Semiparametric Inference for the proportional mean residual life model with right-censored length-biased data

Fangfang Bai[1], Jian Huang[2,3] and Yong Zhou[3,4]

1. University of International Business and Economics,
2. University of Iowa
3. Shanghai University of Finance and Economics and 4. Chinese Academy of Sciences

Supplementary Material

Here, we give the detailed and complete proof of the main results in the paper.

S1 Proof of Theorem 1.

Based on the closed form estimator of $m_0(t)$,

$$\begin{aligned}\widehat{m}_0(t,\beta) &= \frac{\sum_{i=1}^n \delta_i I(Y_i>t)\{Y_i\widehat{S}_C(Y_i-A_i)\}^{-1}(Y_i-t)}{\sum_{i=1}^n \delta_i I(Y_i>t)\{Y_i\widehat{S}_C(Y_i-A_i)\}^{-1}\exp(\beta^T X_i)}\\ &= \Big(\sum_{i=1}^n \frac{\delta_i I(Y_i>t)(Y_i-t)}{Y_i S_C(Y_i-A_i)} + \sum \frac{\delta_i I(Y_i>t)(Y_i-t)}{Y_i\widehat{S}_C S_C}\big(S_C\widehat{S}_C(Y_i-A_i)\big)\Big)\\ &/ \ \Big(\sum_{i=1}^n \frac{\delta_i I(Y_i>t)\exp(\beta^T X_i)}{Y_i S_C(Y_i-A_i)} + \sum \frac{\delta_i I(Y_i>t)\exp(\beta^T X_i)}{Y_i\widehat{S}_C S_C}(S_C(Y_i-A_i)\\ &\quad -\widehat{S}_C(Y_i-A_i))\Big).\end{aligned}$$

Since the following processes:

$$\frac{1}{n}\sum_{i=1}^n \frac{\delta_i I(Y_i>t)}{Y_i S_C(Y_i-A_i)}(Y_i-t) \tag{S1.1}$$

$$\frac{1}{n}\sum_{i=1}^n \frac{\delta_i I(Y_i>t)}{Y_i S_C(Y_i-A_i)}\exp(\beta^T X_i) \tag{S1.2}$$

can be written as the product of monotone function in t and β. And monotone functions have pseudodimension 1(Pollard (1990), page15; Bilias, Gu & Ying (1997), Lemma A.2). Thus by Pollard (1990), page38 and Bilisa , Gu & Ying (1997), Lemma A.1 the processes (S1.1),(S1.2) are manageable. Together with the uniformly consistency of $\widehat{S}_C(t)$ to $S(t)$ in $t\in[0,\tau]$(Flemming & Harrington (1990)), we can obtain that $\widehat{m}_0(t,\beta)$ converges almost surely and uniformly to $m_0(t,\beta)$ in $t\in[0,\tau]$ and $\beta\in\mathcal{B}=\{\beta:||\beta-\beta_0||\le\varepsilon\}$. Here, $m_0^*(t)=m_0(t,\beta_0)$.

Therefore, in order to prove the existence and uniqueness of $\widehat{\beta}_1$ and $\widehat{m}_0(t)$, it suffices to show that there exist a unique solution to $U(\beta) = 0$.

Since $\widehat{m}_0(t, \beta)$ satisfies:

$$\sum_{i=1}^{n} \frac{\delta_i I(Y_i > t)}{Y_i \widehat{S}_C(Y_i - A_i)} \Big[(Y_i - t) - \widehat{m}_0(t, \beta) \exp(\beta^T X_i) \Big] = 0,$$

differentiating it with respect to β, we derive the following equation:

$$\frac{\partial \widehat{m}_0(t, \beta)}{\partial \beta} = - \frac{\sum_{i=1}^{n} \delta_i I(Y_i > t) \{Y_i \widehat{S}_C(Y_i - A_i)\}^{-1} \exp(\beta^T X_i) X_i}{\sum_{i=1}^{n} \delta_i I(Y_i > t) \{Y_i \widehat{S}_C(Y_i - A_i)\}^{-1} \exp(\beta^T X_i)} \widehat{m}_0(t, \beta). \tag{S1.3}$$

