]

By the arguments similar to those in Appendix A2, we obtain that

\[
n^{1/2} \left\{ \sum_{j=1}^{n} Y_j(u) \right\}^{-1} \sum_{i=1}^{n} d\tilde{M}_i(u; \beta, \Lambda_0) \text{ converges weakly to a Gaussian process, and thus is tight.}
\]

\[
\int_0^t \sum_{i=1}^{n} Y_i(u)\tilde{Z}_i(u) du \text{ converges almost surely to } \int_0^t e^T(u) du, \text{ and thus}
\]

\[
\sum_{i=1}^{n} Y_i(u)\tilde{Z}_i(u) / \sum_{j=1}^{n} Y_j(u) du \text{ is tight. Since tightness in } l^\infty[0, \tau] \text{ equipped with the uniform metric has the additivity property, and } n^{1/2}(\hat{\beta}_c - \beta) \text{ is asymptotic normal, we obtain that}
\]

\[
n^{1/2} \left\{ \hat{\Lambda}_0(t; \hat{\beta}_c) - \Lambda_0(t) \right\} \text{ is tight.}
\]

It follows from Appendix A4 that

\[
n^{1/2}(\hat{\beta}_c - \beta) = n^{-1/2} D_c^{-1} \sum_{i=1}^{n} \left[ \int_0^t \{ \tilde{Z}_i(t) - e(t) \} d\tilde{M}_i(t; \beta, \Lambda_0) \right.
\]

\[+ \rho_0 \rho_1^{-1} \sum_{r=1}^{n} (W_{ir} - \bar{W}_{ir}) \otimes^2 E(S_i)\beta - \rho_0 E(S_i) \Sigma_1 \beta + \frac{S_i \Sigma_1 \beta}{n_i} \bigg] + o_p(1).
\]

Thus, \( n^{1/2} \left\{ \hat{\Lambda}_0(t; \hat{\beta}_c) - \Lambda_0(t) \right\} \) is asymptotically equivalent to \( n^{1/2} \sum_{i=1}^{n} \Psi_i(t) \) uniformly in \( t \).

Similar to the proof of the asymptotic normality of \( \hat{\beta}_c \), it is shown that \( n^{-1/2} \sum_{i=1}^{n} \Psi_i(t) \) satisfies the Lindeberg condition, and thus the multivariate Lindeberg-Feller Central Limit Theorem applies to \( n^{1/2} \left\{ \hat{\Lambda}_0(t; \hat{\beta}_c) - \Lambda_0(t) \right\} \). Together with the fact that \( n^{1/2} \left\{ \hat{\Lambda}_0(t; \hat{\beta}_c) - \Lambda_0(t) \right\} \) is tight, the weak convergence result is proved.
Appendix A6: Asymptotic properties of $\hat{\beta}_c$ and $\hat{\Lambda}_0(t; \hat{\beta}_c)$ when $\Sigma_1$ is simply known

Now, we investigate the asymptotic property of $\hat{\beta}_c$ when $\Sigma_1$ is simply known. In this case, note that

$$n^{-1/2}U_2 = n^{-1/2} \int_0^r \left\{ \sum_{i=1}^n Y_i(t) \Sigma_1 \beta / n_i \right\} dt$$

$$= n^{-1/2} \sum_{i=1}^n S_i \Sigma_1 \beta / n_i.$$

It then follows from the proof of Theorem 1 that the asymptotic distribution of $\hat{\beta}_c$ is almost identical as the one in Theorem 1, with the only difference that $\Sigma_2 \beta - \rho_0 E(S_i) \Sigma_1 \beta$ is removed from $\Sigma_c$.

Now, we investigate the asymptotic property of $\hat{\Lambda}_0(t; \hat{\beta}_c)$ when $\Sigma_1$ is simply known. In this case, note that

$$n^{1/2}(\hat{\beta}_c - \beta) = n^{-1/2} D_c^{-1} \sum_{i=1}^n \left[ \int_0^r \left\{ \hat{Z}_i(t) - e(t) \right\} d\tilde{M}_i(t; \beta, \Lambda_0) + \frac{S_i \Sigma_1 \beta}{n_i} \right] + o_p(1).$$

It then follows from the proof of Theorem 2 that the $n^{1/2}\{\hat{\Lambda}_0(t; \hat{\beta}_c) - \Lambda_0(t)\}$ is tight, and the limiting process is almost identical as the one in Theorem 2, with the only difference that $\Sigma_2 \beta - \rho_0 E(S_i) \Sigma_1 \beta$ is removed from $\Psi_i(t)$.

