Online Appendix A. Proof of Theorem 1

To prove Theorem 1, we assume that $w_i(t;\alpha)$ is a function of $\left(\Delta_i, T_i, \bar{A}_i, V_i, \bar{L}_i\right)$ for each $i = 1, \ldots, n$ and $\left(\hat{\beta}', \hat{\alpha}^{a'}, \hat{\alpha}^{c'}\right)'$ as well as $S_w^{(0)}(t;\beta,\alpha)$, $S_w^{(1)}(t;\beta,\alpha)$,$S_w^{(0)}(t;\hat{\beta},\hat{\alpha})$, $S_w^{(1)}(t;\hat{\beta},\hat{\alpha})$ and $w_i(t;\hat{\alpha})$ are functions of the random variables $\left(\Delta_i, T_i, \bar{A}_i, V_i, \bar{L}_i\right)_{i=1}^n$. We will denote with $Y_i(t) = 1 - C_i(t)$, if an individual is still under observation. We further assume the following regularity conditions:

(A) $\int_0^\infty \tilde{\lambda}_0(t)dt < \infty$

(B) There exists a neighborhood U of $\left(\beta', \alpha'\right)'$ such that for each $j = 0, 1, 2$:

$$\sup_{t\in(0,\infty),(\beta^*,\alpha^*)\in U} \left\| S_w^{(j)}(t;\beta^*,\alpha^*) - s_w^{(j)}(t;\beta^*,\alpha^*)\right\| \xrightarrow{P} 0$$

(C) There exists a neighborhood U of $\left(\beta', \alpha'\right)'$ such that for $\left(\beta^{*'}, \alpha^{*'}\right)' \in U$ and $t \in (0,\infty)$:

$$\nabla_\beta s_w^{(0)}(t;\beta^*,\alpha^*) = s_w^{(1)}(t;\beta^*,\alpha^*),$$

$$\nabla_\beta s_w^{(1)}(t;\beta^*,\alpha^*)' = s_w^{(2)}(t;\beta^*,\alpha^*),$$

$s_w^{(0)}(t;\beta^*,\alpha^*)$, $s_w^{(1)}(t;\beta^*,\alpha^*)$, $s_w^{(2)}(t;\beta^*,\alpha^*)$ are continuous functions of $(\beta^{*'},\alpha^{*'})' \in U$, uniformly in $t \in (0,\infty)$, $s_w^{(0)}$, $s_w^{(1)}$ and $s_w^{(2)}$ are bounded on $U \times (0,\infty)$ and $s_w^{(0)}$ is bounded away from zero on $U \times (0,\infty)$,

(D) There exists a neighborhood U of $\left(\beta', \alpha'\right)'$ such that for all $\left(\beta^{*'}, \alpha^{*'}\right)' \in U$:

- $\mathbb{E}\left(\nabla_{\alpha^{(a)}} u_{w,1}(\beta^*,\alpha^*)'\right)$, $\mathbb{E}\left(u_{\alpha^{(a)},1}(\alpha^{a*}) u_{\alpha^{(a)},1}(\alpha^{a*})'\right)$, $\mathbb{E}\left(\nabla_{\alpha^{(c)}} u_{w,i}(\beta^*,\alpha^*)'\right)$ and $\mathbb{E}\left(u_{\alpha^{(c)},1}(\alpha^{c*}) u_{\alpha^{(c)},1}(\alpha^{c*})'\right)$ exist and $\nabla_{\alpha^{(a)}}\mathbb{E}\left(\nabla_{\alpha^{(a)}} u_{w,1}(\beta^*,\alpha^*)'\right)'$, $\nabla_{\alpha^{(a)}}\mathbb{E}\left(u_{\alpha^{(a)},1}(\alpha^{a*}) u_{\alpha^{(a)},1}(\alpha^{a*})'\right)$, $\nabla_{\alpha^{(c)}}\mathbb{E}\left(\nabla_{\alpha^{(c)}} u_{w,i}(\beta^*,\alpha^*)'\right)'$ and $\nabla_{\alpha^{(c)}}\mathbb{E}\left(u_{\alpha^{(c)},1}(\alpha^{c*}) u_{\alpha^{(c)},1}(\alpha^{c*})'\right)$ are bounded.