Let $\widehat{A}(\beta) = -\frac{1}{n} \frac{\partial U(\beta)}{\partial \beta^T}$. Making use of (S1.3),

$$\begin{aligned} \widehat{A}(\beta) &= \frac{1}{n} \sum_{i=1}^{n} \int_0^{\tau} \frac{\delta_i I(Y_i > t)}{Y_i \widehat{S}_C(Y_i - A_i)} X_i \Big[\widehat{m}_0(t, \beta) \exp(\beta^T X_i) X_i^T + \frac{\partial \widehat{m}_0(t, \beta)}{\partial \beta^T} \exp(\beta^T X_i) \Big] dH(t) \\ &= \frac{1}{n} \sum_{i=1}^{n} \int_0^{\tau} \frac{\delta_i I(Y_i > t)}{Y_i \widehat{S}_C(Y_i - A_i)} \Big[X_i^T - \frac{\sum_{i=1}^{n} \delta_i I(Y_i > t) \{Y_i \widehat{S}_C(Y_i - A_i)\}^{-1} \exp(\beta^T X_i) X_i^T}{\sum_{i=1}^{n} \delta_i I(Y_i > t) \{Y_i \widehat{S}_C(Y_i - A_i)\}^{-1} \exp(\beta^T X_i)} \Big] \\ &\quad X_i \widehat{m}_0(t, \beta) \exp(\beta^T X_i) dH(t) \\ &= \frac{1}{n} \sum_{i=1}^{n} \int_0^{\tau} \frac{\delta_i I(Y_i > t)}{Y_i \widehat{S}_C(Y_i - A_i)} \big(X_i - \overline{X}(t, \beta) \big)^{\otimes 2} \widehat{m}_0(t, \beta) \exp(\beta^T X_i) dH(t), \end{aligned}$$

where

$$\overline{X}(t, \beta) = \frac{\sum_{i=1}^{n} \delta_i I(Y_i > t) \{Y_i \widehat{S}_C(Y_i - A_i)\}^{-1} \exp(\beta^T X_i) X_i}{\sum_{i=1}^{n} \delta_i I(Y_i > t) \{Y_i \widehat{S}_C(Y_i - A_i)\}^{-1} \exp(\beta^T X_i)},$$

which is always nonnegative definite. It follows from the expression and some methods used aforementioned, $\overline{X}(t, \beta)$ converges to some nonrandom function $\overline{x}(t, \beta)$ uniformly in $t \in [0, \tau]$. Together with the uniformly convergence of $\widehat{m}_0(t, \beta)$ to $m_0(t, \beta)$ in $t \in [0, \tau]$ and $\beta \in \mathcal{B} = \{\beta : ||\beta - \beta_0|| \leq \varepsilon\}$, we conclude that $\widehat{A}(\beta)$ converges uniformly to a nonrandom function $A(\beta)$ uniformly in $\beta \in \mathcal{B} = \{\beta : ||\beta - \beta_0|| \leq \varepsilon\}$. Denote $B = A(\beta_0)$ and $S_{\widetilde{T}}(t|X)$ is the the survival function of $\widetilde{T}$, where

$$A(\beta) = E \Big[\int_0^{\tau} \mu^{-1}(X) S_{\widetilde{T}}(t|X) (X - \overline{x}(t, \beta))^{\otimes 2} m_0(t, \beta) \exp(\beta^T X) dh(t) \Big].$$

It can be checked easily that $\frac{1}{n} U(\beta_0)$ converges to 0 almost surely. By condition (A4), A is nonsingular. On the other hand, $\widehat{A}(\beta)$ converges uniformly to a nonrandom function $A(\beta)$ uniformly in $\beta \in \mathcal{B} = \{\beta : ||\beta - \beta_0|| \leq \varepsilon\}$, thus there exists a small neighborhood of β_0 in which $\widehat{A}(\beta)$, especially $\widehat{A}(\beta_0)$ is nonsingular for sufficient large n. Therefore it follows from the inverse function theorem(Rudin (1976)) that within a small neighborhood of β_0, there exists a unique solution $\widehat{\beta}$ to the equation $U(\beta) = 0$ when n is large enough. Furthermore by the nonnegative definiteness of $\widehat{A}(\beta)$ in the entire domain of β, the solution $\widehat{\beta}$ is global uniqueness. Hence, there exists a unique estimator $\widehat{\beta}$ and $\widehat{m}_0(t)$, for $t \in [0, \tau]$. Following the proof of the estimator's uniqueness, we can see that $\widehat{\beta}$ is actually strong consistent and then $\widehat{m}_0(t) = \widehat{m}_0(t, \widehat{\beta})$ converges uniformly to $m_0(t)$ almost surely in $t \in [0, \tau]$. □

S2 Proof of Theorem 2.