Appendix A7: Proof of Theorem 3

Let $\tilde{W}_n(t) = n^{-1/2} \sum_{i=1}^n \xi_i \Psi_i(t)$. By the proof of the weak convergence of $n^{1/2}\{\hat{\Lambda}_0(t; \hat{\beta}_c) - \Lambda_0(t)\}$, $n^{-1/2} \sum_{i=1}^n \Psi_i(t)$ converges weakly to $G(t)$ unconditionally. Since weak convergence of $n^{-1/2} \sum_{i=1}^n \Psi_i(t)$ implies that the Donsker condition (van der Vaart and Wellner 1996, Theorem 2.9.6) holds, it then follows from the conditional multiplier Central Limit Theorem (van der Vaart and Wellner 1996, Sec. 2.9) that $\tilde{W}_n(t)$ converges weakly to $G(t)$ in probability conditional on the data. Thus, by Lemma 1 of Pipper and Ritz (2007), it suffices to show that
\[ \sup_{t \in [0, \tau]} |\hat{W}_n(t) - \tilde{W}_n(t)| \xrightarrow{p} 0. \]

Let \( \hat{M}_i(t) \) be the empirical version of \( \tilde{M}_i(t) \), \( i = 1, \ldots, n \). Note that

\[ \sup_{t \in [0, \tau]} |\hat{W}_n(t) - \tilde{W}_n(t)| \leq \sup_{t \in [0, \tau]} |W_n^{(1)}(t)| + \sup_{t \in [0, \tau]} |W_n^{(2)}(t)| + \sup_{t \in [0, \tau]} |W_n^{(3)}(t)|, \]

where

\[ W_n^{(1)}(t) = n^{-1/2} \sum_{i=1}^{n} \xi_i \left\{ \int_{0}^{t} \frac{d\hat{M}_i(u; \beta, \Lambda_0)}{n^{-1} \sum_{i=1}^{n} Y_i(u)} - \int_{0}^{t} \frac{d\tilde{M}_i(u; \beta, \Lambda_0)}{E[Y_i(u)]} \right\}, \]

\[ W_n^{(2)}(t) = \int_{0}^{t} \tilde{Z}^T(u) du \times \hat{D}_c^{-1} \times n^{-1/2} \sum_{i=1}^{n} \xi_i \left[ \int_{0}^{r} \{ \hat{Z}_i(t) - \tilde{Z}(t) \} d\hat{M}_i(t; \beta, \Lambda_0) \right] \]

\[ - \int_{0}^{t} e^T(u) du \times D_c^{-1} \times n^{-1/2} \sum_{i=1}^{n} \xi_i \left[ \int_{0}^{r} \{ \hat{Z}_i(t) - e(t) \} d\tilde{M}_i(t; \beta, \Lambda_0) \right], \]

and

\[ W_n^{(3)}(t) = \int_{0}^{t} \tilde{Z}^T(u) du \times \hat{D}_c^{-1} \times n^{-1/2} \sum_{i=1}^{n} \xi_i \hat{S}_i \hat{\Sigma}_1 \hat{\beta}_c \]

\[ - \int_{0}^{t} e^T(u) du \times D_c^{-1} \times n^{-1/2} \sum_{i=1}^{n} \frac{\xi_i S_i \Sigma_1 \beta_0}{n_i}. \]

Employing the empirical process techniques used in Appendix A2, together with Lemma A.3 and Theorem 2 of Spiekerman and Lin (1998), we show that

\[ \sup_{t \in [0, \tau]} |W_n^{(j)}(t)| \xrightarrow{p} 0, j = 1, 2, 3, \]

and thus \( \sup_{t \in [0, \tau]} |\hat{W}_n(t) - \tilde{W}_n(t)| \xrightarrow{p} 0 \) holds.

