Web-Based Supplementary Materials for "Competing Risks Regression for Stratified Data" by Zhou, B., Latouche, A., Rocha, V., & Fine, J.

Web Appendix A: Assumptions

The following list of conditions are assumed throughout the paper:

1. $\int_0^\tau \lambda_{k0}(t)dt < \infty$.

2. $\mathbf{Z}_{lki}(\cdot)(k = 1, \cdots, s; i = 1, \cdots, n_k)$ have bounded total variations, i.e. for lth component of $\mathbf{Z}_{ki}, l = 1, \ldots, m$, $|\mathbf{Z}_{lki}(0)| + \int_0^\tau |d\mathbf{Z}_{lki}(t)| \leq M$, where M is a constant.

3. $\{N_{ki}(\cdot), Y_{ki}(\cdot), \mathbf{Z}_{ki}(\cdot), i = 1, \ldots, n_k\}_{k=1,\ldots,s}$ are independently distributed; for regular stratified data, $\{N_{ki}(\cdot), Y_{ki}(\cdot), \mathbf{Z}_{ki}(\cdot)\}_{i=1,\ldots,n_k}$ are independently and identically distributed for each stratum k.

4. regularity conditions for different regimes:

   - For regular stratified data, there exists a neighborhood $\mathcal{B}$ of $\boldsymbol{\beta}_0$ and scalar,vector and matrix functions $s_k^{(0)}$, $\boldsymbol{s}_k^{(1)}$ and $\boldsymbol{s}_k^{(2)}$ defined on $\mathcal{B} \times [0, \tau]$ such that for $p = 0, 1, 2$, $\sup_{t\in[0,\tau],\boldsymbol{\beta}\in\mathcal{B}} \|\boldsymbol{S}_k^{(p)}(\boldsymbol{\beta}, t) - \boldsymbol{s}_k^{(p)}(\boldsymbol{\beta}, t)\| \xrightarrow{p} 0$. Define $\boldsymbol{e}_k = \boldsymbol{s}_k^{(1)}/s_k^{(0)}$ and $\boldsymbol{v}_k = \boldsymbol{s}_k^{(2)}/s_k^{(0)} - \boldsymbol{e}_k^{\otimes 2}$.

     The matrix $\boldsymbol{\Sigma}_{k\tau} = \int_0^\tau \boldsymbol{v}_k(\boldsymbol{\beta}_0, t)s^{(0)}(\boldsymbol{\beta}_0, t)\lambda_{0k}(t)dt$ is positive definite.

   - For highly stratified data, both

     $\boldsymbol{I} \equiv \lim_{s\to\infty} s^{-1} \sum_{k=1}^s \mathrm{E}\left[\sum_{i=1}^{n_k} \int_0^\tau \{\mathbf{Z}_{ki}(u) - \overline{\mathbf{Z}}_k(\boldsymbol{\beta}_0, u)\}^{\otimes 2} Y_{ki}^*(u) e^{\boldsymbol{\beta}_0' \mathbf{Z}_{ki}(u)} \frac{d\overline{N}_k(u)}{S_k{}^{(0)}(\boldsymbol{\beta}_0, u)}\right]$

     for censoring complete case and

     $\widetilde{\boldsymbol{I}} \equiv \lim_{s\to\infty} s^{-1} \sum_{k=1}^s \mathrm{E}\left[\sum_{i=1}^{n_k} \int_0^\tau \{\mathbf{Z}_{ki}(u) - \widetilde{\mathbf{Z}}_k(\boldsymbol{\beta}_0, u)\}^{\otimes 2} w_{ki}(u) Y_{ki}(u) e^{\boldsymbol{\beta}_0' \mathbf{Z}_{ki}(u)} \frac{d\overline{N}_k(u)}{\widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta}_0, u)}\right]$

     for right censored case are positive definite.