- $\frac{1}{n}\nabla_{\alpha^{(a)}} U_{\alpha^{(a)}}(\alpha^{a*})'$, $\frac{1}{n}\nabla_{\alpha^{(c)}} U_{\alpha^{(c)}}(\alpha^{c*})'$, $\frac{1}{n}\nabla_{\alpha^{(a)}} U_w(\beta^*,\alpha^{a*},\alpha^{c*})'$ and $\frac{1}{n}\nabla_{\alpha^{(c)}} U_w(\beta^*,\alpha^{a*},\alpha^{c*})'$ exist and $\frac{1}{n}\nabla_{\alpha^{(a)}}\left(\nabla_{\alpha^{(a)}} U_{\alpha^{(a)}}(\alpha^{a*})'\right)'$, $\frac{1}{n}\nabla_{\alpha^{(c)}}\left(\nabla_{\alpha^{(c)}} U_{\alpha^{(c)}}(\alpha^{c*})'\right)'$, $\frac{1}{n}\nabla_{\alpha^{(a)}}\left(\nabla_{\alpha^{(a)}} U_w(\beta^*,\alpha^{a*},\alpha^{c*})'\right)'$ and $\frac{1}{n}\nabla_{\alpha^{(c)}}\left(\nabla_{\alpha^{(c)}} U_w(\beta^*,\alpha^{a*},\alpha^{c*})'\right)'$ are bounded.

(E) Differentiation with respect to $\alpha^{(a)}$ and $\alpha^{(c)}$ is exchangeable with integration as follows:

$$\nabla_{\alpha^{(a)}}\mathbb{E}\left(u_{w,i}(\beta,\alpha^{(a)},\alpha^{(c)})'\right) = \int \nabla_{\alpha^{(a)}}\left(u_{w,i}(\beta,\alpha^{(a)},\alpha^{(c)})' f(\delta_i,t_i,\bar{a}_i,v_i,\bar{l}_i)\right) d(\delta_i,t_i,\bar{a}_i,v_i,\bar{l}_i)$$

$$\nabla_{\alpha^{(c)}}\mathbb{E}\left(u_{w,i}(\beta,\alpha^{(a)},\alpha^{(c)})'\right) = \int \nabla_{\alpha^{(c)}}\left(u_{w,i}(\beta,\alpha^{(a)},\alpha^{(c)})' f(\delta_i,t_i,\bar{a}_i,v_i,\bar{l}_i)\right) d(\delta_i,t_i,\bar{a}_i,v_i,\bar{l}_i)$$

$$\nabla_{\alpha^{(a)}}\mathbb{E}\left(u_{\alpha^{(a)},i}(\alpha^{(a)})'\right) = \int \nabla_{\alpha^{(a)}}\left(u_{\alpha^{(a)},i}(\alpha^{(a)})' f(\delta_i,t_i,\bar{a}_i,v_i,\bar{l}_i)\right) d(\delta_i,t_i,\bar{a}_i,v_i,\bar{l}_i)$$

$$\nabla_{\alpha^{(c)}}\mathbb{E}\left(u_{\alpha^{(c)},i}(\alpha^{(c)})'\right) = \int \nabla_{\alpha^{(c)}}\left(u_{\alpha^{(c)},i}(\alpha^{(c)})' f(\delta_i,t_i,\bar{a}_i,v_i,\bar{l}_i)\right) d(\delta_i,t_i,\bar{a}_i,v_i,\bar{l}_i)$$

Proof.

The proof is based on the proof of Ali et al. [1] of the variance formula for known IPT weights and the proof of Robins et al. [2] of the variance of estimators resulting from IPC weighted generalized estimating equations.

We set $\eta = (\beta, \alpha^{(a)}, \alpha^{(c)})$ and rewrite $U(\eta) = \left(U_w(\beta,\alpha^{(a)},\alpha^{(c)})', U_{\alpha^{(a)}}(\alpha^{(a)})', U_{\alpha^{(c)}}(\alpha^{(c)})'\right)'$. Applying the mean value theorem yields

$$0 = U(\hat{\eta}) = U(\eta) + \left(\nabla_\eta U(\eta^*)'\right)' (\hat{\eta} - \eta) \tag{A1}$$

with $\eta^* = (\beta^*, \alpha^{a*}, \alpha^{c*})$ on the line segment between $\hat{\eta}$ and η. Since $\nabla_\beta U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'$, $\nabla_{\alpha^{(a)}} U_{\alpha^{(a)}}(\alpha^{a*})'$ and $\nabla_{\alpha^{(c)}} U_{\alpha^{(c)}}(\alpha^{c*})'$ are symmetric matrices, it follows that

