Testing Hypotheses of Covariate-Adaptive Randomized Clinical Trials with Time-to-event Outcomes under the AFT model

Hongjian Zhu[1], Li-Xin Zhang[2*], Jing Ning[3] and Lu Wang[4]

[1] *AbbVie Inc.*

[2*] *Zhejiang Gongshang University and Zhejiang University*

[3] *University of Texas MD Anderson Cancer Center*

[4] *University of Texas Health Science Center at Houston*

Supplementary Material

S1 Proof

For a Borel function $h_i = h(T_i, I_i, C_i, \boldsymbol{X}_i, \boldsymbol{Z}_i)$, we let $h_i^{(1)} = h(T_i^{(1)}, 1, C_i, \boldsymbol{X}_i, \boldsymbol{Z}_i)$ and $h_i^{(0)} = h(T_i^{(0)}, 0, C_i, \boldsymbol{X}_i, \boldsymbol{Z}_i)$ be the values of h_i when $I_i = 1, 0$, where

$$\begin{aligned}\log T_i^{(1)} =& \mu_1 + \beta_1 X_{i,1} + \cdots + \beta_{p_1} X_{i,p_1} + \gamma_1 Z_{i,1} + \cdots + \gamma_{p_2} Z_{i,p_2} + \varepsilon_i,\\ \log T_i^{(0)} =& \mu_2 + \beta_1 X_{i,1} + \cdots + \beta_{p_1} X_{i,p_1} + \gamma_1 Z_{i,1} + \cdots + \gamma_{p_2} Z_{i,p_2} + \varepsilon_i.\end{aligned}$$

Let $\delta_i^{(j)}$, $Y_i^{(j)}$ be the values of δ_i, Y_i when $I_i = j$, $j = 1, 0$. Define $\widetilde{E} h_i = qEh_i^{(1)} + (1-q)Eh_i^{(0)}$ to be the expectation of h_i when I_i is com-

*Email:statzlx@zju.edu.cn

pletely random with probability q, i.e., I_i is independent of the other random variables with $P(I_i = 1) = q$, $P(I_i = 0) = 1 - q$. It is clear that $E|h_i| \leq \widetilde{E}|h_i|/(q(1-q))$. For a vector $\boldsymbol{v}$, we denote $\boldsymbol{v}^{\otimes 2} = \boldsymbol{v}\boldsymbol{v}^T$.

To obtain the asymptotic properties, we let $\pi(s) = G(s)\widetilde{P}\left(T_1 \geq s\right)$, $\boldsymbol{B}_1(s) = \widetilde{E}\left[I\{T_1 \geq s\}\underline{\boldsymbol{X}}_1\left(\boldsymbol{Z}_1^T\gamma + \varepsilon_1\right)\right]$, and $\boldsymbol{B}_2(s) = \widetilde{E}\left[I\{T_1 \geq s\}\underline{\boldsymbol{X}}_1^{\otimes 2}\right]$. Moreover, let $\Lambda_c\left(u\right)$ be the cumulative hazard function of the censoring times C. We need the following regularity conditions:

(Ra) ϵ_i and C_i have continuous distributions with $\sup\{t : P(T_1^{(j)} > t) > 0\} \geq \sup\{t : P(C > t) > 0\} \doteq \tau_G$, $j = 1, 0$;

(Rb) $E[f_i^4/G(T_i \wedge \tau_G)|I_i] < \infty$ for $f_i = 1, X_{i,t}, Z_{i,s}$ or ϵ_i, $t = 1, \ldots, p_1$, $s = 1, \ldots, p_2$;

(Rc) $det\big(\int_0^\infty \boldsymbol{B}_2(u)d\Lambda_G(u)\big) < \infty$, $det\big(\int_0^\infty \boldsymbol{B}_1^{\otimes 2}(u)/\pi(u)d\Lambda_G(u)\big) < \infty$.

Suppose that the covariance matrix $Var(\boldsymbol{X}_1)$ is nonsingular. Write $u_i = \gamma_1 Z_{i,1} + \ldots + \gamma_{p_2} Z_{i,p_2} + \epsilon_i$, $\breve{u}_i = E[u_i|W_i] - E[u_i]$. Then we have Theorem 1 in the main paper as follows.

Theorem 1. *Suppose that a covariate-adaptive design satisfies the following three conditions:*

(A) $Cov(X_{i,k}, u_i) =$, $k = 1, \ldots, p_1$;

(B) $\sum_{i=1}^{n}(I_i - q)\breve{u}_i = o_P(\sqrt{n})$;

(C) the within-stratum q-imbalances for all covariates are of order $o(n)$ *in probability, i.e.,* $D_n^{(q)}(t_1, \ldots, t_{p_1}, r_1, r_2, \ldots, r_{p_2}) = o_P(n)$ *for all* $t_k s$ *and* $r_j s$.

Further, suppose that the regularity conditions (Ra)–(Rc) are satisfied.

Then we have the following results:

(i) Under $H_0 : \mu_1 - \mu_2 = 0$,

$$\mathcal{T}(n) \xrightarrow{D} N(0, \tau^2), \text{ with } \tau^2 = \frac{\sigma_{\delta,G}^2}{\sigma_{z,G}^2} \tag{S1.1}$$

where $\sigma_{z,G}^2 = E\big[(u_i - Eu_i)^2 / G(T_i \wedge \tau_G) | H_0\big]$, $\sigma_{\delta,G}^2 = \sigma_{z,G}^2 - E\breve{u}_i^2$.