Web Appendix B: Consistency of $\widehat{\beta}$ for Censoring Complete Highly Stratified Data

Let $C(\boldsymbol{\beta}, t)$ be the logarithm of the partial likelihood evaluated at time t, so we have

$$C(\boldsymbol{\beta}, t) = \sum_{k=1}^{s} \left[\sum_{i=1}^{n_k} \int_0^t \boldsymbol{\beta}' \mathbf{Z}_{ki}(u) dN_{ki}(u) - \int_0^t \left\{ \log \sum_{i'=1}^{n_k} Y^*_{ki'}(u) e^{\boldsymbol{\beta}' \mathbf{Z}_{ki'}(u)} \right\} d\overline{N}_k(u) \right],$$

where $\overline{N}_k = \sum_{i=0}^{n_k} N_{ki}$ and n_k is finite. We also have $C(\boldsymbol{\beta}, \tau) = \log L(\boldsymbol{\beta})$. $\widehat{\boldsymbol{\beta}}$ is the solution to the estimating equation $(\partial/\partial\boldsymbol{\beta}) C(\boldsymbol{\beta}, \tau) = 0$.

Consider the process

$$\begin{aligned} X(\boldsymbol{\beta}, t) &= s^{-1}(C(\boldsymbol{\beta}, t) - C(\boldsymbol{\beta}_0, t)) \\ &= s^{-1} \sum_{k=1}^{s} \left[\sum_{i=1}^{n_k} \int_0^t (\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \mathbf{Z}_{ki}(u) dN_{ki}(u) - \int_0^t \log \left\{ \frac{\sum_{i'=1}^{n_k} Y^*_{ki'}(u) e^{\boldsymbol{\beta}' \mathbf{Z}_{ki'}(u)}}{\sum_{i'=1}^{n_k} Y^*_{ki'}(u) e^{\boldsymbol{\beta}_0' \mathbf{Z}_{ki'}(u)}} \right\} d\overline{N}_k(u) \right] \\ &= s^{-1} \sum_{k=1}^{s} \left[\sum_{i=1}^{n_k} \int_0^t (\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \mathbf{Z}_{ki}(u) dN_{ki}(u) - \int_0^t \log \left\{ \frac{S_k^{(0)}(\boldsymbol{\beta}, u)}{S_k^{(0)}(\boldsymbol{\beta}_0, u)} \right\} d\overline{N}_k(u) \right]. \end{aligned}$$

where $\boldsymbol{S}_k{}^{(p)}(\boldsymbol{\beta}, t) = n_k^{-1} \sum_{i=1}^{n_k} Y^*_{ki}(t) \mathbf{Z}_{ki}(t)^{\otimes p} e^{\boldsymbol{\beta}' \mathbf{Z}_{ki}(t)}$, $p = 0, 1, 2$. Define

$$X_k(\boldsymbol{\beta}, t) = \sum_{i=1}^{n_k} \int_0^t (\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \mathbf{Z}_{ki}(u) dN_{ki}(u) - \int_0^t \log \left\{ \frac{S_k^{(0)}(\boldsymbol{\beta}, u)}{S_k^{(0)}(\boldsymbol{\beta}_0, u)} \right\} d\overline{N}_k(u), \qquad k = 1, \ldots, s.$$

Then, $X(\boldsymbol{\beta}, t) = s^{-1} \sum_{i=1}^{s} X_k(\boldsymbol{\beta}, t)$. For convenience, we suppress t at $t = \tau$ in these expressions. Hence, $X_k(\boldsymbol{\beta}, \tau) = X_k(\boldsymbol{\beta})$, and $X(\boldsymbol{\beta}, \tau) = X(\boldsymbol{\beta})$.

Suppose $\mathcal{X}_k(\boldsymbol{\beta}) = \mathrm{E}\{X_k(\boldsymbol{\beta})\}$ and $\mathcal{X}(\boldsymbol{\beta}) = \lim_{s\to\infty} s^{-1} \sum_{k=1}^{s} \mathcal{X}_k(\boldsymbol{\beta})$.

The conditions 1 and 2 lead to the fact that $d\overline{N}_k(t)$ and $S_k^{(0)}(\boldsymbol{\beta}, t)$ have bounded variation (Lin *et al.* , 2000). This result and the finiteness of strata sizes n_k's can be used to show that $X_k(\boldsymbol{\beta}), k = 1, \ldots, s$ are bounded. As a result, $s^{-1} \sum_{k=1}^{s} \mathrm{E}\{|X_k(\boldsymbol{\beta}) - \mathcal{X}_k(\boldsymbol{\beta})|^2\} \leq \max_{k\in\{1,\ldots,s\}} \mathrm{E}\{|X_k(\boldsymbol{\beta}) - \mathcal{X}_k(\boldsymbol{\beta})|^2\}$ is bounded. Hence, $s^{-2} \sum_{k=1}^{s} \mathrm{E}\{|X_k(\boldsymbol{\beta}) - \mathcal{X}_k(\boldsymbol{\beta})|^2\} \to 0$. In addition,the condition 3 guarantees the independence of $X_k(\boldsymbol{\beta})$'s. Therefore, By Kolmogorov Strong Law of Large Numbers , $X(\boldsymbol{\beta})$ converges to $\mathcal{X}(\boldsymbol{\beta})$ almost surely, for all $\boldsymbol{\beta} \in \mathcal{B}$, where $\mathcal{B}$ is any compact neighborhood of $\boldsymbol{\beta}_0$ (Sen and Singer, 1993, chap.2).

