SUPPLEMENTARY INFORMATION

corresponding to:

"Computationally efficient inference for center effects based on restricted mean survival time"

Xin Wang, Yingchao Zhong, Purna Mukhopadhyay and Douglas E. Schaubel

May 30, 2019

1 Proofs of asymptotic results from Section 3

1.1 Regularity Conditions

We specify the following regularity conditions for $i = 1, \ldots, n$ and $j = 1, \ldots, J$.

(a) $\{\mathcal{O}_1, \ldots, \mathcal{O}_n\}$ are independently and identically distributed.

(b) $P(X_i \geq t) > 0$ for $t \in (0, \tau]$.

(c) $|Z_{ik}| < M_Z < \infty$, where Z_{ik} is the kth component of $\boldsymbol{Z}_i$.

(d) $\Lambda_{0j}^C(\tau) < \infty$ and $\Lambda_{0j}^C(t)$ is absolutely continuous for $t \in (0, \tau]$.

(e) There exists a neighborhood $\mathcal{C}$ of $\boldsymbol{\theta}$ such that for $k = 0, 1, 2, j = 1, \ldots, J$,

$$\sup_{t\in(0,\tau],\boldsymbol{\theta}\in\mathcal{C}} \left\| \boldsymbol{R}_j^{(k)}(t;\boldsymbol{\theta}) - \boldsymbol{r}_j^{(k)}(t;\boldsymbol{\theta}) \right\| \xrightarrow{p} 0,$$

where

$$\boldsymbol{R}_j^{(k)}(t;\boldsymbol{\theta}) = \frac{\sum_{i=1}^n G_{ij} \exp\{\boldsymbol{\theta}'\boldsymbol{Z}_i\} \boldsymbol{Z}_i^{\otimes k} R_i(t)}{\sum_{i=1}^n G_{ij}}, \quad (1)$$

$$\boldsymbol{r}_j^{(k)}(t;\boldsymbol{\theta}) = E\left[G_{ij} \exp\{\boldsymbol{\theta}'\boldsymbol{Z}_i\} \boldsymbol{Z}_i^{\otimes k} R_i(t)\right]. \quad (2)$$

(f) There exists a neighborhood $\mathcal{B}$ of $\boldsymbol{\beta}_0$ such that for $k = 0, 1, 2, j = 1, \ldots, J$,

$$\sup_{t\in(0,\tau],\boldsymbol{\beta}\in\mathcal{B}} \left\| \boldsymbol{S}_j^{(k)}(\boldsymbol{\beta}, \boldsymbol{W}) - \boldsymbol{s}_j^{(k)}(\boldsymbol{\beta}) \right\| \xrightarrow{p} 0,$$

where

$$\boldsymbol{s}_j^{(k)}(\boldsymbol{\beta}) = E\left[G_{ij}\Delta_i^Y W_{ij} \exp\{\boldsymbol{\beta}'\boldsymbol{Z}_i\}\boldsymbol{Z}_i^{\otimes k}\right] = E\left[G_{ij} \exp\{\boldsymbol{\beta}'\boldsymbol{Z}_i\}\boldsymbol{Z}_i^{\otimes k}\right],$$

(g) The matrices $\boldsymbol{A}(\boldsymbol{\beta}_0)$ and $\boldsymbol{\Theta}(\boldsymbol{\theta})$ are each positive definite, where

$$\boldsymbol{A}(\boldsymbol{\beta}) = \sum_{j=1}^J E\left\{G_{ij}\Delta_i^Y W_{ij} Y_i \left(\frac{\boldsymbol{s}_j^{(2)}(\boldsymbol{\beta})}{s_j^{(0)}(\boldsymbol{\beta})} - \overline{\boldsymbol{s}}_j(\boldsymbol{\beta})^{\otimes 2}\right)\right\},$$

$$\boldsymbol{\Theta}(\boldsymbol{\theta}) = \sum_{j=1}^J E\left\{G_{ij} \int_0^\tau \left(\frac{\boldsymbol{r}_j^{(2)}(t;\boldsymbol{\theta})}{r_j^{(0)}(t;\boldsymbol{\theta})} - \overline{\boldsymbol{r}}_j(t;\boldsymbol{\theta})^{\otimes 2}\right) \boldsymbol{r}_j^{(0)}(t;\boldsymbol{\theta})\lambda_{0j}^C(t)dt\right\},$$

with $\overline{\boldsymbol{s}}_j(\boldsymbol{\beta}) = s_j^{(0)}(\boldsymbol{\beta})^{-1} \boldsymbol{s}_j^{(1)}(\boldsymbol{\beta})$ and $\overline{\boldsymbol{r}}_j(t;\boldsymbol{\theta}) = r_j^{(0)}(t;\boldsymbol{\theta})^{-1} \boldsymbol{r}_j^{(1)}(t;\boldsymbol{\theta})$.

1.2 Unbiasedness of Estimating Equation

The unbiasedness of estimating equations $\boldsymbol{\phi}_1(\boldsymbol{\beta}, \boldsymbol{\mu}_0, \widehat{\boldsymbol{W}})$ and $\boldsymbol{\phi}_2(\boldsymbol{\beta}, \boldsymbol{\mu}_0, \widehat{\boldsymbol{W}})$ follows from the consistency of the IPCW weights $\widehat{W}_{ij}$ for W_{ij},and the fact that $\boldsymbol{\phi}_1(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})$ has mean zero. We then obtain