**Appendix A8: Proof of Corrolary 1**

Note that the only difference between \( \hat{\beta}_c \) and \( \tilde{\beta}_c \) is the term

\[ \int_{0}^{r} \left\{ \frac{1}{\sum_{j=1}^{n} Y_j(t)} \right\} \sum_{i=1}^{n} \left\{ Y_i(t) \hat{S}_i \Sigma_1 \beta / n_i \right\} dt \]
in their denominators. By the arguments in Appendix A2, we have

\[ n^{-1/2} \int_0^T \left\{ \frac{1}{\sum_{j=1}^n Y_j(t)} \right\} \sum_{i=1}^n \left\{ Y_i(t) \hat{\Sigma}_1 \beta/n_i \right\} dt = o_p(1). \]

Therefore, \( \hat{\beta}_e \) is asymptotically identical to \( \hat{\beta}_c \), suggesting that the asymptotic normal distribution of \( \hat{\beta}_e \) is identical to that of \( \hat{\beta}_c \). Similarly, the limit process of \( n^{1/2} \left\{ \hat{\Lambda}_0(t; \hat{\beta}_e) - \Lambda_0(t) \right\} \) is identical to that of \( n^{1/2} \left\{ \hat{\Lambda}_0(t; \hat{\beta}_c) - \Lambda_0(t) \right\} \).

**Appendix A9: Generation of Survival Times**

Generating survival times under the Cox model is well studied, see Bender, Augustin, and Blettner (2005), for instance. However, generating survival times under additive hazards models is rarely discussed in the literature. In this Appendix, we describe how we generate survival times in the simulations from the additive hazards model:

\[ \lambda(t; Z) = \lambda_0(t) + \beta^T Z, \]

where \( \lambda_0(t) \) is the baseline hazard function, \( Z \) is the covariate vector, and \( \beta \) is the regression vector. In the simulations, we considered two cases of \( \lambda_0(t) \): \( \lambda_0(t) = 1 \) or \( t \). It is straightforward to generate survival times when \( \lambda_0(t) = 1 \), because the survival time \( T \) follows the exponential distribution in this case.

Here we discuss how survival times are generated for the case with \( \lambda_0(t) = t \). Let \( F(t) \) be the cumulative distribution function of \( T \). Note that in this case, \( \Pr(T > t) = \exp(-t^2/2 - \beta^T Z t) \), and thus

\[ F(t) = \Pr(T \leq t) = 1 - \exp(-t^2/2 - \beta^T Z t). \]

Define \( U = F(T) \). Then \( U \sim UNIF(0,1) \), and

\[ U = 1 - \exp(-T^2/2 - \beta^T Z T). \quad \text{(A3)} \]
It is easily seen that (A3) gives

\[(T + \beta^T Z)^2 = (\beta^T Z)^2 - 2 \log(1 - U).\]

Noting that \((\beta^T Z)^2 - 2 \log(1 - U) > 0\), we obtain that

\[T = \sqrt{(\beta^T Z)^2 - 2 \log(1 - U)} - \beta^T Z.\] (A4)

Therefore, to generate survival time \(T\) under the scenario with \(\lambda_0(t) = t\), we follow the following three steps. In the first step, we specify a value of \(\beta\). In the second step, we generate a random number \(U\) from the uniform distribution \(UNIF(0, 1)\), and generate the covariate \(Z\) from a pre-specified distribution. In the third step, we apply equation (A4) to obtain a survival time \(T\).

**Appendix A10: Additional Simulations**

We provide two additional simulation results summarized in Table A1 and Table A2.

In Table A1, we consider a scenario where the regression calibration estimator does not perform well, compared to the consistent estimators. We generate the data as in Scenario 3 in Section 6.1, except that we set \(\beta_x = 5\), and \(C = 1\). The censoring rate is about 30%. It is clear from Table A1 that the regression calibration estimator of \(\beta_x\) is seriously biased and its coverage probability deviates from the nominal 95% level significantly when measurement error is moderate. In contrast, the proposed consistent estimators perform satisfactorily.