(ii) Under $H_A : \mu_1 - \mu_2 \neq 0$, *consider a sequence of local alternatives, i.e.,* $\mu_2 = \mu_1 - \delta/\sqrt{n}$ *for a fixed* $\delta \neq 0$. *Then*

$$\mathcal{T}(n) \xrightarrow{D} N(\Delta, \tau^2), \text{ with } \Delta = \frac{\delta\sqrt{q(1-q)}}{\sigma_{z,G}}. \tag{S1.2}$$

The following remark is Remark 1 in main paper.

Remark 1. (i) Suppose that the marginal q-imbalances for covariates $Z_1, \ldots, Z_{p_2}$ are of order $o(\sqrt{n})$ in probability, i.e., $D_n^{(q)}(Z, j; r_j) = o_P(\sqrt{n})$, $j = 1, \ldots, p_2$, and that $Z_1, \ldots, Z_{p_2}$ are independent and independent of $\boldsymbol{X}$. Then Assumptions (A) and (B) are satisfied. In this case, $E[\breve{u}_i^2] = \sum_{j=1}^{p_2} \gamma_j^2 Var(E[Z_{i,j} | \tilde{Z}_{i,j}])$.

(ii) Suppose that the within-stratum q-imbalances for $Z_1, \ldots, Z_{p_2}$ are of order $o(\sqrt{n})$ in probability, i.e., $D_n^{(q)}(Z; r_1, \ldots, r_{p_2}) = o_P(\sqrt{n})$ for all r_js, and that $\boldsymbol{Z}$ is independent of $\boldsymbol{X}$. Then Assumptions (A) and (B) are satisfied.

(iii) Suppose the order $o(n)$ in Assumption (C) is strengthened to $o(\sqrt{n})$. Then Assumptions (B) and (C) are satisfied.

Proof. Before proving Theorem 1, we first show this remark. Note that $\breve{u}_i = E[u_i - E[u_i]|W_i] = g(W_i)$ is a function of W_i. If the order $o(n)$ in Assumption (C) is strengthened to $o(\sqrt{n})$, then

$$
\begin{aligned}
&\sum_{i=1}^{n}(I_i - q)\breve{u}_i = \sum_{i=1}^{n}(I_i - q)g(W_i) \\
= &\sum_{t_1,\ldots,t_{p_1},r_1,\ldots,r_{p_2}} (I_i - q)I\{W_i = (x_1^{t_1}, \ldots, x_{p_1}^{t_{p_1}}, z_1^{r_1}, \ldots, z_{p_2}^{r_{p_2}})\} \\
&\qquad \cdot g(x_1^{t_1}, \ldots, x_{p_1}^{t_{p_1}}, z_1^{r_1}, \ldots, z_{p_2}^{r_{p_2}}) \\
= &\sum_{t_1,\ldots,t_{p_1},r_1,\ldots,r_{p_2}} D_n^{(q)}(t_1, \ldots, t_{p_1}, r_1, \ldots, r_{p_2}) g(x_1^{t_1}, \ldots, x_{p_1}^{t_{p_1}}, z_1^{r_1}, \ldots, z_{p_2}^{r_{p_2}}) \\
= &\sum_{t_1,\ldots,t_{p_1},r_1,\ldots,r_{p_2}} o_P(\sqrt{n}) g(x_1^{t_1}, \ldots, x_{p_1}^{t_{p_1}}, z_1^{r_1}, \ldots, z_{p_2}^{r_{p_2}}) = o_P(\sqrt{n}). \qquad \text{(S1.3)}
\end{aligned}
$$

Assumption (B) is satisfied.

When $\boldsymbol{Z}$ is independent of $\boldsymbol{X}$, then $\breve{u}_i = E[u_i - E[u_i]|W_i] = E[u_i - E[u_i]|\tilde{\boldsymbol{Z}}_i] = g(\tilde{\boldsymbol{Z}}_i)$ is a function of $\tilde{\boldsymbol{Z}}_i = (\tilde{Z}_{i,1}, \ldots, \tilde{Z}_{i,p_2})$. Then similar to

(S1.3),

$$\sum_{i=1}^{n}(I_i-q)\breve{u}_i=\sum_{i=1}^{n}(I_i-q)g(\tilde{\boldsymbol{Z}}_i)=o_P(\sqrt{n})$$

if the within-stratum q-imbalances for $Z_1,\ldots,Z_{p_2}$ are of order $o(\sqrt{n})$ in probability.

When $\boldsymbol{X},Z_1,\ldots,Z_{p_2}$ are independent, $\breve{u}_i=E[u_i-E[u_i]|W_i]=E[u_i-E[u_i]|\tilde{\boldsymbol{Z}}_i]=\sum_{i=1}^{p_2}\gamma_j E[Z_{i,j}-E[Z_{i,j}]|\tilde{Z}_{i,j}]$, and $E[Z_{i,j}-E[Z_{i,j}]|\tilde{Z}_{i,j}]=g_j(\tilde{Z}_{i,j})$ is a function of $\tilde{Z}_{i,j}$. Then

$$\begin{aligned}\sum_{i=1}^{n}(I_i-q)\breve{u}_i&=\sum_{j=1}^{q}\gamma_j\sum_{i=1}^{n}(I_i-q)g_j(\tilde{Z}_{i,j})\\ &=\sum_{j=1}^{q}\gamma_j\sum_{r_j=1}^{s_j}D_n^{(q)}(Z,j;r_j)g_j(z_j^{r_j})=o_P(\sqrt{n}).\end{aligned}$$

if the marginal q-imbalances for covariates $Z_1,\ldots,Z_{p_2}$ are of order $o(\sqrt{n})$ in probability. Assumption (B) is also satisfied. Also, $E\breve{u}_i^2=Var(\sum_{i=1}^{p_2}\gamma_j E[Z_{i,j}|\tilde{Z}_{i,j}])=\sum_{j=1}^{p_2}\gamma_j^2 Var(E[Z_{i,j}|\tilde{Z}_{i,j}])$. $\square$.