$$\begin{aligned}
& E\left[G_{ij} W_{ij} \Delta_i^Y \left\{Y_i - \mu_{0j} \exp\{\boldsymbol{\beta}' \boldsymbol{Z}_i\}\right\} \boldsymbol{Z}_i\right] \\
= & E\left(E\left[G_{ij} \frac{I\left(D_i \wedge L \leq C_i\right)}{P\left(C_i > Y_i\right)} \left\{Y_i - \mu_{0j} \exp\{\boldsymbol{\beta}' \boldsymbol{Z}_i\}\right\} \boldsymbol{Z}_i \middle| \boldsymbol{Z}_i\right]\right) \\
= & E\left\{E\left(E\left[G_{ij} \frac{I\left(D_i \wedge L \leq C_i\right)}{P\left(C_i > Y_i\right)} \left\{Y_i - \mu_{0j} \exp\{\boldsymbol{\beta}' \boldsymbol{Z}_i\}\right\} \boldsymbol{Z}_i | \boldsymbol{Z}_i, D_i\right] | \boldsymbol{Z}_i\right)\right\} \\
= & E\left\{E\left(E\left[G_{ij} \frac{I\left(D_i \wedge L \leq C_i\right)}{P\left(C_i > D_i \wedge L\right)} \left\{D_i \wedge L - \mu_{0j} \exp\{\boldsymbol{\beta}' \boldsymbol{Z}_i\}\right\} \boldsymbol{Z}_i | \boldsymbol{Z}_i, D_i\right] | \boldsymbol{Z}_i\right)\right\} \\
= & E\left(E\left[G_{ij} \left\{D_i \wedge L - \mu_{0j} \exp\{\boldsymbol{\beta}' \boldsymbol{Z}_i\}\right\} \boldsymbol{Z}_i | \boldsymbol{Z}_i\right]\right) \\
= & E\left[G_{ij} \left\{E\left(D_i \wedge L | \boldsymbol{Z}_i\right) - \mu_{0j} \exp\{\boldsymbol{\beta}' \boldsymbol{Z}_i\}\right\} \boldsymbol{Z}_i\right] \\
= & \mathbf{0}.
\end{aligned}$$

Since $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2$ are equivalent to $\boldsymbol{\psi}_1, \boldsymbol{\psi}_2$, respectfully, the unbiasedness of the latter also holds.

1.3 Consistency of $\widehat{\boldsymbol{\beta}}$

Using the Inverse Function Theorem, we require the following conditions in order to establish the consistency of $\widehat{\boldsymbol{\beta}}$:

(a) $\partial \boldsymbol{\phi}_1(\boldsymbol{\beta}, \widehat{\boldsymbol{W}}) / \partial \boldsymbol{\beta}'$ exists and is continuous in an open neighborhood of $\boldsymbol{\beta}_0$.

(b) $-n^{-1} \partial \boldsymbol{\phi}_1(\boldsymbol{\beta}, \widehat{\boldsymbol{W}}) / \partial \boldsymbol{\beta}'$ is positive definite at $\boldsymbol{\beta}_0$ with probability 1 as $n \to \infty$.

(c) $-n^{-1} \partial \boldsymbol{\phi}_1(\boldsymbol{\beta}, \widehat{\boldsymbol{W}}) / \partial \boldsymbol{\beta}'$ converges in probability to a fixed function uniformly in an open neighborhood of $\boldsymbol{\beta}_0$.

(d) Asymptotic unbiasedness of the estimating function: $-n^{-1} \boldsymbol{\phi}_1(\boldsymbol{\beta}, \widehat{\boldsymbol{W}}) \xrightarrow{p} \mathbf{0}$.

The only quantity in $\boldsymbol{\phi}_1(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})$ which is a function of $\boldsymbol{\beta}$ is $\{\boldsymbol{Z}_i - \overline{\boldsymbol{S}}_j(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})\}$, and we have

$$\begin{aligned}
& \frac{\partial \left\{\boldsymbol{Z}_i - \overline{\boldsymbol{S}}_j(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})\right\}}{\partial \boldsymbol{\beta}'} \\
= & \frac{\partial}{\partial \boldsymbol{\beta}'} \frac{\sum_{k=1}^n G_{kj} \Delta_k^Y \widehat{W}_{kj} \boldsymbol{Z}_k \exp\{\boldsymbol{\beta}' \boldsymbol{Z}_k\}}{\sum_{k=1}^n G_{kj} \Delta_k^Y \widehat{W}_{kj} \exp\{\boldsymbol{\beta}' \boldsymbol{Z}_k\}} \\
= & \frac{\boldsymbol{S}_j^{(2)}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}}) S_j^{(0)}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}}) - \boldsymbol{S}_j^{(1)}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})^{\otimes 2}}{S_j^{(0)}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})^2} \\
= & \frac{\boldsymbol{s}_j^{(2)}(\boldsymbol{\beta}) s_j^{(0)}(\boldsymbol{\beta}) - \boldsymbol{s}_j^{(1)}(\boldsymbol{\beta})^{\otimes 2}}{s_j^{(0)}(\boldsymbol{\beta})^2} + o_p(1).
\end{aligned}$$

Thus, the derivative of $\boldsymbol{\psi}_1$ is given by

$$\begin{aligned}
-\frac{1}{n} \frac{\boldsymbol{\psi}_1(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})}{\partial \boldsymbol{\beta}'} & = -\frac{1}{n} \sum_{j=1}^J \sum_{i=1}^n G_{ij} \Delta_i^Y \widehat{W}_{ij} Y_i \frac{\partial \left\{\boldsymbol{Z}_i - \overline{\boldsymbol{S}}_j(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})\right\}}{\partial \boldsymbol{\beta}'} \\
& = \frac{1}{n} \sum_{j=1}^J \sum_{i=1}^n G_{ij} \Delta_i^Y \widehat{W}_{ij} Y_i \frac{\boldsymbol{s}_j^{(2)}(\boldsymbol{\beta}) s_j^{(0)}(\boldsymbol{\beta}) - \boldsymbol{s}_j^{(1)}(\boldsymbol{\beta})^{\otimes 2}}{s_j^{(0)}(\boldsymbol{\beta})^2} + o(1) \\
& = \boldsymbol{A}(\boldsymbol{\beta}) + o(1),
\end{aligned}$$

such that Condition (b) holds. Further, Condition (d) holds by the Weak Law of Law of Numbers. Since the remaining conditions hold under mild regularity conditions, it follows that $\widehat{\boldsymbol{\beta}}$ converges in probability to $\boldsymbol{\beta}_0$.