(1) By the expression of $\widehat{m}_0(t,\beta)$, we have

$$\widehat{m}_0(t,\beta) - m_0(t) = \frac{\sum_{i=1}^n \delta_i I(Y_i > t)\{Y_i \widehat{S}_C(Y_i - A_i)\}^{-1}\big[(Y_i - t) - m_0(t)\exp(\beta_0^T X_i)\big]}{\sum_{i=1}^n \delta_i I(Y_i > t)\{Y_i \widehat{S}_C(Y_i - A_i)\}^{-1}\exp(\beta^T X_i)}.$$

Note that $M_i(t) = \frac{\delta_i I(Y_i > t)}{Y_i S_C(Y_i - A_i)}\big[(Y_i - t) - m_0(t)\exp(\beta_0^T X_i)\big]$.

Since

$$\begin{aligned}
U(\beta_0) &= \sum_{i=1}^n \int_0^\tau \frac{\delta_i I(Y_i > t) X_i}{Y_i \widehat{S}_C(Y_i - A_i)}\big[(Y_i - t) - m_0(t)\exp(\beta_0^T X_i) - \exp(\beta_0^T X_i)(\widehat{m}_0(t,\beta_0) \\
&\quad - m_0(t))\big] dH(t) \\
&= \sum_{i=1}^n \int_0^\tau \frac{\delta_i I(Y_i > t)}{Y_i \widehat{S}_C(Y_i - A_i)}\big[(Y_i - t) - m_0(t)\exp(\beta_0^T X_i)\big](X_i \\
&\quad - \overline{X}^*(t)) dH(t) \\
&= \sum_{i=1}^n \int_0^\tau \frac{\delta_i I(Y_i > t)}{Y_i S_C(Y_i - A_i)} \frac{S_C(Y_i - A_i)}{\widehat{S}_C(Y_i - A_i)}\big[(Y_i - t) - m_0(t)\exp(\beta_0^T X_i)\big] \\
&\quad (X_i - \overline{X}^*(t)) dH(t) \\
&= \sum_{i=1}^n \int_0^\tau M_i(t) \frac{S_C(Y_i - A_i)}{\widehat{S}_C(Y_i - A_i)}\{X_i - \overline{x}^*(t)\} dh(t) \\
&\quad + \sum_{i=1}^n \int_0^\tau M_i(t) \frac{S_C(Y_i - A_i)}{\widehat{S}_C(Y_i - A_i)}\{X_i - \overline{x}^*(t)\} d(H(t) - h(t)) \\
&\quad - \sum_{i=1}^n \int_0^\tau M_i(t) \frac{S_C(Y_i - A_i)}{\widehat{S}_C(Y_i - A_i)}\{\overline{X}^*(t) - \overline{x}^*(t)\} dH(t), \qquad \text{(S2.4)}
\end{aligned}$$

where $\overline{X}^*(t) = \frac{\sum \delta_i I(Y_i > t)\{Y_i \widehat{S}_C(Y_i - A_i)\}^{-1}\exp(\beta_0^T X_i) X_i}{\sum \delta_i I(Y_i > t)\{Y_i \widehat{S}_C(Y_i - A_i)\}^{-1}\exp(\beta_0^T X_i)}$ and $\overline{x}^*(t)$ is the limit of $\overline{X}^*(t)$.

On the other hand,

$$\sqrt{n}\frac{\widehat{S}_C(t) - S_C(t)}{S_C(t)} = -\frac{1}{\sqrt{n}}\sum_{i=1}^n \int_0^t \frac{1}{\pi(s)} dM_i^C(s) + o_p(1),$$

where $M_i^C(t) = I(Y_i - A_i \le t, \delta_i = 0) - \int_0^t I(Y_i - A_i \ge s) d\Lambda_C(s)$ is the martingale for the censoring variable, $\Lambda_C(t)$ is the cumulative hazard function, and $\pi(t) = P(Y - A \ge t)$, so we can obtain $\sup_{t\in[0,\tau]} \big|\frac{\widehat{S}_C(t) - S_C(t)}{S_C(t)}\big| = O_p(n^{-1/2})$ and $\sum_{i=1}^n M_i(t) = O_p(n^{1/2})$. Therefore, it is easy to show that