$X_k(\boldsymbol{\beta})$ is random concave by Anderson and Gill (1982). Therefore, $X(\boldsymbol{\beta})$ is a random concave function since it is just a sum of random concave functions. This implies the uniform convergence of $X(\boldsymbol{\beta})$ to $\mathcal{X}(\boldsymbol{\beta})$ on compact subspaces $\boldsymbol{\beta} \in \mathcal{B}$. That is, $\sup_{\boldsymbol{\beta}\in\mathcal{B}} |X(\boldsymbol{\beta}) - \mathcal{X}(\boldsymbol{\beta})| \to_p 0$ (Rockafellar, 1970, Theorem 10.8).

Now by the boundedness conditions and the independence of range of integral on $\boldsymbol{\beta}$, we can evaluate the first and second derivatives of $\mathcal{X}(\boldsymbol{\beta})$ by taking partial derivatives inside the integral and expectation. Clearly, the first derivative evaluated at $\boldsymbol{\beta}_0$ is $\frac{\partial \mathcal{X}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} = 0$, since the expected value of the score function $(\partial/\partial\boldsymbol{\beta})C(\boldsymbol{\beta},\tau)$ equals to 0 at $\boldsymbol{\beta} = \boldsymbol{\beta}_0$. Furthermore,

$$\frac{\partial^2 \mathcal{X}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2} = -\lim_{s\to\infty} s^{-1} \sum_{k=1}^{s} \mathrm{E}\left[\sum_{i=1}^{n_k} \int_0^\tau \{\mathbf{Z}_{ki}(u) - \overline{\mathbf{Z}}_k(\boldsymbol{\beta},u)\}^{\otimes 2} Y_{ki}^*(u) e^{\boldsymbol{\beta}'\mathbf{Z}_{ki}(u)} \frac{d\overline{N}_k(u)}{S_k{}^{(0)}(\boldsymbol{\beta},u)}\right], \quad \text{(A.1)}$$

which equals to $\boldsymbol{I}$ when evaluated at $\boldsymbol{\beta}_0$. By condition 4, (A.1) is negative definite. Therefore, $\mathcal{X}(\boldsymbol{\beta})$ is a concave function of $\boldsymbol{\beta}$ with a unique maximum at $\boldsymbol{\beta}_0$. Thus, the maximizer of $X(\boldsymbol{\beta}) : \widehat{\boldsymbol{\beta}}$ converges in probability to the unique maximum of $\mathcal{X}(\boldsymbol{\beta}) : \boldsymbol{\beta}_0$ (Anderson and Gill, 1982, Corollary II.2).

Web Appendix C.1: Consistency of $\widehat{\beta}$ for Right Censored Highly Stratified Data

When the highly stratified data are right censored, we show the consistency of $\widehat{\boldsymbol{\beta}}$ by modifying the proof in Appendix A. Note that the risk process $Y^*(t) = I(C \geq t)Y(t)$ is replaced by $\widehat{Y}(t) = \widehat{w}(t)Y(t)$, where $\widehat{w}(t) = I(C \geq T \wedge t)\widehat{G}(t)/\widehat{G}(X \wedge t)$. Thus, we have

$$\widetilde{C}(\boldsymbol{\beta},t) = \sum_{k=1}^{s}\left[\sum_{i=1}^{n_k}\int_0^t \boldsymbol{\beta}'\mathbf{Z}_{ki}(u)dN_{ki}(u) - \int_0^t \left\{\log \sum_{i'=1}^{n_k} \widehat{w}_{ki'}(u)Y_{ki'}(t)e^{\boldsymbol{\beta}'\mathbf{Z}_{ki'}(u)}\right\} d\overline{N}_k(u)\right].$$