1.4 Asymptotic Distribution of $\widehat{\boldsymbol{\beta}}$

We first express $\boldsymbol{\psi}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})$ as a sum of independent and identically distributed terms. We replace Y_i in $\boldsymbol{\psi}_1$ with $Y_i - \mu_{ij}$, per the following relationship:

$$
\begin{aligned}
& \sum_{j=1}^{J}\sum_{i=1}^{n} G_{ij}\left\{\boldsymbol{Z}_i - \overline{\boldsymbol{S}}_j(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})\right\}\widehat{W}_{ij}\Delta_i^Y \mu_{ij} \\
= \; & \sum_{j=1}^{J}\sum_{i=1}^{n} G_{ij}\left\{\boldsymbol{Z}_i - \frac{\boldsymbol{S}_j^{(1)}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})}{S_j^{(0)}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})}\right\}\widehat{W}_{ij}\Delta_i^Y \mu_{0j}\exp\{\boldsymbol{\beta}'\boldsymbol{Z}_i\} \\
= \; & \sum_{j=1}^{J} n_j \frac{\boldsymbol{S}_j^{(1)}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})S_j^{(0)}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}}) - \boldsymbol{S}_j^{(1)}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})S_j^{(0)}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})}{S_j^{(0)}(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})} \\
= \; & \boldsymbol{0} \\
\Rightarrow \; & \boldsymbol{\psi}_1(\boldsymbol{\beta}, \boldsymbol{\mu}_0, \widehat{\boldsymbol{W}}) = \sum_{j=1}^{J}\sum_{i=1}^{n} G_{ij}\left\{\boldsymbol{Z}_i - \overline{\boldsymbol{S}}_j(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})\right\}\widehat{W}_{ij}\Delta_i^Y Y_i \\
= \; & \sum_{j=1}^{J}\sum_{i=1}^{n} G_{ij}\left\{\boldsymbol{Z}_i - \overline{\boldsymbol{S}}_j(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})\right\}\widehat{W}_{ij}\Delta_i^Y \left(Y_i - \mu_{ij}\right).
\end{aligned}
$$

The main estimating equation $\boldsymbol{\psi}_1(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})$ can be further separated as the sum of the following three components,

$$
\begin{aligned}
\boldsymbol{\psi}_1(\boldsymbol{\beta}, \widehat{\boldsymbol{W}}) \; = \; & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{J} G_{ij}\left\{\boldsymbol{Z}_i - \overline{\boldsymbol{S}}_j\left(\boldsymbol{\beta}, \boldsymbol{W}\right)\right\}W_{ij}\Delta_i^Y\left(Y_i - \mu_{ij}\right) && (3) \\
+ \; & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{J} G_{ij}\left\{\boldsymbol{Z}_i - \overline{\boldsymbol{S}}_j\left(\boldsymbol{\beta}, \boldsymbol{W}\right)\right\}\left(\widehat{W}_{ij} - W_{ij}\right)\Delta_i^Y\left(Y_i - \mu_{ij}\right) && (4) \\
+ \; & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{J} G_{ij}\left\{\overline{\boldsymbol{S}}_j(\boldsymbol{\beta}, \boldsymbol{W}) - \overline{\boldsymbol{S}}_j(\boldsymbol{\beta}, \widehat{\boldsymbol{W}})\right\}\widehat{W}_{ij}\Delta_i^Y\left(Y_i - \mu_{ij}\right). && (5)
\end{aligned}
$$

Through techniques from empirical processes, (5) can be shown to converge in probability to 0, while (3) can be shown to be asymptotically equivalent to

$$
(3) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{J} G_{ij}\left\{\boldsymbol{Z}_i - \overline{\boldsymbol{s}}_j\left(\boldsymbol{\beta}, \boldsymbol{W}\right)\right\}W_{ij}\Delta_i^Y\left(Y_i - \mu_{ij}\right).
$$

Letting $\boldsymbol{\epsilon}_i(\boldsymbol{\beta}, \boldsymbol{W}) = G_{ij}\{\boldsymbol{Z}_i - \overline{\boldsymbol{S}}_j(\boldsymbol{\beta}, \boldsymbol{W})\}W_{ij}\Delta_i^Y(Y_i - \mu_{ij})$, (4) can be written as,

$$
\begin{aligned}
(4) \; = \; & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{J}\boldsymbol{\epsilon}_i(\boldsymbol{\beta}, \boldsymbol{W})\frac{\widehat{W}_{ij} - W_{ij}}{W_{ij}} \\
= \; & \frac{1}{\sqrt{n}^3}\sum_{i=1}^{n}\sum_{j=1}^{J}\sum_{k=1}^{n}\sum_{l=1}^{J} G_{ij}G_{kl}\boldsymbol{\epsilon}_i(\boldsymbol{\beta}, \boldsymbol{W})\boldsymbol{D}_i(\boldsymbol{\theta})'\boldsymbol{\Theta}(\boldsymbol{\theta})^{-1}\boldsymbol{U}_{kl}(\boldsymbol{\theta}) && (6) \\
+ \; & \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{J}\sum_{k=1}^{n}\frac{1}{n_j}G_{ij}G_{kj}\boldsymbol{\epsilon}_i(\boldsymbol{\beta}, \boldsymbol{W})J_{ik}(\boldsymbol{\theta}) + o_p(1), && (7)
\end{aligned}
$$

where we define

$$\begin{aligned}\boldsymbol{D}_i(\boldsymbol{\theta}) &= \sum_{j=1}^{J} G_{ij} \int_0^L \{\boldsymbol{Z}_i - \overline{\boldsymbol{r}}_j(u;\boldsymbol{\theta})\} R_i(u) d\Lambda_{ij}^C(u), \\ J_{ik}(\boldsymbol{\theta}) &= \sum_{j=1}^{J} G_{ij} G_{kj} \int_0^L r_j^{(0)}(u;\boldsymbol{\theta})^{-1} \exp(\boldsymbol{\theta}'\boldsymbol{Z}_i) R_i(u) dM_k^C(u).\end{aligned}$$