$$\begin{aligned}
&\Big|\sum_{i=1}^n \int_0^\tau M_i(t)\Big(\frac{S_C(Y_i - A_i)}{\widehat{S}_C(Y_i - A_i)} - 1\Big)\{\overline{X}^*(t) - \overline{x}^*(t)\} dH(t) \\
&+ \sum_{i=1}^n M_i(t)\{\overline{X}^*(t) - \overline{x}^*(t)\} dH(t)\Big| = o_p(n^{1/2}),
\end{aligned}$$

and

$$\Big|\sum_{i=1}^{n}\int_0^\tau M_i(t)\frac{S_C(Y_i-A_i)}{\widehat{S}_C(Y_i-A_i)}\{X_i-\overline{x}^*(t)\}d(H(t)-h(t))\Big| = o_p(n^{1/2}).$$

Hence,

$$\begin{aligned}(S2.4) \quad &= \sum_{i=1}^{n}\int_0^\tau M_i(t)\{X_i-\overline{x}^*(t)\}dh(t)+\sum_{i=1}^{n}\int_0^\tau M_i(t)\frac{\widehat{S}_C(Y_i-A_i)-S_C(Y_i-A_i)}{S_C(Y_i-A_i)}\\ &\quad \{X_i-\overline{x}^*(t)\}dh(t)+o_p(n^{1/2}).\end{aligned}$$

Based on the martingale representation of $\widehat{S}_C(t)$,

$$\begin{aligned}&\sum_{i=1}^{n}\int_0^\tau M_i(t)\frac{\widehat{S}_C(Y_i-A_i)-S_C(Y_i-A_i)}{S_C(Y_i-A_i)}\{X_i-\overline{x}^*(t)\}dh(t)\\ =& \sum_{i=1}^{n}\int_0^\tau M_i(t)\{X_i-\overline{x}^*(t)\}dh(t)\int_0^{Y_i-A_i}\frac{1}{n\pi(s)}\sum_{j=1}^{n}dM_i^C(s)+o_p(n^{1/2})\\ =& \sum_{i=1}^{n}\int_0^\tau \frac{Q(t)}{\pi(t)}dM_i^C(t)+o_p(n^{1/2}),\end{aligned}$$

where $Q(t)=\frac{1}{n}\sum_{i=1}^n\int_0^\tau M_i(s)\{X_i-\overline{x}^*(s)\}dh(s)I(Y_i-A_i\geq t)$ and $Q(t)$ converges to some nonrandom process $q(t)$.

So by Lemma 1 in Lin et al.(2000), we obtain

$$\frac{1}{\sqrt{n}}U(\beta_0)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i+o_p(1),$$

where $\xi_i=\int_0^\tau M_i(t)\{X_i-\overline{x}^*\}dh(t)+\int_0^\tau\frac{q(t)}{\pi(t)}dM_i^C(t)$. It follows from the multivariate central limit theorem that $n^{-1/2}U(\beta_0)$ is asymptotically normal with mean 0 and covariance matrix $\Sigma=E[\xi_i^{\otimes 2}]$. By the Taylor expansion of $U(\beta)$ at β_0, we have

$$\begin{aligned}n^{1/2}(\widehat{\beta}-\beta_0) &= B^{-1}n^{-1/2}U(\beta_0)+o_p(1)\\ &= B^{-1}n^{-1/2}\sum_{i=1}^{n}\xi_i+o_p(1).\end{aligned}$$

Therefore, $\sqrt{n}(\widehat{\beta}-\beta_0)$ is asymptotically zero-mean normal with covariance matrix $B^{-1}\Sigma B^{-1}$, which can be consistently estimated by $\widehat{B}^{-1}\widehat{\Sigma}\widehat{B}^{-1}$, where

$$\widehat{B}=\frac{1}{n}\sum_{i=1}^{n}\int_0^\tau\frac{\delta_i I(Y_i>t)}{Y_i\widehat{S}_C(Y_i-A_i)}\big(X_i-\overline{X}(t,\widehat{\beta})\big)^{\otimes 2}\widehat{m}_0(t,\widehat{\beta})\exp(\widehat{\beta}^TX_i)dH(t),$$

$$\widehat{\Sigma}=\frac{1}{n}\sum_{i=1}^{n}\widehat{\xi}_i^{\otimes 2},$$

and

$$\widehat{\xi}_i=\int_0^\tau\widehat{M}_i(t)\{X_i-\overline{X}(t,\widehat{\beta})\}dH(t)+\int_0^\tau\frac{\widehat{Q}(t)}{\widehat{\pi}(t)}d\widehat{M}_i^C(t).$$