To prove Theorem 1, we need some lemmas.

Lemma 1. *Let the function $h_i=h(T_i,I_i,C_i,\boldsymbol{X}_i,\boldsymbol{Z}_i)$ be such that $\widetilde{E}|h_i|<\infty$. Under assumption (C) of Theorem 1, $\sum_{i=1}^n I_i h_i/n\xrightarrow{P}qEh_i^{(1)}$, $\sum_{i=1}^n(1-I_i)h_i/n\xrightarrow{P}(1-q)Eh_i^{(0)}$, and $\sum_{i=1}^n h_i/n\xrightarrow{P}\widetilde{E}h_i$.*

In particular, we have the following results:

(1.1) $\frac{1}{n}\sum_{i=1}^n I_i\xrightarrow{P}q$, $\frac{1}{n}\sum_{i=1}^n(1-I_i)\xrightarrow{P}1-q$.

(1.2) $\frac{1}{n}\sum_{i=1}^n I_i X_{i,k} \xrightarrow{P} qEX_k$, $\frac{1}{n}\sum_{i=1}^n (1-I_i) X_{i,k} \xrightarrow{P} (1-q)EX_k$, $\frac{1}{n}\sum_{i=1}^n I_i Z_{i,j} \xrightarrow{P} qEZ_j$, *and* $\frac{1}{n}\sum_{i=1}^n (1-I_i) Z_{i,j} \xrightarrow{P} (1-q)EZ_j$, $k = 1,\dots,p_1$, $j = 1,\dots,p_2$.

(1.3) $\frac{1}{n}\sum_{i=1}^n \frac{I_i\delta_i}{G(Y_i)} f(\boldsymbol{X}_i,\boldsymbol{Z}_i,\epsilon_i) \xrightarrow{P} qE[f(\boldsymbol{X}_1,\boldsymbol{Z}_1,\epsilon_1)]$, $\frac{1}{n}\sum_{i=1}^n \frac{(1-I_i)\delta_i}{G(Y_i)} f(\boldsymbol{X}_i,\boldsymbol{Z}_i,\epsilon_i) \xrightarrow{P} (1-q)E[f(\boldsymbol{X}_1,\boldsymbol{Z}_1,\epsilon_1)]$, *and* $\frac{1}{n}\sum_{i=1}^n \frac{\delta_i}{G(Y_i)} f(\boldsymbol{X}_i,\boldsymbol{Z}_i,\epsilon_i) \xrightarrow{P} E[f(\boldsymbol{X}_1,\boldsymbol{Z}_1,\epsilon_1)]$ *if* $E[|f(\boldsymbol{X}_1,\boldsymbol{Z}_1,\epsilon_1)|] < \infty$.

Proof. Let $\mathscr{F}_{i-1} = \sigma(I_{i-1},\boldsymbol{X}_{i-1},\boldsymbol{Z}_{i-1},C_{i-1},T_{i-1})$ be the history sigma field. We first show that if $g(W_i)$ is a Borel function of W_i with $E|g(W_i)| < \infty$, then

$$\frac{1}{n}\sum_{i=1}^n I_i g(W_i) \xrightarrow{P} qEg(W_i). \tag{S1.4}$$

By assumption (C) of Theorem 1, we have

$$\begin{aligned}
&\sum_{i=1}^n (I_i - q) g(W_i) \\
&= \sum_{t_1,\dots,t_{p_1},r_1,\dots,r_{p_2}} (I_i - q) I\{W_i = (x_1^{t_1},\dots,x_{p_1}^{t_{p_1}},z_1^{r_1},\dots,z_{p_2}^{r_{p_2}})\} \\
&\qquad \cdot g(x_1^{t_1},\dots,x_{p_1}^{t_{p_1}},z_1^{r_1},\dots,z_{p_2}^{r_{p_2}}) \\
&= \sum_{t_1,\dots,t_{p_1},r_1,\dots,r_{p_2}} D_n^{(q)}(t_1,\dots,t_{p_1},r_1,\dots,r_{p_2}) g(x_1^{t_1},\dots,x_{p_1}^{t_{p_1}},z_1^{r_1},\dots,z_{p_2}^{r_{p_2}}) \\
&= \sum_{t_1,\dots,t_{p_1},r_1,\dots,r_{p_2}} o_P(n) g(x_1^{t_1},\dots,x_{p_1}^{t_{p_1}},z_1^{r_1},\dots,z_{p_2}^{r_{p_2}}) = o_P(n).
\end{aligned}$$

On the other hand, by the law of large numbers for i.i.d. random variables, we have $\sum_{i=1}^n g(W_i)/n \xrightarrow{P} Eg(W_i)$. It follows that

$$\frac{1}{n}\sum_{i=1}^n I_i g(W_i) = \frac{1}{2n}\sum_{i=1}^n (I_i - q)g(W_i) + q\frac{1}{n}\sum_{i=1}^n g(W_i)$$
$$\xrightarrow{P} qEg(W_i).$$

The proof of (S1.4) is complete.