Since $\widehat{w}(t) \to w(t) = I(C \geq T \wedge t)G(t)/G(X \wedge t)$, where $G(t) = Pr(C_{ki} \geq T)$, $k = 1, \ldots, s$; $i =$

$1, \ldots, n_k$, we further have

$$\widetilde{C}(\boldsymbol{\beta}, t) = \sum_{k=1}^{s} \left[\sum_{i=1}^{n_k} \int_0^t \boldsymbol{\beta}' \mathbf{Z}_{ki}(u) dN_{ki}(u) - \int_0^t \left\{ \log \sum_{i'=1}^{n_k} w_{ki'}(u) Y_{ki'}(u) e^{\boldsymbol{\beta}' \mathbf{Z}_{ki'}(u)} \right\} d\overline{N}_k(u) \right] + o_p(1).$$

Hence,

$$\begin{aligned}
&\widetilde{X}(\boldsymbol{\beta}, t) \\
=& s^{-1}(\widetilde{C}(\boldsymbol{\beta}, t) - \widetilde{C}(\boldsymbol{\beta}_0, t)) \\
=& s^{-1} \sum_{k=1}^{s} \left[\sum_{i=1}^{n_k} \int_0^t (\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \mathbf{Z}_{ki}(u) dN_{ki}(u) - \int_0^t \log \left\{ \frac{\sum_{i'=1}^{n_k} w_{ki'}(u) Y_{ki'}(u) e^{\boldsymbol{\beta}' \mathbf{Z}_{ki'}(u)}}{\sum_{i'=1}^{n_k} w_{ki'}(u) Y_{ki'}(u) e^{\boldsymbol{\beta}_0' \mathbf{Z}_{ki'}(u)}} \right\} d\overline{N}_k(u) \right] + o_p(1) \\
=& s^{-1} \sum_{k=1}^{s} \left[\sum_{i=1}^{n_k} \int_0^t (\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \mathbf{Z}_{ki}(u) dN_{ki}(u) - \int_0^t \log \left\{ \frac{\widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta}, u)}{\widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta}_0, u)} \right\} d\overline{N}_k(u) \right] + o_p(1),
\end{aligned}$$

where

$$\widetilde{\mathbf{S}}_k^{(p)}(\boldsymbol{\beta}_0, t) = n_k^{-1} \sum_{i=1}^{n_k} w_{ki}(t) Y_{ki}(t) \mathbf{Z}_{ki}(t)^{\otimes p} e^{\boldsymbol{\beta}_0' \mathbf{Z}_{ki}(t)}, \ p = 0, 1, 2.$$

Correspondingly,

$$\begin{aligned}
\widetilde{X}_k(\boldsymbol{\beta}, t) &= \sum_{i=1}^{n_k} \int_0^t (\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \mathbf{Z}_{ki}(u) dN_{ki}(u) - \int_0^t \log \left\{ \frac{\widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta}, u)}{\widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta}_0, u)} \right\} d\overline{N}_k(u), \\
\widetilde{\mathcal{X}}_k(\boldsymbol{\beta}) &= \mathrm{E}\left\{ \widetilde{X}_k(\boldsymbol{\beta}) \right\}, \quad \text{and} \\
\widetilde{\mathcal{X}}(\boldsymbol{\beta}) &= \lim_{s \to \infty} s^{-1} \sum_{k=1}^{s} \widetilde{\mathcal{X}}_k(\boldsymbol{\beta}).
\end{aligned}$$

Since $w(t)Y(t) \leq I(C \geq t)Y(t)$, $\widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta}, t) \leq S_k^{(0)}(\boldsymbol{\beta}, t)$. We can show $\widetilde{X}_k(\boldsymbol{\beta}), k = 1, \ldots, s$ are bounded using the boundedness results from appendix A. Therefore, $s^{-1} \sum_{k=1}^{s} \mathrm{E}\left\{ |\widetilde{X}_k(\boldsymbol{\beta}) - \widetilde{\mathcal{X}}_k(\boldsymbol{\beta})|^2 \right\}$ $\leq \max_{k \in \{1, \ldots, s\}} \mathrm{E}\left\{ |\widetilde{X}_k(\boldsymbol{\beta}) - \widetilde{\mathcal{X}}_k(\boldsymbol{\beta})|^2 \right\}$ is bounded. Hence, $s^{-2} \sum_{k=1}^{s} \mathrm{E}\left\{ |\widetilde{X}_k(\boldsymbol{\beta}) - \widetilde{\mathcal{X}}_k(\boldsymbol{\beta})|^2 \right\} \to 0$. In addition, condition 3 guarantees the independence of $\widetilde{X}_k(\boldsymbol{\beta})$'s. Therefore, By Kolmogorov Strong Law of Large Numbers, $\widetilde{X}(\boldsymbol{\beta})$ converges to $\widetilde{\mathcal{X}}(\boldsymbol{\beta})$ almost surely, for all $\boldsymbol{\beta} \in \mathcal{B}$, where $\mathcal{B}$ is any compact neighborhood of $\boldsymbol{\beta}_0$ (Sen and Singer, 1993, chap.2).