(2) We begin by showing the weak convergence of $\widehat{m}_0(t)$. Note that

$$\begin{aligned}\sqrt{n}(\widehat{m}_0(t)-m_0(t)) &= \sqrt{n}(\widehat{m}_0(t,\widehat{\beta})-m_0(t)) \\ &= \sqrt{n}(\widehat{m}_0(t,\widehat{\beta})-\widehat{m}_0(t,\beta_0))+\sqrt{n}(\widehat{m}_0(t,\beta_0)-m_0(t)).\end{aligned} \tag{S2.5}$$

Denote:

$$\Phi(t)=\frac{1}{n}\sum_{i=1}^{n}\delta_i I(Y_i>t)\{Y_i\widehat{S}_C(Y_i-A_i)\}^{-1}\exp(\beta_0^T X_i),$$

and

$$R(t,\mu)=\frac{1}{n}\sum_{i=1}^{n}M_i(t)I(Y_i-A_i\geq\mu).$$

The first part in (S2.5) can be written as follows:

$$\begin{aligned}\sqrt{n}(\widehat{m}_0(t,\widehat{\beta})-\widehat{m}_0(t,\beta_0)) &= -\overline{x}^*(t)m_0(t)\sqrt{n}(\widehat{\beta}-\beta_0)+o_p(1), \\ &= -\overline{x}^*(t)m_0(t)\frac{1}{\sqrt{n}}\sum_{i=1}^{n}B^{-1}\xi_i+o_p(1),\end{aligned}$$

and the second part in (S2.5) can be written as follows:

$$\begin{aligned}\sqrt{n}(\widehat{m}_0(t,\beta_0)-m_0(t)) &= \Phi(t)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[M_i(t)+M_i(t)\frac{S_C(Y_i-A_i)-\widehat{S}_C(Y_i-A_i)}{\widehat{S}_C(Y_i-A_i)}\big] \\ &= \phi(t)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[M_i(t)+M_i(t)\int_0^\tau\frac{I(\mu\leq Y_i-A_i)}{n\pi(\mu)}\sum_{j=1}^{n}dM_j^C(\mu)\big] \\ &\quad +o_p(1) \\ &= \phi(t)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big[M_i(t)+\int_0^\tau\frac{r(t,\mu)}{\pi(\mu)}dM_i^C(\mu)\big]+o_p(1),\end{aligned}$$

where $\phi(t)$ and $r(t,\mu)$ is the corresponding nonrandom limit of $\Phi(t)$ and $R(t,\mu)$.

Therefore,

$$\sqrt{n}(\widehat{m}_0(t)-m_0(t))=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\psi_i(t)+o_p(1),$$

where

$$\psi_i(t)=\phi(t)^{-1}[M_i(t)+\int_0^\tau\frac{r(t,\mu)}{\pi(\mu)}dM_i^C(\mu)]-\overline{x}(t,\beta_0)m_0(t)B^{-1}\xi_i.$$

Since the terms in last equation are independent zero-mean random variables for every fixed t, the multivariate central limit theorem implies that the finite dimensional distribution of the process $\sqrt{n}\{\widehat{m}_0(t)-m_0(t)\}(0\leq t\leq\tau)$ converges to a zero-mean Gaussian process. In order to prove the weak convergence, it suffices to show the tightness. It reduces to the tightness of $n^{-1/2}\sum_{i=1}^{n}M_i(t)$. By Assumption (A1), $m_0(t)$ is of bounded variation. Since $M_i(t)$ can be written as the sum or product of monotone functions of t and is thus manageable (Pollard (1990), page 38; Bilisa, Gu and Ying (1997), Lemmas A.1-A.2), it follows from the functional central limit theorem(Pollard (1990), page 53; Lin et al.(2000), page 726) that $n^{-1/2}\sum_{i=1}^{n}M_i(t)$ is tight. Therefore $n^{1/2}\{\widehat{m}_0(t)-m_0(t)\}(0\leq t\leq\tau)$ is tight and converges weakly to a zero-mean Gaussian process with the covariance function $\Gamma(s,t)=E\{\psi_i(s)\psi_i(t)\}$ at (s,t), which can be consistently estimated by $\widehat{\Gamma}(s,t)$. $\square$

S3 Proof of Theorem 3.