$$\left(\nabla_\eta U(\eta^*)'\right)' = \begin{pmatrix} \nabla_\beta U_w(\beta^*, \alpha^{a*}, \alpha^{c*})' & (\nabla_{\alpha^{(a)}} U_w(\beta^*, \alpha^{a*}, \alpha^{c*})')' & (\nabla_{\alpha^{(c)}} U_w(\beta^*, \alpha^{a*}, \alpha^{c*})')' \\ 0 & \nabla_{\alpha^{(a)}} U_{\alpha^{(a)}}(\alpha^{a*})' & 0 \\ 0 & 0 & \nabla_{\alpha^{(c)}} U_{\alpha^{(c)}}(\alpha^{c*})' \end{pmatrix}.$$

Equation (A1) is then equivalent to

$$\begin{aligned} 0 &= U_w(\beta, \alpha^{(a)}, \alpha^{(c)}) + \nabla_\beta U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'(\hat{\beta} - \beta) + \left(\nabla_{\alpha^{(a)}} U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)' (\hat{\alpha}^{(a)} - \alpha^{(a)}) \\ &\quad + \left(\nabla_{\alpha^{(c)}} U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)' (\hat{\alpha}^{(c)} - \alpha^{(c)}), \\ 0 &= U_{\alpha^{(a)}}(\alpha^{(a)}) + \nabla_{\alpha^{(a)}} U_{\alpha^{(a)}}(\alpha^{a*})'(\hat{\alpha}^{(a)} - \alpha^{(a)}) \text{ and} \\ 0 &= U_{\alpha^{(c)}}(\alpha^{(c)}) + \nabla_{\alpha^{(c)}} U_{\alpha^{(c)}}(\alpha^{c*})'(\hat{\alpha}^{(c)} - \alpha^{(c)}) \end{aligned}$$

which can be rewritten as

$$\begin{aligned} \hat{\beta} - \beta &= -\left(\nabla_\beta U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)^{-1} U_w(\beta, \alpha^{(a)}, \alpha^{(c)}) \\ &\quad - \left(\nabla_\beta U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)^{-1} \left(\nabla_{\alpha^{(a)}} U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)' (\hat{\alpha}^{(a)} - \alpha^{(a)}) \\ &\quad - \left(\nabla_\beta U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)^{-1} \left(\nabla_{\alpha^{(c)}} U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)' (\hat{\alpha}^{(c)} - \alpha^{(c)}), \\ \hat{\alpha}^{(a)} - \alpha^{(a)} &= -\left(\nabla_{\alpha^{(a)}} U_{\alpha^{(a)}}(\alpha^{a*})'\right)^{-1} U_{\alpha^{(a)}}(\alpha^{(a)}) \text{ and} \\ \hat{\alpha}^{(c)} - \alpha^{(c)} &= -\left(\nabla_{\alpha^{(c)}} U_{\alpha^{(c)}}(\alpha^{c*})'\right)^{-1} U_{\alpha^{(c)}}(\alpha^{(c)}), \end{aligned}$$

respectively. Inserting the last two equations into the first one and multiplying by $\sqrt{n}$ leads to

$$\begin{aligned} \sqrt{n}(\hat{\beta} - \beta) = \left(\frac{1}{n}\nabla_\beta U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)^{-1} \Bigg(&-\frac{1}{\sqrt{n}} U_w(\beta, \alpha^{(a)}, \alpha^{(c)}) + \\ &\frac{1}{n}\left(\nabla_{\alpha^{(a)}} U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)' \left(\frac{1}{n}\nabla_{\alpha^{(a)}} U_{\alpha^{(a)}}(\alpha^{a*})'\right)^{-1} \frac{1}{\sqrt{n}} U_{\alpha^{(a)}}(\alpha^{(a)}) + \\ &\frac{1}{n}\left(\nabla_{\alpha^{(c)}} U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)' \left(\frac{1}{n}\nabla_{\alpha^{c*}} U_{\alpha^{(c)}}(\alpha^{c*})'\right)^{-1} \frac{1}{\sqrt{n}} U_{\alpha^{(c)}}(\alpha^{(c)}) \Bigg). \end{aligned} \tag{A2}$$