The uniform convergence is showed next.

$$\frac{\partial^2 \widetilde{X}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2} = -s^{-1}\sum_{k=1}^{s}\left[\sum_{i=1}^{n_k}\int_0^{\tau}\left\{\mathbf{Z}_{ki}(u)-\widetilde{\mathbf{Z}}_k(\boldsymbol{\beta},u)\right\}^{\otimes 2} w_{ki}(u)Y_{ki}(u)e^{\boldsymbol{\beta}'\mathbf{Z}_{ki}(u)}\frac{d\overline{N}_k(u)}{\widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta},u)}\right], \quad \text{(A.2)}$$

where $\widetilde{\mathbf{Z}}_k(\boldsymbol{\beta}_0,t) = \widetilde{\mathbf{S}}_k^{(1)}(\boldsymbol{\beta}_0,t)/\widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta}_0,t)$. Since $w_{ki}(u)Y_{ki}(u), d\overline{N}_k(u)$ and $\widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta},u)$ are all nonnegative, (A.2) is negative semidefinite. Thus, $\widetilde{X}(\boldsymbol{\beta})$ is a random concave function.

Now by the boundedness conditions and the independence of range of integral on $\boldsymbol{\beta}$, we can evaluate the first and second derivatives of $\widetilde{\mathcal{X}}(\boldsymbol{\beta})$ by taking partial derivatives inside the integral and expectation. Clearly, the first derivative evaluated at $\boldsymbol{\beta}_0$ is $\frac{\partial \widetilde{\mathcal{X}}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} = 0$, since the expected value of the score function $(\partial/\partial\boldsymbol{\beta})\widetilde{C}(\boldsymbol{\beta},\tau)$ equals to 0 at $\boldsymbol{\beta}=\boldsymbol{\beta}_0$. Furthermore,

$$\frac{\partial^2 \widetilde{\mathcal{X}}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2} = -\lim_{s\to\infty} s^{-1}\sum_{k=1}^{s}\mathrm{E}\left[\sum_{i=1}^{n_k}\int_0^{\tau}\{\mathbf{Z}_{ki}(u)-\widetilde{\mathbf{Z}}_k(\boldsymbol{\beta},u)\}^{\otimes 2} w_{ki}(u)Y_{ki}(u)e^{\boldsymbol{\beta}'\mathbf{Z}_{ki}(u)}\frac{d\overline{N}_k(u)}{\widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta},u)}\right], \quad \text{(A.3)}$$

which equals to $\widetilde{\boldsymbol{I}}$ when evaluated at $\boldsymbol{\beta}_0$. By condition 4, (A.3) is negative definite. Therefore, $\widetilde{\mathcal{X}}(\boldsymbol{\beta})$ is a concave function of $\boldsymbol{\beta}$ with a unique maximum at $\boldsymbol{\beta}_0$. Thus, the maximizer of $\widetilde{X}(\boldsymbol{\beta}) : \widehat{\boldsymbol{\beta}}$ converges in probability to the unique maximum of $\widetilde{\mathcal{X}}(\boldsymbol{\beta}) : \boldsymbol{\beta}_0$ (Anderson and Gill, 1982, Corollary II.2).