Now, note that $I_i h_i = I_i h_i^{(1)}$. Since we have

$$\frac{1}{n}\left|\sum_{i=1}^n I_i h_i^{(1)} I\{|h_i^{(1)}| \geq M\}\right| \leq \frac{1}{n}\sum_{i=1}^n |h_i^{(1)}| I\{|h_i^{(1)}| \geq M\}$$
$$\xrightarrow{P} E\big[|h_1^{(1)}| I\{|h_1^{(1)}| \geq M\}\big] \to 0 \quad \text{as } M \to \infty$$

by the law of large numbers for i.i.d. random variables, we can assume without loss of generality that the h_is are bounded. Let $g_i = E[h_i^{(1)}|W_i] = g(W_i)$. Then

$$E\Big[I_i\big(h_i^{(1)} - g_i\big)\big|\mathscr{F}_{i-1}, W_i\Big] = E\Big[I_i\big|\mathscr{F}_{i-1}, W_i\Big]E\Big[h_i^{(1)} - g_i\big|W_i\Big] = 0.$$

Thus, $E\Big[I_i\big(h_i^{(1)} - g_i\big)\big|\mathscr{F}_{i-1}\Big] = 0$. It follows that $\{I_i(h_i^{(1)} - g_i), i = 1, 2\ldots\}$ is a sequence of bounded martingale differences. By the law of large numbers for martingale differences,

$$\frac{1}{n}\sum_{i=1}^n I_i(h_i^{(1)} - g_i) \xrightarrow{P} 0.$$

On the other hand, by (S1.4),

$$\frac{1}{n}\sum_{i=1}^n I_i g_i \xrightarrow{P} qE[g(W_i)] = qE\big[E[h_i^{(1)}|W_i]\big] = qEh_i^{(1)}.$$

It follows that $\frac{1}{n}\sum_{i=1}^{n} I_i h_i \xrightarrow{P} qEh_i^{(1)}$. The proof of $\sum_{i=1}^{n}(1-I_i)h_i \xrightarrow{P} (1-q)Eh_i^{(0)}$ is similar.

The conclusions of (1.1) and (1.2) follow immediately. For (1.3), note that $\delta_i^{(j)}, \boldsymbol{X}_i, \boldsymbol{Z}_i$ are independent given $T_i^{(j)}$ (Stute (1993)), i.e., $E\big[\delta_i^{(j)}|\mathscr{F}_{i-1}, \boldsymbol{X}_i, \boldsymbol{Z}_i, T_i^{(j)}\big] = E\big[\delta_i^{(j)}|T_i^{(j)}\big] = G\big(T_i^{(j)}\big)$, $j = 1, 0$, and then $E\big[\delta_i|\boldsymbol{X}_i, \boldsymbol{Z}_i, T_i, I_i\big] = E\big[\delta_i|T_i\big] = G\big(T_i\big)$. Thus,

$$E\Big[\delta_i^{(j)}/G\big(Y_i^{(j)}\big)|\boldsymbol{X}_i, \boldsymbol{Z}_i, T_i^{(j)}\Big] = E\Big[\delta_i^{(j)}/G\big(T_i^{(j)}\big)|\boldsymbol{X}_i, \boldsymbol{Z}_i, T_i^{(j)}\Big] = 1,$$

$j = 1, 0$, and

$$E\Big[\frac{\delta_i}{G\big(Y_i\big)}|\boldsymbol{X}_i, \boldsymbol{Z}_i, I_i, T_i\Big] = 1. \tag{S1.5}$$

It follows that if $f(\boldsymbol{X}_i, \boldsymbol{Z}_i, \epsilon_i)$ is a function of $(\boldsymbol{X}_i, \boldsymbol{Z}_i, \epsilon_i)$, then

$$E\left[\frac{\delta_i^{(j)}}{G\big(Y_i^{(j)}\big)} f(\boldsymbol{X}_i, \boldsymbol{Z}_i, \epsilon_i)\right] = E\left[f(\boldsymbol{X}_i, \boldsymbol{Z}_i, \epsilon_i)\right], j = 1, 0.$$

Hence, (1.3) holds. □

Lemma 2. *Under assumption (C) of Theorem 1 and regularity condition (Ra), we have a martingale integral representation for* $(\widehat{G} - G)/\widehat{G}$ *such that*

$$\sqrt{n}\{\widehat{G}(t) - G(t)\}/\widehat{G}(t) \tag{S1.6}$$

$$= -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^{\infty}\frac{I\{s \le t\}dM_i^c(s)}{\pi(s)} + o_p(1), \quad \textit{uniformly in } t \le \tau$$

for any $\tau < \tau_G$, *where*

$$\pi(s) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} I\{Y_i \ge s\} = \widetilde{P}\left(Y_1 \ge s\right) \tag{S1.7}$$

in probability, and $M_i^c(s) = I\{Y_i \le s, \delta_i = 0\} - \int_0^s I\{Y_i \ge u\}\, d\Lambda_c(u)$ *(*$s \ge 0$*) is a continuous-time martingale with predictable variation process*

$$\langle M_i^c, M_i^c\rangle(t) = \int_0^t I\{Y_i \ge u\}\, d\Lambda_G(u).$$

Proof. This result is known when $T_i, i = 1, 2, \ldots$ are i.i.d. random variables (Gill (1980), p. 37). We want to show it under the covariate-adaptive design. Note that the adaptive allocation depends only on the covariates. Therefore, given $\mathscr{D} = \sigma(\boldsymbol{X}_i, \boldsymbol{Z}_i, I_i; i = 1, 2, \ldots)$, we have that $T_i, C_i, i = 1, 2, \ldots$ are independent. Note that $C_1, C_2, \ldots$ are i.i.d. random variables with cumulative hazard function Λ_G. We consider C_i to be the survival time and T_i the censoring time. It follows that for given $\mathscr{D}$, $M(s) = \sum_{i=1}^n M_i^c(s)$ $(s \ge 0)$ is a martingale with predictable variation process

$$\langle M, M\rangle(t) = \int_0^t Y(s) d\Lambda_G(s), \quad Y(t) = \sum_{i=1}^n I\{Y_i(t) \ge t\}.$$