Using the same arguments as in Lin et al. [3] and the definition of $\tilde{U}_w(\beta, \alpha^{(a)}, \alpha^{(c)})$ in (13), we may replace $\frac{1}{\sqrt{n}} U_w(\beta, \alpha^{(a)}, \alpha^{(c)})$ by $\frac{1}{\sqrt{n}} \tilde{U}_w(\beta, \alpha^{(a)}, \alpha^{(c)}) + o_p(1)$, with $o_p(1)$ denoting a random variable converging to zero in probability. Analogously to Andersen and Gill [4], using regularity assumptions (A)-(C), $\left(\frac{1}{n}\nabla_\beta U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)$ converges in probability to $H_{\beta\beta}$. By the law of large numbers and regularity assumption (D), $\left(\frac{1}{n}\nabla_{\alpha^{(a)}} U_{\alpha^{(a)}}(\alpha^{a*})\right)'$ converges in probability to $H_{\alpha^{(a)}\alpha^{(a)}} = \mathbb{E}\left(\nabla_{\alpha^{(a)}} u_{1,\alpha^{(a)}}(\alpha^{(a)})'\right)$, $\left(\frac{1}{n}\nabla_{\alpha^{(c)}} U_{\alpha^{(c)}}(\alpha^{c*})\right)'$ converges in probability to $H_{\alpha^{(c)}\alpha^{(c)}} = \mathbb{E}(\nabla_{\alpha^{(c)}} u_{1,\alpha^{(c)}}(\alpha^{(c)})')$, $\frac{1}{n}\left(\nabla_{\alpha^{(a)}} U_w(\beta^*, \alpha^{a*}, \alpha^{c*})'\right)'$ converges in probability to $H_{\beta\alpha^a}$ and $\frac{1}{n}\left(\nabla_{\alpha^{(c)}} U_w(\beta^*, \alpha^{a*}, \alpha^{c*})\right)'$

converges in probability to $H_{\beta\alpha^{(c)}}$. Then, (A2) can be rewritten as

$$\sqrt{n}(\hat{\beta}-\beta) = H_{\beta\beta}^{-1}\left(-\frac{1}{\sqrt{n}}\tilde{U}_w(\beta,\alpha^{(a)},\alpha^{(c)}) + H_{\beta\alpha^{(a)}}H^{-1}_{\alpha^{(a)}\alpha^{(a)}}\frac{1}{\sqrt{n}}U_{\alpha^{(a)}}(\alpha^{(a)}) + H_{\beta\alpha^{(c)}}H^{-1}_{\alpha^{(c)}\alpha^{(c)}}\frac{1}{\sqrt{n}}U_{\alpha^{(c)}}(\alpha^{(c)})\right) + o_p(1)$$
$$= H_{\beta\beta}^{-1}\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} -\tilde{u}_{w,i}(\beta,\alpha^{(a)},\alpha^{(c)}) + H_{\beta\alpha^{(a)}}H^{-1}_{\alpha^{(a)}\alpha^{(a)}}u_{\alpha^{(a)},i}(\alpha^{(a)}) + H_{\beta\alpha^{(a)}}H^{-1}_{\alpha^{(c)}\alpha^{(c)}}u_{\alpha^{(c)},i}(\alpha^{(c)})\right) + o_p(1).$$

Due to the central limit theorem, it holds that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} -\tilde{u}_{w,i}(\beta,\alpha^{(a)},\alpha^{(c)}) + H_{\beta\alpha^{(a)}}H^{-1}_{\alpha^{(a)}\alpha^{(a)}}u_{\alpha^{(a)},i}(\alpha^{(a)}) + H_{\beta\alpha^{(a)}}H^{-1}_{\alpha^{(c)}\alpha^{(c)}}u_{\alpha^{(c)},i}(\alpha^{(c)})$$