The first component can be simplified as

$$(6) = \frac{1}{\sqrt{n}} \boldsymbol{K}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{W}) \boldsymbol{\Theta}(\boldsymbol{\theta})^{-1} \sum_{i=1}^{n}\sum_{j=1}^{J} G_{ij} \boldsymbol{U}_{ij}(\boldsymbol{\theta}) + o_p(1),$$

with the following additional definitions,

$$\boldsymbol{K}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{W}) = \sum_{j=1}^{J} \boldsymbol{K}_j(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{W}) = \sum_{j=1}^{J} E\left\{G_{ij} \boldsymbol{\epsilon}_i(\boldsymbol{\beta},\boldsymbol{W}) \boldsymbol{D}_i(\boldsymbol{\theta})'\right\}.$$

Let $\boldsymbol{H}_j(t;\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{W}) = E\{\boldsymbol{\epsilon}_i(\boldsymbol{\beta},\boldsymbol{W}) \exp\{\boldsymbol{\theta}'\boldsymbol{Z}_i\} R_i(t)\}$, such that (7) can be represented as

$$\begin{aligned}(7) &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n}\sum_{j=1}^{J} G_{kj} \int_0^L \frac{\sum_{i=1}^{n} G_{ij} \boldsymbol{\epsilon}_i(\boldsymbol{\beta},\boldsymbol{W}) \exp\{\boldsymbol{\theta}'\boldsymbol{Z}_i\} R_i(u)}{n_j r_j^{(0)}(u;\boldsymbol{\theta})} dM_{kj}^C(u) + o_p(1) \\ &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n}\sum_{j=1}^{J} G_{kj} \int_0^L \frac{\boldsymbol{H}_j(u;\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{W})}{r_j^{(0)}(u;\boldsymbol{\theta})} dM_{kj}^C(u) + o_p(1).\end{aligned}$$

Combining (6) and (7), we can write (4) as

$$\frac{1}{\sqrt{n}}(4) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n}\sum_{j=1}^{J} G_{ij} \left\{ \boldsymbol{K}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{W}) \boldsymbol{\Theta}(\boldsymbol{\theta})^{-1} \boldsymbol{U}_{ij}(\boldsymbol{\theta}) + \int_0^L \frac{\boldsymbol{H}_j(u;\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{W})}{r_j^{(0)}(u;\boldsymbol{\theta})} dM_{ij}^C(u) \right\},$$

asymptotically. In summary, our estimating equation $n^{-1/2}\boldsymbol{\psi}_1(\boldsymbol{\beta},\widehat{\boldsymbol{W}})$ can be written as a sum of n i.i.d. terms at a difference of $o_p(1)$; i.e., $n^{-1/2}\boldsymbol{\psi}_1(\boldsymbol{\beta},\widehat{\boldsymbol{W}}) = n^{-1/2}\sum_{i=1}^{n}\sum_{j=1}^{J} G_{ij}\boldsymbol{b}_{ij}(\boldsymbol{\beta},\widehat{\boldsymbol{W}})$, where

$$\begin{aligned}\boldsymbol{b}_{ij}(\boldsymbol{\beta},\boldsymbol{W}) &= \{\boldsymbol{Z}_i - \overline{\boldsymbol{s}}_j(\boldsymbol{\beta},\boldsymbol{W})\} W_{ij} \Delta_i^Y (Y_i - \mu_{ij}) + \boldsymbol{K}(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{W}) \boldsymbol{\Theta}(\boldsymbol{\theta})^{-1} \boldsymbol{U}_{ij}(\boldsymbol{\theta}) \\ &\quad + \int_0^L \frac{\boldsymbol{H}_j(u;\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{W})}{r_j^{(0)}(u;\boldsymbol{\theta})} dM_{ij}^C(u).\end{aligned}$$

Next, we consider the asymptotic distribution of $\boldsymbol{\psi}_1$. As such,

$$\frac{1}{\sqrt{n}} \boldsymbol{\psi}_1(\boldsymbol{\beta},\widehat{\boldsymbol{W}}) \xrightarrow{D} \text{Normal}\left(\boldsymbol{0}, \boldsymbol{B}(\boldsymbol{\beta},\boldsymbol{W})\right),$$

where

$$\boldsymbol{B}(\boldsymbol{\beta},\boldsymbol{W}) = E\left[\sum_{j=1}^{J} G_{ij} \{\boldsymbol{b}_{ij}(\boldsymbol{\beta},\boldsymbol{W})\}^{\otimes 2}\right].$$