Web Appendix C.2: Asymptotic Normality of $s^{-\frac{1}{2}}\boldsymbol{U}_1(\boldsymbol{\beta}_0,t)$ for Highly Stratified Right Censored Data

Rewrite equation (3): $\boldsymbol{U}_1(\boldsymbol{\beta}_0,t) = \sum_{k=1}^{s}\sum_{i=1}^{n_k}\int_0^t\left\{\mathbf{Z}_{ki}(u)-\widehat{\mathbf{Z}}_k(\boldsymbol{\beta}_0,u)\right\}\widehat{w}_{ki}(u)dN_{ki}(\boldsymbol{\beta}_0,u)$ as

$$\begin{aligned}\boldsymbol{U}_1(\boldsymbol{\beta}_0,t) &= \sum_{k=1}^{s}\sum_{i=1}^{n_k}\int_0^t\left\{\mathbf{Z}_{ki}(u)-\widetilde{\mathbf{Z}}_k(\boldsymbol{\beta}_0,u)\right\}\widehat{w}_{ki}(u)dN_{ki}(\boldsymbol{\beta}_0,u) \quad \text{(A.4)}\\ &+\sum_{k=1}^{s}\sum_{i=1}^{n_k}\int_0^t\left\{\widetilde{\mathbf{Z}}_k(\boldsymbol{\beta}_0,u)-\widehat{\mathbf{Z}}_k(\boldsymbol{\beta}_0,u)\right\}\widehat{w}_{ki}(u)dN_{ki}(\boldsymbol{\beta}_0,u),\end{aligned}$$

where

$$\widehat{\mathbf{S}}_k^{(p)}(\boldsymbol{\beta}, t) = n_k^{-1} \sum_{i=1}^{n_k} \widehat{w}_{ki}(t) Y_{ki}(t) \mathbf{Z}_{ki}(t)^{\otimes p} e^{\boldsymbol{\beta}' \mathbf{Z}_{ki}(t)}, \; p = 0, \ldots, 2,$$

$$\widehat{\mathbf{Z}}_k(\boldsymbol{\beta}, t) = \widehat{\mathbf{S}}_k^{(1)}(\boldsymbol{\beta}, t) / \widehat{S}_k^{(0)}(\boldsymbol{\beta}, t),$$

$$\widetilde{\mathbf{S}}_k^{(p)}(\boldsymbol{\beta}, t) = n_k^{-1} \sum_{i=1}^{n_k} w_{ki}(t) Y_{ki}(t) \mathbf{Z}_{ki}(t)^{\otimes p} e^{\boldsymbol{\beta}' \mathbf{Z}_{ki}(t)}, \; p = 0, \ldots, 2,$$

$$\widetilde{\mathbf{Z}}_k(\boldsymbol{\beta}, t) = \widetilde{\mathbf{S}}_k^{(1)}(\boldsymbol{\beta}, t) / \widetilde{S}_k^{(0)}(\boldsymbol{\beta}, t).$$

Clearly,

$$\text{first part of } (A.4) = \sum_{k=1}^{s} \sum_{i=1}^{n_k} \int_0^t \left\{ \mathbf{Z}_{ki}(u) - \widetilde{\mathbf{Z}}_k(\boldsymbol{\beta}_0, u) \right\} w_{ki}(u) dN_{ki}(\boldsymbol{\beta}_0, u)$$

since $\widehat{w}_{ki}(u) dN_{ki}(\boldsymbol{\beta}_0, u) = w_{ki}(u) dN_{ki}(\boldsymbol{\beta}_0, u)$.

We can now write

$$s^{-1} \boldsymbol{U}_1(\boldsymbol{\beta}_0, t) = s^{-1} \sum_{k=1}^{s} \widetilde{\boldsymbol{U}}_{1k}(\boldsymbol{\beta}_0, t) + \boldsymbol{H}(\boldsymbol{\beta}_0, t)$$

where $\widetilde{\boldsymbol{U}}_{1k}(\boldsymbol{\beta}_0, t) = \sum_{i=1}^{n_k} \int_0^t \left\{ \mathbf{Z}_{ki}(u) - \widetilde{\mathbf{Z}}_k(\boldsymbol{\beta}_0, u) \right\} w_{ki}(u) dN_{ki}(\boldsymbol{\beta}_0, u)$, and $\boldsymbol{H}(\boldsymbol{\beta}_0, t)$ is the second part of (A.4).