By the maximum partial likelihood theory in Flemming & Harrington (1991), we can obtain the uniform consistency of $\widehat{\alpha}$ and $\widehat{\Lambda}_0(t)$ in $[0, \tau]$. Then similar arguments as that in the proof of Theorem 1 can be used to obtain the conclusion of Theorem 3. Here, we omit the details. □

S4 Proof of Theorem 4.

(1) Since C follows cox proprotional hazards model, from Flemming & Harrington (1991), we can obtain:

$$\begin{aligned}\widehat{\alpha} - \alpha_0 &= \Omega^{-1}\frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau}(Z_i - \overline{z}(t))dM_i^d(t) + o_p(n^{-1/2}),\\ \widehat{\Lambda}_0(t) - \Lambda_0(t) &= \frac{1}{n}\sum_{i=1}^{n}\int_0^{t}\frac{M_i^d(u)}{s^{(0)}(u)} - \int_0^t \overline{z}(u)' d\Lambda_0(u)(\widehat{\alpha} - \alpha_0) + o_p(n^{-1/2}),\end{aligned} \tag{S4.6}$$

where $M_i^d(t) = N_i^C(t) - \int_0^t Y_i(u)\exp(\alpha_0^T Z_i)d\Lambda_0(u)$, $s^{(k)}(t;\alpha) = \lim_{n\to\infty} S^{(k)}(t;\alpha)$, $s^{(0)}(t) = S^{(0)}(t;\alpha_0)$, $\overline{z}(t;\alpha) = \lim_{n\to\infty}\overline{Z}(t;\alpha)$, $\overline{z}(t) = \overline{z}(t;\alpha_0)$, $\Omega = \lim_{n\to\infty}\widehat{\Omega}$.

Similar to the proof of (S2.4), we obtain

$$\begin{aligned}U_d(\beta_0) &\triangleq \sum_{i=1}^{n}\int_0^{\tau}\frac{\delta_i I(Y_i > t)X_i}{Y_i\widehat{S}_C(Y_i - A_i|Z_i)}\big[(Y_i - t) - m_0(t)\exp(\beta_0^T X_i) - \exp(\beta_0^T X_i)\\ &\quad (\widehat{m}_{0d}(t,\beta_0) - m_0(t))\big]dH(t)\\ &= \sum_{i=1}^{n}\int_0^{\tau} M_i^*(t)(X_i - \overline{x}_d^*(t))dh(t) + \sum_{i=1}^{n}\int_0^{\tau} M_i^*(t)(X_i - \overline{x}_d^*(t))dh(t)\\ &\quad \frac{S_C(Y_i - A_i|Z_i) - \widehat{S}_C(Y_i - A_i|Z_i)}{S_C(Y_i - A_i|Z_i)} + o_p(n^{1/2}),\end{aligned} \tag{S4.7}$$

where $M_i^*(t) = \frac{\delta_i I(Y_i>t)}{Y_i S_C(Y_i - A_i|Z_i)}\big[(Y_i - t) - m_0(t)\exp(\beta_0^T X_i)\big]$, $\overline{x}_d^*(t) = \lim_{n\to\infty}\overline{X}_d^*(t)$, $\overline{X}_d^*(t) = \frac{\sum \delta_i I(Y_i>t)\{Y_i\widehat{S}_C(Y_i-A_i|Z_i)\}^{-1}\exp(\beta_0^T X_i)X_i}{\sum \delta_i I(Y_i>t)\{Y_i\widehat{S}_C(Y_i-A_i|Z_i)\}^{-1}\exp(\beta_0^T X_i)}$, and $\widehat{S}(t|Z_i) = \exp(-\exp(\widehat{\alpha}^T Z_i)\widehat{\Lambda}_0(t))$.