With respect to the asymptotic distribution of $\widehat{\boldsymbol{\beta}}$, a Taylor expansion of $\boldsymbol{\psi}_1(\boldsymbol{\beta},\widehat{\boldsymbol{W}})$ around $\boldsymbol{\beta}$ yields

$$\begin{aligned}&\boldsymbol{0} = \boldsymbol{\psi}_1(\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{W}}) \approx \boldsymbol{\psi}_1(\boldsymbol{\beta},\widehat{\boldsymbol{W}}) + \frac{\partial \boldsymbol{\psi}_1(\boldsymbol{\beta},\widehat{\boldsymbol{W}})}{\partial \boldsymbol{\beta}'} \left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \\ \Rightarrow\quad &\sqrt{n}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \left[\frac{1}{n} \frac{\partial \boldsymbol{\psi}_1(\boldsymbol{\beta},\widehat{\boldsymbol{W}})}{\partial \boldsymbol{\beta}'}\right]^{-1} \frac{1}{\sqrt{n}} \boldsymbol{\psi}_1(\boldsymbol{\beta},\widehat{\boldsymbol{W}}) = \boldsymbol{A}(\boldsymbol{\beta})^{-1} \frac{1}{\sqrt{n}} \boldsymbol{\psi}_1(\boldsymbol{\beta},\widehat{\boldsymbol{W}}) + o\left(\frac{1}{\sqrt{n}}\right).\end{aligned}$$

Finally, applying the Delta Method,

$$\sqrt{n}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{D} \text{Normal}\left(\boldsymbol{0}, \boldsymbol{A}(\boldsymbol{\beta})^{-1} \boldsymbol{B}(\boldsymbol{\beta},\boldsymbol{W}) \boldsymbol{A}(\boldsymbol{\beta})^{-1}\right).$$

1.5 Asymptotic distributions of $\widehat{\boldsymbol{\mu}}_0$ and $\widehat{\boldsymbol{\eta}}$

Relating $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\mu}}_0$, we define the functions:

$$\overline{d}_j(\boldsymbol{\beta}) = \frac{\sum_{i=1}^n G_{ij}\widehat{W}_{ij}\Delta_i^Y Y_i}{\sum_{i=1}^n G_{ij}\widehat{W}_{ij}\Delta_i^Y \exp\{\boldsymbol{\beta}'\boldsymbol{Z}_i\}},$$

$$d_j(\boldsymbol{\beta}) = \lim_{n\to 0}\overline{d}_j(\boldsymbol{\beta}) = \frac{E\left\{G_{ij}\widehat{W}_{ij}\Delta_i^Y Y_i\right\}}{E\left\{G_{ij}\widehat{W}_i\Delta_i^Y \exp(\boldsymbol{\beta}'\boldsymbol{Z}_i)\right\}} = \frac{\mu_{0j}E\left\{\exp(\boldsymbol{\beta}'\boldsymbol{Z}_i)\right\}}{E\left\{\exp(\boldsymbol{\beta}'\boldsymbol{Z}_i)\right\}},$$

and $\overline{\boldsymbol{d}}(\boldsymbol{\beta}) = [\overline{d}_1(\boldsymbol{\beta}),\ldots,\overline{d}_J(\boldsymbol{\beta})]'$, $\boldsymbol{d}(\boldsymbol{\beta}) = [d_1(\boldsymbol{\beta}),\ldots,d_J(\boldsymbol{\beta})]'$. Then $\widehat{\boldsymbol{\mu}}_0 = \overline{\boldsymbol{d}}(\widehat{\boldsymbol{\beta}})$ and $\boldsymbol{\mu}_0 = \boldsymbol{d}(\boldsymbol{\beta})$. The difference between the estimated and true value of μ_{0j} can be written as:

$$\begin{aligned}
\sqrt{n}\left(\widehat{\boldsymbol{\mu}}_0 - \boldsymbol{\mu}_0\right) &= \sqrt{n}\left\{\overline{\boldsymbol{d}}(\widehat{\boldsymbol{\beta}}) - \overline{\boldsymbol{d}}(\boldsymbol{\beta})\right\} + \sqrt{n}\left\{\overline{\boldsymbol{d}}(\boldsymbol{\beta}) - \boldsymbol{d}(\boldsymbol{\beta})\right\}\\
&= \frac{\partial\overline{\boldsymbol{d}}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}\sqrt{n}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) + \sqrt{n}\left\{\overline{\boldsymbol{d}}(\boldsymbol{\beta}) - \boldsymbol{d}(\boldsymbol{\beta})\right\} + o_p(1),
\end{aligned}$$

where $\partial\overline{\boldsymbol{d}}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}$ is a $J\times p$ matrix with jth row given by

$$\begin{aligned}
\left(\frac{\partial\overline{\boldsymbol{d}}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}\right)_{j*} &= -\frac{\left\{\sum_{i=1}^n G_{ij}\widehat{W}_{ij}\Delta_i^Y Y_i\right\}\left\{\sum_{i=1}^n G_{ij}\widehat{W}_{ij}\Delta_i^Y \exp(\boldsymbol{\beta}'\boldsymbol{Z}_i)\boldsymbol{Z}_i'\right\}}{\left\{\sum_{i=1}^n G_{ij}\widehat{W}_{ij}\Delta_i^Y \exp(\boldsymbol{\beta}'\boldsymbol{Z}_i)\right\}^2}\\
&= -\frac{\left[\mu_{0j}E\left\{G_{ij}\exp(\boldsymbol{\beta}'\boldsymbol{Z}_i)\right\}\right]\left[E\left\{G_{ij}\exp(\boldsymbol{\beta}'\boldsymbol{Z}_i)\boldsymbol{Z}_i'\right\}\right]}{\left[E\left\{G_{ij}\exp(\boldsymbol{\beta}'\boldsymbol{Z}_i)\right\}\right]^2} + o_p(1)\\
&= -\mu_{0j}\overline{\boldsymbol{s}}_j(\boldsymbol{\beta})' + o_p(1).
\end{aligned}$$