A first-order Taylor expansion of $\widetilde{\mathbf{Z}}_k(\boldsymbol{\beta}_0, u) - \widehat{\mathbf{Z}}_k(\boldsymbol{\beta}_0, u)$ w.r.t. $\widehat{w}_{ki}(u)$ around $w_{ki}(u), i = 1, \ldots, n_k$ gives:

$$\widehat{\mathbf{Z}}_k(\boldsymbol{\beta}_0, u) - \widetilde{\mathbf{Z}}_k(\boldsymbol{\beta}_0, u) \approx \sum_{j=1}^{n_k} \boldsymbol{A}_j(\boldsymbol{\beta}_0, u) \{ \widehat{w}_{kj}(u) - w_{kj}(u) \}$$

where

$$\begin{aligned} \boldsymbol{A}_{kj}(\boldsymbol{\beta}_0, u) &= \frac{Y_{kj}(u) \mathbf{Z}_{kj}(u) e^{\boldsymbol{\beta}_0' \mathbf{Z}_{kj}(t)}}{n \widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta}_0, u)} - \frac{Y_{kj}(u) e^{\boldsymbol{\beta}_0' \mathbf{Z}_{kj}(t)}}{n \widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta}_0, u)} \times \frac{\widetilde{\mathbf{S}}_k^{(1)}(\boldsymbol{\beta}_0, t)}{\widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta}_0, t)} \\ &= \frac{Y_{kj}(u) e^{\boldsymbol{\beta}_0' \mathbf{Z}_{kj}(t)}}{n \widetilde{\mathbf{S}}_k^{(0)}(\boldsymbol{\beta}_0, u)} \times \left\{ \mathbf{Z}_{kj}(u) - \widetilde{\mathbf{Z}}_k(\boldsymbol{\beta}_0, u) \right\} \end{aligned}$$

Therefore,

$$\boldsymbol{H}(\boldsymbol{\beta}_0, t) = -s^{-1}\sum_{k=1}^{s}\sum_{i=1}^{n_k}\int_0^t \sum_{j=1}^{n_k} \boldsymbol{A}_{kj}(\boldsymbol{\beta}_0, u)\{\widehat{w}_{kj}(u) - w_{kj}(u)\}w_{ki}(u)dN_{ki}(u) + O_p(1/s).$$

Now write

$$\begin{aligned}\widehat{w}_{kj}(u) - w_{kj}(u) &= I(C_{kj} \geq T_{kj} \wedge u)\left\{\frac{\widehat{G}(u)}{\widehat{G}(X_{kj})} - \frac{G(u)}{G(X_{kj})}\right\} \\ &= -I(X_{xj} < u)w_{kj}(u)\int_{X_{kj}}^{u} \frac{dM_{\bullet}^c(y)}{nY^c(y)} + o_p(1),\end{aligned}$$

where $Y^c(y) = \frac{1}{n}\sum_{j=1}^{n} I(X_j \geq y) \longrightarrow \pi(y)$ and $M^c(y) = I(X \leq y, \Delta = 0) - \int_0^y I(X \geq t)d\Lambda^c(t)$ is the martingale associated with the censoring process. Since we assume no strata effect on the censoring process, a single index can be used, such that $M_{\bullet}^c(y) = \sum_{l=1}^{n} M_l^c(y)$.

This gives

$$\begin{aligned}\boldsymbol{H}(\boldsymbol{\beta}_0, t) &= s^{-1}\sum_{k=1}^{s}\sum_{i=1}^{n_k}\int_0^t \sum_{j=1}^{n_k} \boldsymbol{A}_{kj}(\boldsymbol{\beta}_0, u)I(X_{xj} < u)w_{kj}(u)\int_{X_{kj}}^{u}\frac{dM_{\bullet}^c(y)}{nY^c(y)}w_{ki}(u)dN_{ki}(u) + O_p(1/s) \\ &= s^{-1}\sum_{k=1}^{s}\sum_{i=1}^{n_k}\int_0^t \sum_{j=1}^{n_k} \boldsymbol{A}_{kj}(\boldsymbol{\beta}_0, u)w_{kj}(u)\int_0^{\infty} I(X_{kj} < y \leq u)\frac{dM_{\bullet}^c(y)}{nY^c(y)}w_{ki}(u)dN_{ki}(u) + O_p(1/s) \\ &= n^{-1}\int_0^{\infty} s^{-1}\sum_{k=1}^{s}\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}\int_0^t \boldsymbol{A}_{kj}(\boldsymbol{\beta}_0, u)I(X_{kj} < y \leq u)w_{kj}(u)w_{ki}(u)dN_{ki}(u)\frac{dM_{\bullet}^c(y)}{Y^c(y)} + O_p(1/s).\end{aligned}$$