Similarly to (3.2.15) of (Gill (1980), p. 37; Shen et al. (2009)),

$$\begin{aligned} &\frac{G(t) - \widehat{G}(t)}{G(t)} = \frac{\widehat{F}_G(t) - F_G(t)}{1 - F_G(t)} \\ &= \int_0^t \frac{\widehat{G}(u-)}{G(u)} \frac{J(u)}{Y(u)} dM(u) - I\{\zeta < t\} \frac{\widehat{G}(\zeta)(G(\zeta) - G(t))}{G(t)G(\zeta)}, \end{aligned} \tag{S1.8}$$

where $\zeta = \zeta_n = \inf\{t > 0 : Y(t) = 0\}$ is a stopping time and $J(s) =$

$I\{Y(s) > 0\}$. For the second term of (S1.8), note that

$$\begin{aligned} P(\zeta < \tau|\mathscr{D}) =& P(Y_i < \tau, i = 1, \ldots, n|\mathscr{D}) = \prod_{i=1}^{n} P(Y_i < \tau|\mathscr{D}) \\ \leq& \left(1 - \min_{j=0,1} P(T_1^{(j)} \geq \tau)G(\tau)\right)^n \doteq (1-a)^n. \end{aligned}$$

Hence, $P(I\{\xi_n < \tau\} \neq 0, i.o.) = 0$ since $\sum_n (1-a)^n < \infty$. Given $\mathscr{D}$, the first term of (S1.8), denoted by $\widetilde{M}(t)$, is a martingale with predictable variation process

$$\langle \widetilde{M}, \widetilde{M} \rangle(t) = \int_0^t \frac{\widehat{G}^2(u-)}{G^2(u)} \frac{J(u)}{Y(u)} d\Lambda_G(u) \leq \frac{\Lambda_G(\tau)}{G^2(\tau)} \frac{J(\tau)}{Y(\tau)}.$$

Note that $Y(\tau)/n \xrightarrow{P} \pi(\tau) = \widetilde{P}(T_1 \geq \tau)G(\tau) > 0$, which implies that $\langle \widetilde{M}, \widetilde{M} \rangle(\tau) \xrightarrow{P} 0$. By Lenglart's inequality ((Lenglart, 1977)), for any $\delta > 0$ and $\epsilon > 0$,

$$P\left(\sup_{t \leq \tau} |\widetilde{M}(t)| > \epsilon|\mathscr{D}\right) \leq \frac{\eta}{\epsilon^2} + P\left(\langle \widetilde{M}, \widetilde{M} \rangle(\tau) > \eta|\mathscr{D}\right).$$

Hence,

$$P\left(\sup_{t \leq \tau} |\widetilde{M}(t)| > \epsilon\right) \leq \frac{\eta}{\epsilon^2} + P\left(\langle \widetilde{M}, \widetilde{M} \rangle(\tau) > \eta\right) \to 0$$

as $n \to \infty$ and then $\eta \to 0$. That is, $\sup_{t \leq \tau} |\widetilde{M}(t)| \xrightarrow{P} 0$. It follows that

$$\sup_{t \leq \tau} \frac{|G(t) - \widehat{G}(t)|}{G(t)} \xrightarrow{P} 0.$$

Now, let

$$\overline{M}(t) = \int_0^t \frac{1}{\sqrt{n}} \left(\frac{\widehat{G}(u-)}{G(u)} \frac{nJ(u)}{Y(u)} - \frac{1}{\pi(u)} \right) dM(u).$$

Then given $\mathscr{D}$, $\overline{M}$ is also a martingale with predictable variation process

$$\langle \overline{M}, \overline{M} \rangle(t) = \int_0^t \left(\frac{\widehat{G}(u-)}{G(u)} \frac{nJ(u)}{Y(u)} - \frac{1}{\pi(u)} \right)^2 \frac{Y(u)}{n} d\Lambda_G(u).$$

Note that $\frac{\widehat{G}(u-)}{G(u)} \xrightarrow{P} 1$ for almost all u and $\frac{Y(u)}{n} \xrightarrow{P} \pi(u)$, which implies that $\langle \overline{M}, \overline{M} \rangle(\tau) \xrightarrow{P} 0$. By applying Lenglart's inequality again we have $\sup_{t \leq \tau} |\overline{M}(t)| \xrightarrow{P} 0$. We conclude that

$$\begin{aligned}
\frac{\sqrt{n}(G(t) - \widehat{G}(t))}{G(t)} =& \frac{1}{\sqrt{n}} \int_0^t \frac{dM(u)}{\pi(u)} + \overline{M}(t) + o_P(1) \\
=& \frac{1}{\sqrt{n}} \int_0^t \frac{dM(u)}{\pi(u)} + o_P(1) \text{ uniformly in } t \leq \tau.
\end{aligned}$$

(S1.6) is proved. □

Lemma 3. (Theorem 3 in the main paper) *Under assumption (C) of Theorem 1 and the regularity conditions (Ra)–(Rc), $\widehat{\boldsymbol{\beta}}$ is a consistent estimate of $\boldsymbol{\beta}^*$, where $\boldsymbol{\beta}^* = (\mu_1 + E[u_1], \mu_2 + E[u_2], \beta_1, \ldots, \beta_{p_1})^T$.*