Thus, we can write $\partial\overline{\boldsymbol{d}}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}$ as $-\boldsymbol{\mu}_M\overline{\boldsymbol{s}}(\boldsymbol{\beta})$, where

$$\begin{aligned}
\boldsymbol{M} &= \begin{pmatrix} \mu_{01} & \ldots & 0\\ \vdots & \ddots & \vdots\\ 0 & \ldots & \mu_{0J}\end{pmatrix},\\
\overline{\boldsymbol{s}}(\boldsymbol{\beta}) &= \begin{pmatrix} \overline{\boldsymbol{s}}_1(\boldsymbol{\beta})'\\ \vdots\\ \overline{\boldsymbol{s}}_J(\boldsymbol{\beta})'\end{pmatrix}.
\end{aligned}$$

Now, the jth element the second part is given by

$$\begin{aligned}
\sqrt{n}\left\{\overline{\boldsymbol{d}}(\boldsymbol{\beta}) - \boldsymbol{d}(\boldsymbol{\beta})\right\}_j &= \frac{\sqrt{n}}{n_j}\frac{\sum_{i=1}^n G_{ij}\widehat{W}_{ij}\Delta_i^Y\left\{Y_i - \mu_{0j}\exp(\boldsymbol{\beta}'\boldsymbol{Z}_i)\right\}}{\frac{1}{n_j}\sum_{i=1}^n G_{ij}\widehat{W}_{ij}\Delta_i^Y\exp(\boldsymbol{\beta}'\boldsymbol{Z}_i)}\\
&= n^{-1/2}\widehat{\rho}_j^{-1}\sum_{i=1}^n\frac{G_{ij}\widehat{W}_{ij}\Delta_i^Y}{s_j^{(0)}(\boldsymbol{\beta})}\left(Y_i - \mu_{ij}\right)\\
&\xrightarrow{p} n^{-1/2}\rho_j^{-1}s_j^{(0)}(\boldsymbol{\beta})^{-1}\sum_{i=1}^n G_{ij}W_i\Delta_i^Y\left(Y_i - \mu_{ij}\right),
\end{aligned}$$

where $\widehat{\rho}_j = n_j/n$ and $\rho_j = E[g_i]$. Therefore, we obtain

$$\begin{aligned}
\sqrt{n}\left(\widehat{\boldsymbol{\mu}}_0 - \boldsymbol{\mu}_0\right) &= \frac{1}{\sqrt{n}}\sum_{i=1}^n\left\{\begin{pmatrix}\{\rho_1 s_1^{(0)}(\boldsymbol{\beta})\}^{-1}G_{i1}W_i\Delta_i^Y\left(Y_i - \mu_{ij}\right)\\ \vdots\\ \{\rho_J s_J^{(0)}(\boldsymbol{\beta})\}^{-1}G_{iJ}W_i\Delta_i^Y\left(Y_i - \mu_{ij}\right)\end{pmatrix}\right.\\
&\quad\left. - \sum_{j=1}^J G_{ij}\boldsymbol{M}\overline{\boldsymbol{s}}(\boldsymbol{\beta})\boldsymbol{b}_i\left(\boldsymbol{\beta},\boldsymbol{W}\right)\right\} + o_p(1).
\end{aligned}$$

Let

$$\boldsymbol{v}_{\mu i}(\boldsymbol{\beta}, \boldsymbol{W}) = \begin{pmatrix} \{\rho_1 s_1^{(0)}(\boldsymbol{\beta})\}^{-1} G_{i1} W_i \Delta_i^Y (Y_i - \mu_{ij}) \\ \vdots \\ \{\rho_J s_J^{(0)}(\boldsymbol{\beta})\}^{-1} G_{iJ} W_i \Delta_i^Y (Y_i - \mu_{ij}) \end{pmatrix} - \sum_{j=1}^{J} G_{ij} \boldsymbol{M} \overline{\boldsymbol{s}}(\boldsymbol{\beta}) \boldsymbol{b}_{ij}(\boldsymbol{\beta}, \boldsymbol{W}),$$

then

$$\sqrt{n}\,(\widehat{\boldsymbol{\mu}}_0 - \boldsymbol{\mu}_0) \xrightarrow{D} \text{Normal}\left(0, E\left\{\boldsymbol{v}_{\mu i}(\boldsymbol{\beta}, \boldsymbol{W})^{\otimes 2}\right\}\right).$$

Next, connecting $\widehat{\boldsymbol{\mu}}_0$ to $\widehat{\boldsymbol{\eta}}$, we rescale $\widehat{\boldsymbol{\mu}}_0$ to $\widehat{\boldsymbol{\eta}} = \boldsymbol{\mu}/(\boldsymbol{\mu}'\boldsymbol{w})$ with a pre-specified weight vector $\boldsymbol{w}$. Define a new function $\boldsymbol{\zeta}(\boldsymbol{\mu}_0) = \boldsymbol{\mu}_0\{\boldsymbol{\mu}_0'\boldsymbol{w}\}^{-1}$. Then $\boldsymbol{\eta} = \boldsymbol{\zeta}(\boldsymbol{\mu}_0)$ and $\widehat{\boldsymbol{\eta}} = \boldsymbol{\zeta}(\widehat{\boldsymbol{\mu}}_0)$. Since

$$\frac{\partial \boldsymbol{\zeta}(\boldsymbol{\mu}_0)}{\partial \boldsymbol{\mu}_0} = \frac{\boldsymbol{\mu}_0'\boldsymbol{w}\boldsymbol{I}_J - \boldsymbol{\mu}_0\boldsymbol{w}'}{(\boldsymbol{\mu}_0'\boldsymbol{w})^2}.$$

By the Delta Method,

$$\begin{aligned} \sqrt{n}\,(\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta}) &= \sqrt{n}\,\{\boldsymbol{\zeta}(\widehat{\boldsymbol{\mu}}_0) - \boldsymbol{\zeta}(\boldsymbol{\mu}_0)\} \\ &\xrightarrow{D} \text{Normal}\left(\boldsymbol{0}, \frac{\boldsymbol{\mu}_0'\boldsymbol{w}\boldsymbol{I}_J - \boldsymbol{\mu}_0\boldsymbol{w}'}{(\boldsymbol{\mu}_0'\boldsymbol{w})^2} E\left\{\boldsymbol{v}_{\mu i}(\boldsymbol{\beta}, \boldsymbol{W})^{\otimes 2}\right\} \frac{\boldsymbol{\mu}_0'\boldsymbol{w}\boldsymbol{I}_J - \boldsymbol{w}\boldsymbol{\mu}_0'}{(\boldsymbol{\mu}_0'\boldsymbol{w})^2}\right). \end{aligned}$$

.