Accordingly, $\boldsymbol{H}(\boldsymbol{\beta}_0, \infty) = n^{-1}\sum_{l=1}^{n}\int_0^{\infty} Y^c(y)^{-1}\mathbf{Q}(\boldsymbol{\beta}_0, y)dM_l^c(y) + O_p(1/s)$, where

$$\mathbf{Q}(\boldsymbol{\beta}_0, y) = s^{-1}\sum_{k=1}^{s}\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}\int_0^{\infty} \boldsymbol{A}_{kj}(\boldsymbol{\beta}_0, u)I(X_{kj} < y \leq u)w_{kj}(u)w_{ki}(u)dN_{ki}(u).$$

Since $\{N_{ki}, Y_{ki}, Z_{ki}, i = 1, \ldots, n_k, n_k\}, k = 1, \ldots, s$ are i.i.d. for highly stratified data,

$$\mathbf{Q}(\boldsymbol{\beta}_0, y) \longrightarrow \mathbf{q}(y) = \mathrm{E}\left[\sum_{i=1}^{n_k}\sum_{j=1}^{n_k}\int_0^t \boldsymbol{A}_{kj}(\boldsymbol{\beta}_0, u)I(X_{kj} < y \leq u)w_{kj}(u)w_{ki}(u)dN_{ki}(u)\right].$$

Let $\boldsymbol{\psi}_k = \sum_{i=1}^{n_k}\int_0^{\infty}\mathbf{q}(y)/\{\overline{m}\pi(y)\}dM_{ki}^c(y)$, $\boldsymbol{H}(\boldsymbol{\beta}_0, \infty) = s^{-1}\sum_{k=1}^{s}\boldsymbol{\psi}_k + O_p(1/s)$, $\boldsymbol{\eta}_k = \widetilde{\boldsymbol{U}}_{1k}(\boldsymbol{\beta}_0, t)$.

Then,

$$s^{-1}\boldsymbol{U}_1(\boldsymbol{\beta}_0, \tau) = s^{-1}\sum_{k=1}^{s}(\boldsymbol{\eta}_k + \boldsymbol{\psi}_k) + o_p(1),$$

which is approximately a sum of s i.i.d. distributed random variables. By multivariate central limit theorem, $s^{-\frac{1}{2}}\boldsymbol{U}_1(\boldsymbol{\beta}_0,\tau)$ is asymptotically normal with covariance matrix $\boldsymbol{\Sigma}_h = \mathrm{E}\left\{(\boldsymbol{\eta}_k+\boldsymbol{\psi}_k)(\boldsymbol{\eta}_k+\boldsymbol{\psi}_k)^T\right\}$. This can be estimated empirically by

$$\frac{1}{s}\sum_{k=1}^{s}(\widehat{\boldsymbol{\eta}}_k+\widehat{\boldsymbol{\psi}}_k)^{\otimes 2},$$

where $\widehat{\boldsymbol{\psi}}_k=\sum_{l=1}^{n} I(K_l=k)\int_0^\infty \widehat{\mathbf{Q}}(\widehat{\boldsymbol{\beta}},y)/\{\overline{m}Y^c(y)\}d\widehat{M}_l^c(y)$. Here $\widehat{M}^c(y)$ is defined analogously to $M^c(y)$ with $\Lambda^c(t)$ replaced by $\widehat{\Lambda}^c(t)$, and $\widehat{\mathbf{Q}}(\widehat{\boldsymbol{\beta}},y)$ is defined analogously to $\mathbf{Q}(\widehat{\boldsymbol{\beta}},y)$, with $\widetilde{\mathbf{Z}}$ replaced by $\widehat{\mathbf{Z}}$ and w_{ki} replaced by $\widehat{w}_{ki}$. Further, $\widehat{\Lambda}_j^c(t)=\int_0^t\{nY^c(u)\}^{-1}\sum_{l=1}^{n} dI(X_l\le u,\Delta_l=0)$ and $\widehat{\boldsymbol{\eta}}_k$ is defined analogously to $\widetilde{\boldsymbol{U}}_{1k}(\widehat{\boldsymbol{\beta}},\tau)$, with $\widetilde{\mathbf{Z}}$ replaced by $\widehat{\mathbf{Z}}$ and w_{ki} replaced by $\widehat{w}_{ki}$.