Proof. Recall that $u_i = \boldsymbol{Z}_i^T \boldsymbol{\gamma} + \epsilon_i$ and

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \left\{ \sum_{i=1}^n \frac{\delta_i \underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}(Y_i)} \right\}^{-1} \sum_{i=1}^n \frac{\delta_i \underline{\boldsymbol{X}}_i u_i}{\widehat{G}(Y_i)}$$

. Then

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}^* + \left\{ \sum_{i=1}^n \frac{\delta_i \underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}(Y_i)} \right\}^{-1} \sum_{i=1}^n \frac{\delta_i \underline{\boldsymbol{X}}_i (u_i - E[u_i])}{\widehat{G}(Y_i)}.$$

Without loss of generality, we can assume that $E[u_i] = 0$. Otherwise, we can replace u_i by $u_i - Eu_i$, μ_1 and μ_2 by $\mu_1 + Eu_i$ and $\mu_2 + Eu_i$ respectively.

We can write

$$\sum_{i=1}^{n}\frac{\delta_i\underline{\boldsymbol{X}}_i u_i}{\widehat{G}\left(Y_i\right)}=\sum_{i=1}^{n}\left\{\frac{1}{\widehat{G}\left(Y_i\right)}-\frac{1}{G\left(Y_i\right)}\right\}\delta_i\underline{\boldsymbol{X}}_i u_i+\sum_{i=1}^{n}\frac{\delta_i\underline{\boldsymbol{X}}_i u_i}{G\left(Y_i\right)}.$$

From the martingale integral representation (S1.6) for $(\widehat{G}-G)/\widehat{G}$, we have

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}\frac{G\left(Y_i\right)-\widehat{G}\left(Y_i\right)}{\widehat{G}\left(Y_i\right)G\left(Y_i\right)}\delta_i\underline{\boldsymbol{X}}_i u_i\\ =&\frac{1}{n\sqrt{n}}\sum_{i=1}^{n}\left\{\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\int_0^\infty\frac{I\left\{s\leq Y_i\right\}dM_j^c(s)}{\pi(s)}\right\}\frac{\delta_i\underline{\boldsymbol{X}}_i u_i}{G\left(Y_i\right)}+o_p\left(n^{-1/2}\right).\end{aligned}$$

Also, by Lemma 1,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{I\{Y_i\geq s\}\delta_i\underline{\boldsymbol{X}}_i u_i}{G\left(Y_i\right)}\xrightarrow{P}\widetilde{E}\left[\frac{I\left\{Y_1\geq s\right\}\delta_1\underline{\boldsymbol{X}}_1 u_1}{G\left(Y_1\right)}\right]=\boldsymbol{B}_1(s), \qquad \text{(S1.9)}$$

where $\boldsymbol{B}_1(s)=\widetilde{E}\left[I\left\{T_1\geq s\right\}\underline{\boldsymbol{X}}_1 u_1\right]$. Hence,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{G\left(Y_i\right)-\widehat{G}\left(Y_i\right)}{\widehat{G}\left(Y_i\right)G\left(Y_i\right)}\delta_i\underline{\boldsymbol{X}}_i u_i=\frac{1}{n}\sum_{i=1}^{n}\int_0^\infty\frac{\boldsymbol{B}_1(s)}{\pi(s)}dM_i^c(s)+o_P\left(n^{-1/2}\right).$$

Note that $\int_0^\infty\frac{\boldsymbol{B}_1(s)}{\pi(s)}dM_i^c(s)$ is a function of Y_i and δ_i with mean zero. By Lemma 1,

$$\frac{1}{n}\sum_{i=1}^{n}\int_0^\infty\frac{\boldsymbol{B}_1(s)}{\pi(s)}dM_i^c(s)\xrightarrow{P}\mathbf{0}.$$

By Lemma 1 again,

$$\frac{1}{n}\sum_{i=1}^{n}\delta_i\underline{\boldsymbol{X}}_i u_i/G(Y_i) = \frac{1}{n}\begin{bmatrix}\sum_{i=1}^{n}\delta_i I_i u_i/G(Y_i)\\ \sum_{i=1}^{n}\delta_i(1-I_i)u_i/G(Y_i)\\ \sum_{i=1}^{n}\delta_i\boldsymbol{X}_i u_i/G(Y_i)\end{bmatrix} \xrightarrow{P} \begin{bmatrix} qE[u_1]\\ (1-q)E[u_1]\\ E[\boldsymbol{X}_1 u_1]\end{bmatrix} = \mathbf{0}$$

by the assumption that $Eu_i = 0$ and $Cov(\boldsymbol{X}_i, u_i) = 0$. Combining these we have

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i\underline{\boldsymbol{X}}_i u_i}{\widehat{G}(Y_i)} \xrightarrow{P} \mathbf{0}. \tag{S1.10}$$

Next, we consider

$$\sum_{i=1}^{n}\frac{\delta_i\underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}(Y_i)} = \sum_{i=1}^{n}\left\{\frac{1}{\widehat{G}(Y_i)} - \frac{1}{G(Y_i)}\right\}\delta_i\underline{\boldsymbol{X}}_i^{\otimes 2} + \sum_{i=1}^{n}\frac{\delta_i\underline{\boldsymbol{X}}_i^{\otimes 2}}{G(Y_i)}.$$