Using functional delta method in Van der Vaart & Wellner (1996) and(S4.6), (S4.7) becomes:

$$\begin{aligned}U_d(\beta_0) &= \sum_{i=1}^{n}\int_0^{\tau} M_i^*(t)(X_i - \overline{x}_d^*(t))dh(t) + \sum_{i=1}^{n}\Big\{\int_0^{\tau}\frac{Q_d(t)}{s^{(0)}(t)}dM_i^d(t) +\\ &\quad \big(D_d - \int_0^{\tau} Q_d(t)\overline{z}^T(t)d\Lambda_0(t)\big)\Omega^{-1}\int_0^{\tau}(Z_i - \overline{z}(t))dM_i^d(t)\Big\} + o_p(n^{1/2}),\end{aligned}$$

where

$$
\begin{aligned}
Q_d(t) &= \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} M_i^*(t)(X_i - \overline{x}_d^*(t))dh(t)\exp(\alpha_0^T Z_i)I(Y_i - A_i \geq t),\\
D_d &= \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} M_i^*(t)(X_i - \overline{x}_d^*(t))dh(t)\Lambda_0(Y_i - A_i)\exp(\alpha_0^T Z_i)Z_i^T.
\end{aligned}
$$

Followed by the uniform strong law of large numbers, we know that $Q_d(t)$ and D_d converges almost surely to nonrandom function $q_d(t)$ and D_d^* in $[0, \tau]$.

Therefore,

$$
\frac{1}{\sqrt{n}}U_d(\beta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i^d + o_p(1),
$$

where

$$
\begin{aligned}
\xi_i^d &= \int_0^{\tau} M_i^*(t)(X_i - \overline{x}_d^*(t))dh(t) + \int_0^{\tau}\frac{q_d(t)}{s^{(0)}(t)}dM_i^d(t)\\
&\quad + \big[D_d^* - \int_0^{\tau} q_d(t)\overline{z}^T(t)d\Lambda_0(t)\big]\Omega^{-1}\int_0^{\tau}(Z_i - \overline{z}(t))dM_i^d(t).
\end{aligned}
$$

It then follows from the multivariate central limit theorem that $\frac{1}{\sqrt{n}}U_d(\beta_0)$ is asymptotically normal with mean 0 and covaraite matrix $\Sigma_d = E[\xi_i^{d\otimes 2}]$.

Let $B_d(\beta_0) = -\frac{1}{n}\sum_{i=1}^{n}\frac{\partial U_d(\beta)}{\partial \beta^T}|_{\beta=\beta_0}$ and $B_d(\beta_0)$ converges to a nonrandom matrix B_d by the law of large number.

By the Taylor expansion of $U_d(\beta)$ at β_0, we have

$$
\begin{aligned}
\sqrt{n}(\widehat{\beta}_d - \beta_0) &= B_d^{-1}n^{-1/2}U_d(\beta_0) + o_p(1),\\
&= \frac{1}{\sqrt{n}}B_d^{-1}\sum_{i=1}^{n}\xi_i^d + o_p(1).
\end{aligned}
$$

Hence, $\sqrt{n}(\widehat{\beta}_d - \beta_0)$ is asymptotically zero-mean normal with covariance matrix $B_d^{-1}\Sigma_d B_d^{-1}$, which can be consistently estimated by $\widehat{B}_d^{-1}\widehat{\Sigma}_d\widehat{B}_d^{-1}$, where

$$
\begin{aligned}
\widehat{B}_d &= \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau}\frac{\delta_i I(Y_i > t)}{Y_i\widehat{S}_C(Y_i - A_i|Z_i)}\big(X_i - \overline{X}_d(t,\widehat{\beta}_d)\big)^{\otimes 2}\widehat{m}_{0d}(t)\exp(\widehat{\beta}_d^T X_i)dH(t),\\
\widehat{\xi}_i^d &= \int_0^{\tau}\widehat{M}_i^*(t)\{X_i - \overline{X}_d(t,\widehat{\beta}_d)\}dH(t) + \int_0^{\tau}\frac{\widehat{Q}_d(t)}{S^{(0)}(t;\widehat{\alpha})}d\widehat{M}_i^d(t) + \big[\widehat{D}_d\\
&\quad - \int_0^{\tau}\widehat{Q}_d(t)\overline{Z}^T(t;\widehat{\alpha})d\widehat{\Lambda}_0(t)\big]\widehat{\Omega}^{-1}\int_0^{\tau}(Z_i - \overline{Z}(t;\widehat{\alpha}))d\widehat{M}_i^d(t),
\end{aligned}
$$

and

$$
\widehat{\Sigma}_d = \frac{1}{n}\sum_{i=1}^{n}\widehat{\xi}_i^{d\otimes 2}.
$$