converges in distribution to a mean zero Gaussian variable B with variance

$$\begin{aligned} Var(B) &= \mathbb{E}\left(\left(-\tilde{u}_{w,1}(\beta,\alpha^{(a)},\alpha^{(c)}) + H_{\beta\alpha^{(a)}}H^{-1}_{\alpha^{(a)}\alpha^{(a)}}u_{\alpha^{(a)},1}(\alpha^{(a)}) + H_{\beta\alpha^{(a)}}H^{-1}_{\alpha^{(c)}\alpha^{(c)}}u_{\alpha^{(c)},1}(\alpha^{(c)})\right)^{\otimes 2}\right) \\ &= \mathbb{E}\left(\tilde{u}_{w,1}(\beta,\alpha^{(a)},\alpha^{(c)})\tilde{u}_{w,1}(\beta,\alpha^{(a)},\alpha^{(c)})'\right) \\ &+ H_{\beta\alpha^{(a)}}H^{-1}_{\alpha^{(a)}\alpha^{(a)}}\mathbb{E}\left(u_{\alpha^{(a)},1}(\alpha^{(a)})u_{\alpha^{(a)},1}(\alpha^{(a)})'\right)H^{-1}_{\alpha^{(a)}\alpha^{(a)}}H'_{\beta\alpha^{(a)}} \\ &+ H_{\beta\alpha^{(c)}}H^{-1}_{\alpha^{(c)}\alpha^{(c)}}\mathbb{E}\left(u_{\alpha^{(c)},1}(\alpha^{(c)})u_{\alpha^{(c)},1}(\alpha^{(c)})'\right)H^{-1}_{\alpha^{(c)}\alpha^{(c)}}H'_{\beta\alpha^{(c)}} \\ &- 2\mathbb{E}\left(\tilde{u}_{w,1}(\beta,\alpha^{(a)},\alpha^{(c)})u_{\alpha^{(a)},1}(\alpha^{(a)})'\right)H^{-1}_{\alpha^{(a)}\alpha^{(a)}}H'_{\beta\alpha^{(a)}} \\ &- 2\mathbb{E}\left(\tilde{u}_{w,1}(\beta,\alpha^{(a)},\alpha^{(c)})u_{\alpha^{(c)},1}(\alpha^{(c)})'\right)H^{-1}_{\alpha^{(c)}\alpha^{(c)}}H'_{\beta\alpha^{(c)}} \\ &+ 2H_{\beta\alpha^{(a)}}H^{-1}_{\alpha^{(a)}\alpha^{(a)}}\mathbb{E}\left(u_{\alpha^{(a)},1}(\alpha^{(a)})u_{\alpha^{(c)},1}(\alpha^{(c)})'\right)H^{-1}_{\alpha^{(c)}\alpha^{(c)}}H'_{\beta\alpha^{(c)}}, \end{aligned} \tag{A3}$$

whereby we use the standard notation $v^{\otimes 2} = vv'$. Thus, to prove the theorem, it remains to show that

$$Var(B) = F_\beta - H_{\beta\alpha^{(a)}}F^{-1}_{\alpha^{(a)}}H'_{\beta\alpha^{(a)}} - H_{\beta\alpha^{(c)}}F^{-1}_{\alpha^{(c)}}H'_{\beta\alpha^{(c)}}.$$

We will first show that $H_{\beta\alpha^{(a)}} = \mathbb{E}\left(u_{w,1}(\beta,\alpha^{(a)},\alpha^{(c)})u_{\alpha^{(a)},1}(\alpha^{(a)})'\right)$.
Following the argumentation outlined in Lawless et al. [5, chapter 7.1.3], it can be shown that $\mathbb{E}(U_w(\beta,\alpha))$ and thus $\mathbb{E}\left(u_{w,1}(\beta,\alpha^{(a)},\alpha^{(c)})\right)$ as a function of the true parameter value is always zero. Therefore

$$\nabla_{\alpha^{(a)}}\mathbb{E}\left(u_{w,1}(\beta,\alpha^{(a)},\alpha^{(c)})'\right) = 0. \tag{A4}$$

Furthermore, the joint density of $(\Delta_1, T_1, \bar{A}_1, V_1, \bar{L}_1)$ depends on $\alpha^{(a)}$ only through the conditional densities $Pr\left(A_1(k) = a_1(k)|Y_1(k) = 1, \bar{A}_1(k-1) = \bar{a}_1(k-1), V_1 = v_1, \bar{L}_1(k) = \bar{l}_1(k); \alpha^{(a)}\right)$, $k = 1,\dots,K$. Using (1) of the main document, we may write the joint density of $(\Delta_1, Y_1, \bar{A}_1, \bar{L}_1)$ as