Similarly, by noting that

$$\frac{1}{n}\sum_{i=1}^{n}\frac{I\{Y_i \geq s\}\delta_i\underline{\boldsymbol{X}}_i^{\otimes 2}}{G(Y_i)} \xrightarrow{P} \widetilde{E}\left[\frac{I\{Y_1 \geq s\}\delta_1\underline{\boldsymbol{X}}_1^{\otimes 2}}{G(Y_1)}\right] = \boldsymbol{B}_2(s),$$

where $\boldsymbol{B}_2(s) = \widetilde{E}\left[I\{T_1 \geq s\}\underline{\boldsymbol{X}}_1^{\otimes 2}\right]$, we have

$$\frac{1}{n}\sum_{i=1}^{n}\frac{G(Y_i) - \widehat{G}(Y_i)}{\widehat{G}(Y_i)G(Y_i)}\delta_i\underline{\boldsymbol{X}}_i^{\otimes 2} = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\infty}\frac{\boldsymbol{B}_2(s)}{\pi(s)}dM_i^c(s) + o_P(1)$$

and

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i \underline{\boldsymbol{X}}_i^{\otimes 2}}{G\left(Y_i\right)} \\
&\quad = \frac{1}{n}\frac{\delta_i}{G\left(Y_i\right)}\begin{bmatrix} diag(I_i, 1-I_i) & (I_i, 1-I_i)^T \boldsymbol{X}_i \\ (I_i, 1-I_i)\boldsymbol{X}_i^T & \boldsymbol{X}_i^{\otimes 2} \end{bmatrix} \\
&\quad \xrightarrow{P} \begin{bmatrix} diag(q, 1-q) & (q, 1-q)^T E\boldsymbol{X}_1^T \\ E\boldsymbol{X}_1(q, 1-q) & E\boldsymbol{X}_1^{\otimes 2} \end{bmatrix} \doteq \Gamma_{\boldsymbol{\beta}}. \qquad \text{(S1.11)}
\end{aligned}$$

The inverse matrix of $\Gamma_{\boldsymbol{\beta}}$ is

$$\Gamma_{\boldsymbol{\beta}}^{-1} = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^T & (Var(\boldsymbol{X}_1))^{-1} \end{bmatrix},$$

$\Gamma_{11} = diag(\frac{1}{q}, \frac{1}{1-q}) + (1,1)^T(1,1)E\boldsymbol{X}_1^T(Var(\boldsymbol{X}_1))^{-1}E\boldsymbol{X}_1$, $\Gamma_{12} = -(1,1)^T E\boldsymbol{X}_1^T(Var(\boldsymbol{X}_1))^{-1}$.

Note that $\int_0^\infty \frac{\boldsymbol{B}_2(s)}{\pi(s)} dM_i^c(s)$ is a function of Y_i and δ_i, with mean zero. By Lemma 1,

$$\frac{1}{n}\sum_{i=1}^{n}\int_0^\infty \frac{\boldsymbol{B}_2(s)}{\pi(s)} dM_i^c(s) \xrightarrow{P} \mathbf{0}.$$

By Slutsky's theorem

$$\widehat{\Gamma}_{\boldsymbol{\beta}} = \frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i \underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}\left(Y_i\right)} \xrightarrow{P} \Gamma_{\boldsymbol{\beta}}. \qquad \text{(S1.12)}$$

Hence,

$$\left\{\sum_{i=1}^{n}\frac{\delta_i \underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}\left(Y_i\right)}\right\}^{-1}\sum_{i=1}^{n}\frac{\delta_i \underline{\boldsymbol{X}}_i u_i}{\widehat{G}\left(Y_i\right)} \xrightarrow{P} \mathbf{0}.$$

It follows that

$$\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} \xrightarrow{P} \mathbf{0}.$$

The proof of Lemma 3 is complete. □

Lemma 4. (Theorem 4 in the main paper) *Under the conditions of Theorem 1,*

$$\sqrt{n}\big(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\big) \xrightarrow{D} N\big(0, \Gamma_{\boldsymbol{\beta}}^{-1}\Sigma_{\boldsymbol{\beta}}\Gamma_{\boldsymbol{\beta}}^{-1}\big), \tag{S1.13}$$

where $\Gamma_{\boldsymbol{\beta}}$ *and* $\Sigma_{\boldsymbol{\beta}}$ *are defined as in* (S1.11) *and*

$$\Sigma_{\boldsymbol{\beta}} = \widetilde{E}\Big[\frac{\underline{\boldsymbol{X}}_i^{\otimes 2}u_i^2}{G(T_i \wedge \tau_G)}\Big] - q(1-q)\boldsymbol{L}\boldsymbol{L}^T E(\breve{u}_i)^2 - \int_0^\infty \frac{\boldsymbol{B}_1^{\otimes 2}(u)}{\pi(u)} d\Lambda_c(u), \tag{S1.14}$$

respectively, with the estimators $\widehat{\Gamma}_{\boldsymbol{\beta}}, \widehat{\Sigma}_{\boldsymbol{\beta}}$ *given by* (S1.12) *and* (S1.22), *respectively.*

Proof. Without loss of generality, we assume $E[u_i] = 0$. Note that

$$\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} = \left\{\frac{1}{n}\sum_{i=1}^n \frac{\delta_i \underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}(Y_i)}\right\}^{-1} \frac{1}{n}\sum_{i=1}^n \frac{\delta_i \underline{\boldsymbol{X}}_i u_i}{\widehat{G}(Y_i)}.$$

We have shown in the proof of Lemma 3 that

$$\widehat{\Gamma}_{\boldsymbol{\beta}} = \frac{1}{n}\sum_{i=1}^n \frac{\delta_i \underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}(Y_i)} \xrightarrow{P} \Gamma_{\boldsymbol{\beta}}. \tag{S1.15}$$