In Table A2, we consider a scenario where we generate covariates \(X_i\) and \(V_i\) from normal distributions, and one of the regression parameter is negative. Specifically, we generate \(Z_i =\)
We set the true parameter as \( \beta = (0.5, -0.5)^T \) and \( \lambda_0(t) = 1 \). We generate \( C_i \) from a uniform distribution so that the censoring rate is about 30\%. The measurement error model is the same as the one in Table A1. We consider \( n = 200 \) and 1000 simulation runs for each parameter configuration. The case that \( 1 + \beta^T Z_i \leq 0 \) occurs in a small fraction (about 2\% \( \sim \) 3\%) among the 200 subjects in each simulation run. Due to that \( \lambda(t; Z_i) = 1 + \beta^T Z_i \) needs to be positive, we drop this small fraction. To maintain the sample size \( n = 200 \), we replace this fraction of data with additionally generated data that satisfies \( 1 + \beta^T Z_i > 0 \). The performance of the methods are similar to those in Section 6.

**Appendix A11: Proof of Lemma 1**

Using the proof of Theorem 1 and SULLN, we can show that

\[
n^{-1/2} U_c(\beta, t) = n^{-1/2} \sum_{i=1}^{n} A_i(t) + o_p(1)
\]

uniformly in \( t \). Note that by the derivations in Appendix A4,

\[
n^{1/2}(\beta_c - \beta) = D_c^{-1}(\tau)n^{-1/2} \sum_{i=1}^{n} A_i(\tau) + o_p(1).
\]

Thus, applying the Taylor series expansion gives

\[
n^{-1/2} U_c(\beta, t) = n^{-1/2} U_c(\beta, t) + \left[ n^{-1} \frac{\partial U_c(\beta, t)}{\partial \beta} \right] n^{1/2}(\beta_c - \beta)
\]

\[
= n^{-1/2} \sum_{i=1}^{n} A_i(t) - D_c(t)D_c^{-1}(\tau)n^{-1/2} \sum_{i=1}^{n} A_i(\tau) + o_p(1)
\]

\[
= n^{-1/2} \sum_{i=1}^{n} \left\{ A_i(t) - D_c(t)D_c^{-1}(\tau)A_i(\tau) \right\} + o_p(1)
\]
uniformly in $t$. The tightness of $n^{-1/2} \sum_{i=1}^{n} \{A_i(t) - D_c(t)D_c^{-1}(\tau)A_i(\tau)\}$ can be shown using the arguments similar to those in the proof of Theorem 2. The proof then follows.

**Appendix A12: Proof of Theorem 4**

The arguments are similar to those in Appendix A7 by adopting the conditional multiplier Central Limit Theorem, and thus omitted.

**Reference**

Andersen, P. K., and Gill, R. D. (1994), Cox’s Regression Model for Counting Processes: A Large Sample Study, *The Annals of Statistics*, 10, 1100-1120.

Bender, R., Augustin, T., and Blettner, M. (2005), Generating Survival Times to Simulate Cox Proportional Hazards Models, *Statistics in Medicine*, 24, 1713-1723.

Bilias, Y., Gu, M., and Ying, Z. (1997), Towards a General Asymptotic Theory for Cox Model with Staggered Entry, *Annals of Statistics*, 25, 662-682.

Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M. (2006), *Measurement Error in Nonlinear Models: A Modern Perspective*, 2nd ed., Chapman & Hall/CRC, Boca Raton.

Hu, C., and Lin, D. Y. (2004), Semiparametric Failure Time Regression with Replicates of Mismeasured Covariates, *Journal of the American Statistical Association*, 99, 105-118.

Kulich, M., and Lin, D. Y. (2000), Additive Hazards Regression with Covariate Measurement Error, *Journal of the American Statistical Association*, 95, 238-248.

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: Series B*, 62, 711-730.

Pipper, C. B., and Ritz, C. (2007), Checking the Grouped Data Version of the Cox Model for Interval-Grouped Survival Data, *Scandinavian Journal of Statistics*, 34, 405-418.