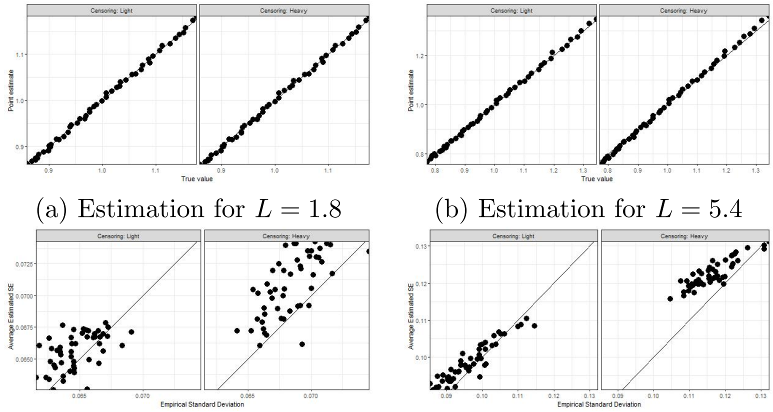
2 Simulation results for small center sizes

j	n_j	BIAS	ESD	ASE	CP(%)
1	25	0.027	0.183	0.176	93
3	25	0.020	0.171	0.173	93
5	25	0.018	0.175	0.170	92
7	25	0.006	0.172	0.170	94
9	25	0.011	0.166	0.169	94
11	50	0.010	0.120	0.123	94
13	50	0.014	0.119	0.122	95
15	50	0.009	0.117	0.121	95
17	50	-0.001	0.116	0.120	95
19	50	0.000	0.119	0.120	95
21	50	0.000	0.116	0.117	94
23	50	0.003	0.119	0.120	95
25	50	0.008	0.117	0.117	93
27	50	0.010	0.116	0.118	94
29	50	-0.002	0.116	0.117	94
31	50	0.001	0.112	0.116	95
33	50	-0.000	0.111	0.116	95
35	50	-0.011	0.111	0.116	94
37	50	-0.006	0.116	0.117	95
39	50	-0.007	0.108	0.113	95
41	25	-0.035	0.153	0.151	90
43	25	-0.027	0.156	0.146	90
45	25	-0.032	0.157	0.147	89
47	25	-0.017	0.147	0.146	92
49	25	-0.030	0.157	0.149	89

3 Simulation results for traditional method

L	Censoring	Parameter	n	BIAS	ESD	ASE	CP(%)
1.8	$\approx$ 15%	β_1 = -0.1315	2500	0.002	0.010	0.009	94
			5000	0.001	0.007	0.007	95
			10000	0.001	0.005	0.005	94
		β_2 = 0.2644	2500	-0.001	0.010	0.010	95
			5000	0.001	0.006	0.007	96
			10000	0.001	0.005	0.005	94
	$\approx$ 30%	β_1 = -0.1315	2500	0.003	0.010	0.010	94
			5000	0.002	0.007	0.007	95
			10000	0.001	0.005	0.005	96
		β_2 = -0.2644	2500	0.002	0.010	0.010	94
			5000	0.001	0.007	0.007	96
			10000	0.001	0.005	0.005	94
5.4	$\approx$ 15%	β_1 = -0.2273	2500	0.004	0.014	0.014	93
			5000	0.002	0.010	0.010	95
			10000	0.002	0.007	0.007	95
		β_2 = -0.4559	2500	0.003	0.013	0.014	95
			5000	0.002	0.009	0.010	96
			10000	0.001	0.007	0.007	94
	$\approx$ 30%	β_1 = -0.2273	2500	0.016	0.018	0.018	83
			5000	0.009	0.013	0.013	87
			10000	0.005	0.009	0.010	90
		β_2 = -0.4559	2500	0.004	0.016	0.016	94
			5000	0.003	0.011	0.012	95
			10000	0.002	0.008	0.008	94

Figure 1: Simulation results for conventional method: $\widehat{\boldsymbol{\eta}}$ for $L = 1.8$ and $L = 5.4$