$$\begin{aligned} & f\left(\delta_1, t_1, \bar{a}_1, v_1\bar{l}_1; \beta, \alpha^{(a)}, \alpha^{(c)}\right) \\ &= Pr\left(\bar{A}_1 = \bar{a}_1|\Delta_1 = \delta_1, T_1 = t_1, V_1 = v_1, \bar{L}_1 = \bar{l}_1; \beta, \alpha^{(a)}, \alpha^{(c)}\right) \times f\left(\delta_1, t_1, v_1, \bar{l}_1; \beta, \alpha^{(c)}\right) \\ &= \prod_{k=1}^{K} Pr\left(A_1(k) = a_1(k)|Y_1(k) = 1, \bar{A}_1(k-1) = \bar{a}_1(k-1), V_1 = v_1, \bar{L}_1(k) = \bar{l}_1(k); \alpha^{(a)}\right) \\ & \times f\left(\delta_1, t_1, v_1, \bar{l}_1; \beta, \alpha^{(c)}\right). \end{aligned}$$

The derivative of $f\left(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1; \beta, \alpha^{(a)}, \alpha^{(c)}\right)$ is then calculated as

$$\begin{aligned}
&\nabla_{\alpha^{(a)}} f\left(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1; \beta, \alpha^{(a)}, \alpha^{(c)}\right) \\
&= \left(\nabla_{\alpha^{(a)}} \prod_{k=1}^{K} Pr\left(A_1(k) = a_1(k) | Y_1(k) = 1, \bar{A}_1(k-1) = \bar{a}_1(k-1), V_1 = v_1, \bar{L}_1(k) = \bar{l}_1(k); \alpha^{(a)}\right)\right) \\
&\times f\left(\delta_1, t_1, \bar{l}_1; \beta, \alpha^{(c)}\right).
\end{aligned}$$

Applying the product rule for derivatives, we obtain

$$\begin{aligned}
&\nabla_{\alpha^{(a)}} f\left(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1; \beta, \alpha^{(a)}, \alpha^{(c)}\right) \\
&= \left(\nabla_{\alpha^{(a)}} \log \prod_{k=1}^{K} Pr\left(A_1(k) = a_1(k) | Y_1(k) = 1, \bar{A}_1(k-1) = \bar{a}_1(k-1), V_1 = v_1, \bar{L}_1(k) = \bar{l}_1(k); \alpha^{(a)}\right)\right) \\
&\times \prod_{k=1}^{K} Pr\left(A_1(k) = a_1(k) | Y_i(k) = 1, \bar{A}_1(k-1) = \bar{a}_1(k-1), V_1 = v_1, \bar{L}_1(k) = \bar{l}_1(k); \alpha^{(a)}\right) \\
&\times f\left(\delta_1, t_1, v_1, \bar{l}_1; \beta, \alpha^{(c)}\right).
\end{aligned}$$

Since the first factor is by definition equal to of $u_{\alpha^{(a)},1}(\alpha^{(a)})$ in (4) of the main document and the two other factors simplify to $f\left(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1; \beta, \alpha^{(a)}, \alpha^{(c)}\right)$, we get

$$\nabla_{\alpha^{(a)}} f\left(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1; \beta, \alpha^{(a)}, \alpha^{(c)}\right) = u_{\alpha^{(a)},1}(\alpha^{(a)}) f\left(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1; \beta, \alpha^{(a)}, \alpha^{(c)}\right) \tag{A5}$$

Finally, under regularity assumption (E), (A4) can be expressed using (A5) as

$$\begin{aligned}
0 &= \nabla_{\alpha^{(a)}} \mathbb{E}\left(u_{w,1}(\beta, \alpha^{(a)}, \alpha^{(c)})'\right) \\
&= \int \nabla_{\alpha^{(a)}} \left(u_{w,1}(\beta, \alpha^{(a)}, \alpha^{(c)})' f\left(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1\right)\right) d(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1) \\
&= \int \nabla_{\alpha^{(a)}} u_{w,1}(\beta, \alpha^{(a)}, \alpha^{(c)})' f\left(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1\right) d(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1) \\
&+ \int u_{\alpha^{(a)},1}(\alpha^{(a)}) u_{w,1}(\beta, \alpha^{(a)}, \alpha^{(c)})' f\left(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1\right) d(\delta_1, t_1, \bar{a}_1, v_1, \bar{l}_1)
\end{aligned}$$

and thus

$$H_{\beta\alpha^{(a)}} = \mathbb{E}\left(\nabla_{\alpha^{(a)}} u_{w,1}(\beta, \alpha^{(a)}, \alpha^{(c)})'\right)' = -\mathbb{E}\left(u_{w,1}(\beta, \alpha^{(a)}, \alpha^{(c)}) u_{\alpha^{(a)},1}(\alpha^{(a)})'\right). \tag{A6}$$