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

Spiekerman, C. F., and Lin, D. Y. (1998), Marginal Regression Models for Multivariate Failure Time Data, *Journal of the American Statistical Association*, 93, 1164-1175.

van der Vaart, A. W. (1998), *Asymptotic Statistics*, Cambridge University Press, New York.

van der Vaart, A. W., and Wellner, J. A. (1996), *Weak Convergence and Empirical Process*, Springer-Verlag, New York.
Table A1: Comparison of the performance of various estimators when $\beta_x = 5$

| $\sigma$ | Method | Estimating $\beta_x$ | Bias | EVE | MVE | MSE | MCP(%) | Estimating $\beta_v$ | Bias | EVE | MVE | MSE | MCP(%) |
|----------|--------|----------------------|------|-----|-----|-----|--------|----------------------|------|-----|-----|-----|--------|
| 0.25     | $\hat{\beta}_{nv}$ | -0.291 0.638 0.579 0.723 89.3 | 0.068 0.324 0.326 0.328 95.8 |
|          | $\hat{\beta}_{szs}$ | 0.135 0.856 0.764 0.874 92.3 | -0.043 0.344 0.348 0.346 95.3 |
|          | $\hat{\beta}_{rc}$  | -0.126 0.682 0.620 0.697 91.7 | -0.030 0.326 0.330 0.327 95.3 |
|          | $\hat{\beta}_c$      | 0.125 0.845 0.756 0.859 92.4 | -0.040 0.343 0.348 0.344 95.4 |
|          | $\hat{\beta}_e$      | 0.136 0.851 0.762 0.869 92.3 | -0.043 0.343 0.348 0.345 95.3 |
| 0.75     | $\hat{\beta}_{nv}$ | -2.105 0.246 0.243 4.675 4.3 | 0.549 0.374 0.370 0.675 84.7 |
|          | $\hat{\beta}_{szs}$ | 0.321 1.879 2.202 1.980 95.6 | -0.126 0.853 0.835 0.868 96.2 |
|          | $\hat{\beta}_{rc}$  | -1.187 0.427 0.424 1.835 50.6 | 0.000 0.420 0.391 0.419 94.3 |
|          | $\hat{\beta}_c$      | 0.257 1.704 1.905 1.769 95.1 | -0.082 0.730 0.756 0.736 95.9 |
|          | $\hat{\beta}_e$      | 0.341 1.743 2.046 1.858 95.8 | -0.111 0.760 0.787 0.771 96.2 |

* Bias: finite sample biases;  
  * Eve: empirical variances;  
  * Mve: average of the model-based variance estimates;  
  * dMse: mean square errors;  
  * MCP: model-based coverage probability.

Table A2: Comparison of the performance of various estimators for normally distributed covariates

| $\sigma$ | Method | Estimating $\beta_x$ | Bias | EVE | MVE | MSE | MCP(%) | Estimating $\beta_v$ | Bias | EVE | MVE | MSE | MCP(%) |
|----------|--------|----------------------|------|-----|-----|-----|--------|----------------------|------|-----|-----|-----|--------|
| 0.25     | $\hat{\beta}_{nv}$ | -0.014 0.010 0.010 0.010 92.7 | 0.007 0.010 0.010 0.010 93.8 |
|          | $\hat{\beta}_{szs}$ | 0.011 0.012 0.011 0.012 93.6 | -0.008 0.011 0.011 0.011 94.2 |
|          | $\hat{\beta}_{rc}$  | 0.009 0.011 0.011 0.011 93.1 | -0.005 0.011 0.010 0.011 94.3 |
|          | $\hat{\beta}_c$      | 0.011 0.011 0.011 0.012 93.4 | -0.007 0.011 0.011 0.011 94.4 |
|          | $\hat{\beta}_e$      | 0.012 0.012 0.011 0.012 93.5 | -0.008 0.011 0.011 0.011 94.4 |
| 0.75     | $\hat{\beta}_{nv}$ | -0.146 0.007 0.007 0.029 53.8 | 0.083 0.010 0.009 0.016 83.4 |
|          | $\hat{\beta}_{szs}$ | 0.029 0.022 0.021 0.022 95.0 | -0.018 0.016 0.015 0.016 94.1 |
|          | $\hat{\beta}_{rc}$  | 0.004 0.016 0.014 0.016 92.8 | 0.001 0.012 0.011 0.012 93.1 |
|          | $\hat{\beta}_c$      | 0.022 0.020 0.019 0.021 94.9 | -0.014 0.015 0.014 0.015 94.1 |
|          | $\hat{\beta}_e$      | 0.028 0.021 0.020 0.022 94.9 | -0.018 0.015 0.015 0.016 94.3 |