(a) Estimation for $L = 1.8$ (b) Estimation for $L = 5.4$

(c) Standard Error for $L = 1.8$ (d) Standard Error for $L = 5.4$

4 R code for proposed methods

```
computeWY = function(f, dat, ZC){
  N = nrow(dat)
  J = length(unique(dat$center))
  ModC = coxph(formula = f, data = data.frame(dat))
  gammaCHat = matrix(coef(ModC), ncol = 1)
  BaseCumHaz = data.frame(basehaz(ModC, centered = F), center = 1)
  BaseCumHaz = BaseCumHaz[order(BaseCumHaz$time, decreasing = T),]
  tmp = levels(BaseCumHaz$strata)
  for(j in 1:J){
    BaseCumHaz[BaseCumHaz$strata == tmp[j], "center"] = j
  }
  WYhat = vector(mode = "numeric", length = N)
  for(j in 1:J){
    WhichObs = which(dat[, "center"] == j)
    CumHazi = BaseCumHaz[BaseCumHaz[, "center"] == j,
                         c("hazard", "time")]
    if(min(CumHazi$time) > 0)
      CumHazi = rbind(CumHazi, c(0, 0))
    WYhat[WhichObs] = sapply(dat[WhichObs, "Y"],
                             function(t){
                               CumHazi[which(CumHazi[, "time"] <= t)[1],
                                       "hazard"]
                             })
  }
  dat = cbind(dat,
              WYhat = exp(WYhat * exp(as.matrix(dat[, ZC]) %*% gammaCHat)) *
                dat[, "DeltaY"],
              newWt = exp(WYhat * exp(as.matrix(dat[, ZC]) %*% gammaCHat)) *
                dat[, "DeltaY"] * dat[, "Y"])
  return (dat)
}
```

```
compute_se = function(dat, ZD, compute_eta = T){
  J = length(unique(dat$center))
  dat = as.matrix(dat)
  r = dat[, "WYhat"] * dat[, "HRHat"]
  zbar0 = aggregate(r, by = list(dat[, "center"]), mean)[, -1]
  zbar1 = aggregate(dat[, ZD] * r, by = list(dat[, "center"]), mean)[, -1]
  zbar2 = aggregate(data.frame(t(apply(as.matrix(dat[, ZD]), 1, tcrossprod))) * r,
                    by = list(dat[, "center"]), mean)[, -1]
  ### epsilon
  zbar10 = zbar1 / zbar0
  epsilon = data.frame(dat[, ZD] - zbar10[dat[, "center"], ]) *
    dat[, "WYhat"] * (dat[, "Y"] - mujHat[dat[, "center"]] * dat[, "HRHat"])
  ### outside matrix A: zbar2*zbar0 - zbar1^2 / zbar0^2
  r = (zbar2 * zbar0 - t(apply(as.matrix(zbar1), 1, tcrossprod))) / zbar0^2
  A = colSums(r[(dat[, "center"]),] * dat[, "newWt"], na.rm = T) / N

  ### middle matrix B: the fraction part: zbar1/zbar0
  B1 = rowMeans(apply(epsilon, 1, tcrossprod), na.rm = T)
  Ainv = solve(matrix(A, sqrt(length(A))))
  AsyVar = Ainv %*% matrix(B1, sqrt(length(A))) %*% Ainv
  SEbeta = sqrt(diag(AsyVar / N))
  # save.image(paste(proj_path, "/data/result.rda", sep = ""))

  center_ind = paste("center", 1:J, sep = "")
  dat = cbind(dat,model.matrix(~factor(dat[,"center"])+0))
  colnames(dat)[42: (42+J-1)] <- center_ind
  r = dat[, "WYhat"] * (dat[, center_ind] *
       ((dat[, "Y"] - dat[, "HRHat"] %*% matrix(mujHat, nrow = 1)) %*%
       matrix(N / (njs * zbar0), J, J))) -
    as.matrix(epsilon) %*% t(diag(mujHat, J, J) %*% as.matrix(zbar10) %*% Ainv)
  Vmu = matrix(rowMeans(apply(r, 1, tcrossprod), na.rm = T), J)
  SEmu = sqrt(diag(Vmu / N))

  if(compute_eta){
    r = (diag(sum(mujHat * ws), J, J) - matrix(mujHat, J, 1) %*% matrix(ws, 1, J)) /
      sum(mujHat * ws)^2
    SEeta = sqrt(diag(r %*% Vmu %*% t(r) / N))
    return (list(SEbeta = SEbeta, SEmu = SEmu, SEeta = SEeta))
  }else{
    return (list(SEbeta = SEbeta, SEmu = SEmu))
  }
}


Main script
library(survival)
library(geepack)
set.seed(327299)
L = 5

source("computeWY.R")
source("compute_se.R")
load("dat.rda")
# reduce to only the first 50 centers
```

```
dat = dat[dat[,"center"]%in% c(1:50),]

N = nrow(dat)
J = length(unique(dat$center))
njs = aggregate(dat$center, list(dat$center), length)$x
ws = rep(1/J, J)
time_start = Sys.time()

### Define parameters
ZD = c("female","prd_diab","prd_htn")
ZC = ZD
n_ZD = length(ZD)

### estimate W(Y)_hat
f = Surv(fol_time_yr, 1-dead) ~
  female+prd_diab+prd_htn+strata(center)
dat = computeWY(f, dat, ZC)
time_censoring = Sys.time()

### point estimation
dat = data.frame(dat, newX = 1, newDelta = 1)
# time our estimation procedure
ModD = coxph(Surv(newX, newDelta) ~
               female+prd_diab+prd_htn+
               strata(center) + offset(-log(Y)) + cluster(USRDS_ID),
             weights = newWt,
             ties = "breslow",
             data = dat[which(dat[, "newWt"] > 0),])
betaHat = matrix(ModD$coef, ncol = 1)
dat = data.frame(dat, HRHat = exp(as.matrix(dat[, ZD]) %*% betaHat))
mujHat = (aggregate(dat[, "WYhat"] * dat[, "Y"], by = list(dat[, "center"]),
                    sum))[, 2] /
  (aggregate(dat[, "WYhat"] * dat[, "HRHat"], by = list(dat[, "center"]),
             sum))[, 2]
etaHat = mujHat / sum(ws * mujHat)
time_death = Sys.time()

### compute SE
betaHat_summary = summary(ModD)$coef
rm(list = c("ModD"))
#center_ind = paste("center", 1:J, sep = "")
r = compute_se(dat, ZD)
SEbeta = r$SEbeta
SEmu = r$SEmu
SEeta = r$SEeta
time_se = Sys.time()
```