(2) We begin to show the weak convergence of $\widehat{m}_{0d}(t)$. Since

$$\begin{aligned}\sqrt{n}(\widehat{m}_{0d}(t)-m_0(t)) &= \sqrt{n}(\widehat{m}_{0d}(t,\widehat{\beta})-\widehat{m}_{0d}(t,\beta_0))+\sqrt{n}(\widehat{m}_{0d}(t,\beta_0)-m_0(t)),\\ &= -\overline{x}_d^*(t)m_0(t)\frac{1}{\sqrt{n}}\sum_{i=1}^n B_d^{-1}\xi_i^d+\sqrt{n}(\widehat{m}_{0d}(t,\beta_0)-m_0(t))+o_p(1).\end{aligned}$$

Let

$$\begin{aligned}\Phi_d(t) &= \frac{1}{n}\sum_{i=1}^n \delta_i I(Y_i>t)\{Y_i\widehat{S}_C(Y_i-A_i|Z_i)\}^{-1}\exp(\beta_0^T X_i),\\ R_d(t,u) &= \frac{1}{n}\sum_{i=1}^n M_i^*(t)\exp(\alpha_0' Z_i)I(Y_i-A_i\geq u),\\ R_1(t) &= \frac{1}{n}\sum_{i=1}^n M_i^*(t)\Lambda_0(Y_i-A_i)\exp(\alpha_0^T Z_i)Z_i'.\end{aligned}$$

Using similar methods as that in (S4.7), we obtain:

$$\begin{aligned}&\sqrt{n}(\widehat{m}_{0d}(t,\beta_0)-m_0(t))=\Phi_d(t)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n\big[M_i^*(t)+M_i^*(t)(1-\frac{\widehat{S}_C(Y_i-A_i|Z_i)}{S_C(Y_i-A_i|Z_i)})\big],\\ =\ &\Phi_d(t)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n\Big[M_i^*(t)+\int_0^\tau\frac{R_d(t,u)}{s^{(0)}(u)}dM_i^d(u)+\big[R_1(t)-\int_0^\tau R_d(t,u)\overline{z}^T(u)d\Lambda_0(u)\big],\\ &\Omega^{-1}\int_0^\tau(Z_i-\overline{z}(t))dM_i^d(t)\Big]+o_p(1),\\ =\ &\phi_d(t)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n\Big[M_i^*(t)+\int_0^\tau\frac{r_d(t,u)}{s^{(0)}(u)}dM_i^d(u)+\big[r_1(t)-\int_0^\tau r_d(t,u)\overline{z}^T(u)d\Lambda_0(u)\big],\\ &\Omega^{-1}\int_0^\tau(Z_i-\overline{z}(t))dM_i^d(t)\Big]+o_p(1).\end{aligned}$$

Here, $\phi_d(t)$, $r_d(t,u)$ and $r_1(t)$ are the limits of $\Phi_d(t)$, $R_d(t,u)$ and $R_1(t)$ and all are nonrandom functions.

Therefore,

$$\sqrt{n}(\widehat{m}_{0d}(t)-m_0(t))=\frac{1}{\sqrt{n}}\sum_{i=1}^n\psi_i^d(t)+o_p(1),$$

where

$$\begin{aligned}\psi_i^d(t) &= \phi_d^{-1}(t)\Big\{M_i^*(t)+\int_0^\tau\frac{r_d(t,u)}{s^{(0)}(u)}dM_i^d(u)+\big[r_1(t)-\int_0^\tau r_d(t,u)\overline{z}^T(u)d\Lambda_0(u)\big],\\ &\quad\Omega^{-1}\int_0^\tau(Z_i-\overline{z}(t))dM_i^d(t)\Big\}-\overline{x}_d^*(t)m_0(t)B_d^{-1}\xi_i^d.\end{aligned}$$

Then using similar arguments as that in the proof of Theorem 2, we complete the proof. Here, we omit the details. □

References

Bilias, Y., Gu, M., and Ying, Z. (1997). Towards a general asymptotic theory for cox model with staggered entry. *The Annals of Statistics* **25**, 662-682.

Flemming, T. R. and Harrington, D. P. (1991). Counting Processes and Survival Analysis. Wiley, New York.

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 Royal Statistical Society B* **62**, 711-730.

Pollard, D. (1990). *Empirical Processes: Theory and Applications.* Institute of Mathematical Statistics, CA.

Rudin, W. B. (1976). *Principles of Mathematical Analysis.* 3rd Edition, McGraw-Hill.

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence and empirical processes.* Springer Verlag, New York.