Analogously,

$$H_{\beta\alpha^{(c)}} = -\mathbb{E}\left(u_{w,1}(\beta, \alpha^{(a)}, \alpha^{(c)}) u_{\alpha^{(c)},1}(\alpha^{(c)})'\right), \tag{A7}$$

$$H_{\alpha^{(a)}\alpha^{(a)}} = \mathbb{E}\left(\nabla_{\alpha^{(a)}} u_{\alpha^{(a)},1}(\alpha^{(a)})'\right) = -\mathbb{E}\left(u_{\alpha^{(a)},1}(\alpha^{(a)}) u_{\alpha^{(a)},1}(\alpha^{(a)})'\right) \tag{A8}$$

and

$$H_{\alpha^{(c)}\alpha^{(c)}} = -\mathbb{E}\left(u_{\alpha^{(c)},1}(\alpha^{(c)}) u_{\alpha^{(c)},1}(\alpha^{(c)})'\right). \tag{A9}$$

Prepared using simauth.cls

Furthermore, since $u_{\alpha^{(a)},1}(\alpha^{(a)})$ does not depend on α_c,

$$\mathbb{E}\left(u_{\alpha^{(a)},1}(\alpha^{(a)})u_{\alpha^{(c)},1}(\alpha^{(c)})'\right) = \mathbb{E}\left(\nabla_{\alpha^{(c)}}u_{\alpha^{(a)},1}(\alpha^{(a)})\right) = \mathbb{E}(0) = 0. \tag{A10}$$

Plugging (A6), (A7), (A8), (A9) and (A10) into (A3) yields

$$\begin{aligned} Var(B) = {} & \mathbb{E}\left(\tilde{u}_{w,1}(\beta, \alpha^{(a)}, \alpha^{(c)})\tilde{u}_{w,1}(\beta, \alpha^{(a)}, \alpha^{(c)})'\right) \\ & + H_{\beta\alpha^{(a)}}\mathbb{E}\left(u_{\alpha^{(a)},1}(\alpha^{(a)})u_{\alpha^{(a)},1}(\alpha^{(a)})'\right)^{-1} H'_{\beta\alpha^{(a)}} \\ & + H_{\beta\alpha^{(c)}}\mathbb{E}\left(u_{\alpha^{(c)},1}(\alpha^{(c)})u_{\alpha^{(c)},1}(\alpha^{(c)})'\right)^{-1} H'_{\beta\alpha^{(c)}} \\ & - 2H_{\beta\alpha^{(a)}}\mathbb{E}\left(u_{\alpha^{(a)},1}(\alpha^{(a)})u_{\alpha^{(a)},1}(\alpha^{(a)})'\right)^{-1} H'_{\beta\alpha^{(a)}} \\ & - 2H_{\beta\alpha^{(c)}}\mathbb{E}\left(u_{\alpha^{(c)},1}(\alpha^{(c)})u_{\alpha^{(c)},1}(\alpha^{(c)})'\right)^{-1} H'_{\beta\alpha^{(c)}}, \end{aligned}$$

This simplifies to

$$Var(B) = F_\beta - H_{\beta\alpha^{(a)}}F^{-1}_{\alpha^{(a)}}H'_{\beta\alpha^{(a)}} - H_{\beta\alpha^{(c)}}F^{-1}_{\alpha^{(c)}}H'_{\beta\alpha^{(c)}},$$

which completes the proof.

References

1. Ali RA, Ali MA, Wei Z. On computing standard errors for marginal structural Cox models. *Lifetime Data Analysis* 2014; **20**(1):106–131.
2. Robins JM, Rotnitzky A, Zhao LP. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association* 1995; **90**(429):106–121.
3. Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association* 1989; **84**(408):1074–1078.
4. Anderson PK, Gill RD. Cox's regression model for counting processes: a large sample study. *The Annals of Statistics* 1982; **10**(4):1100–1120.
5. Lawless JF. *Statistical Models and Methods for Lifetime Data* (2nd edn). Wiley: Hoboken, New Jersey, 2011.