Let

$$\boldsymbol{A}_n = \frac{1}{n}\sum_{i=1}^n \frac{\delta_i \underline{\boldsymbol{X}}_i u_i}{\widehat{G}(Y_i)}, \quad \boldsymbol{B}_n = \left\{\widehat{\Gamma}_{\boldsymbol{\beta}}^{-1} - \Gamma_{\boldsymbol{\beta}}^{-1}\right\}\boldsymbol{A}_n.$$

Then

$$\begin{aligned}\sqrt{n}\boldsymbol{A}_n &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \left\{\frac{1}{\widehat{G}(Y_i)} - \frac{1}{G(Y_i)}\right\}\delta_i \underline{\boldsymbol{X}}_i u_i + \frac{1}{\sqrt{n}}\sum_{i=1}^n \frac{\delta_i \underline{\boldsymbol{X}}_i u_i}{G(Y_i)} \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n \left[\int_0^\infty \frac{\boldsymbol{B}_1(s)}{\pi(s)} dM_i^c(s) + \frac{\delta_i \underline{\boldsymbol{X}}_i u_i}{G(Y_i)}\right] + o_P(1).\end{aligned}$$

The first term in the above bracket is a sequence of martingale differences. However, the second term needs to be modified. Let

$$\boldsymbol{V}_{i,1} = \int_0^\infty \frac{\boldsymbol{B}_1(s)}{\pi(s)} dM_i^c(s), \quad \boldsymbol{V}_{i,2} = \frac{\delta_i \underline{\boldsymbol{X}}_i u_i}{G\left(Y_i\right)} - \big(I_i - q\big)\boldsymbol{L}\breve{u}_i.$$

By assumption (B),

$$\sum_{i=1}^n (I_i - q)E[u_i|W_i] = \sum_{i=1}^n (I_i - q)\breve{u}_i = o_P(\sqrt{n}).$$

It follows that

$$\sqrt{n}\boldsymbol{A}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\boldsymbol{V}_{i,1} + \boldsymbol{V}_{i,2}\right) + o_P\left(1\right).$$

Since for given $\mathscr{D} = \sigma(I_j, \boldsymbol{X}_j, \boldsymbol{Z}_j, j = 1, \ldots\}$, $T_1, C_1, T_2, C_2, \ldots$ are independent and $M_i^c(t)$ $(t \geq 0)$ is a martingale, it is easy to show that

$$E[\boldsymbol{V}_{i,1}|\mathscr{F}_{i-1}, \boldsymbol{X}_i, \boldsymbol{Z}_i, I_i] = E\big[E[\boldsymbol{V}_{i,1}|\mathscr{D}]\mathscr{F}_{i-1}, \boldsymbol{X}_i, \boldsymbol{Z}_i, I_i\big] = 0. \qquad \text{(S1.16)}$$

Furthermore, $E[\delta_i X_{i,t} u_i / G(Y_i)|\mathscr{F}_{i-1}, W_i] = E[X_{i,t} u_i | W_i]$ and

$$E\Big[\delta_i I_i u_i/G(Y_i) - (I_i - q)\breve{u}_i\Big|\mathscr{F}_{i-1}, W_i\Big] = E\Big[I_i(u_i - \breve{u}_i) + q\breve{u}_i\Big|\mathscr{F}_{i-1}, W_i\Big]$$

$$= E[I_i|\mathscr{F}_{i-1}, W_i]\big(E[u_i - \breve{u}_i|W_i]\big) + q\breve{u}_i = q\breve{u}_i = qE[u_i|W_i].$$

Similarly,

$$E\big[\delta_i(1 - I_i)u_i/G(Y_i) + (I_i - q)\breve{u}_i\big|\mathscr{F}_{i-1}, W_i\big]$$

$$= (1 - q)\breve{u}_i = (1 - q)E[u_i|W_i].$$

It follows that

$$E[\boldsymbol{V}_{i,2}|\mathscr{F}_{i-1}] = E[(q, 1-q, \boldsymbol{X}_i^T)^T u_i] = \big(qEu_i, (1-q)Eu_i, E[\boldsymbol{X}_i^T u_i]\big) = 0$$

by the assumptions that $E[u_i] = 0$ and $Cov(\boldsymbol{X}_i, u_i) = 0$. Hence, $\{\boldsymbol{V}_{i,1} + \boldsymbol{V}_{i,2}, \mathscr{F}_{i-1}; i = 1, 2, \ldots, n\}$ is a sequence of martingale differences. Next, we verify the conditions for the central limit theorem of martingale differences.

It can be checked that

$$\begin{aligned}\widetilde{E}[\boldsymbol{V}_{i,1}^{\otimes 2}] =&\widetilde{E}\left[\int_0^\infty \frac{\boldsymbol{B}_1^{\otimes 2}(u)}{\pi^2(u)} I\{Y_i \geq u\} d\Lambda_c(u)\right] \\ =& \int_0^\infty \frac{\boldsymbol{B}_1^{\otimes 2}(u)}{\pi(u)} d\Lambda_c(u) \doteq \Sigma_{\boldsymbol{\beta},G},\end{aligned} \tag{S1.17}$$

$$\begin{aligned}\widetilde{E}[\boldsymbol{V}_{i,2}^{\otimes 2}] =&\widetilde{E}\Big[\frac{\delta_i^2 \underline{\boldsymbol{X}}_i^{\otimes 2} u_i^2}{G^2(Y_i)}\Big] - q(1-q)\boldsymbol{L}\boldsymbol{L}^T E(\breve{u}_i)^2 \\ =&\widetilde{E}\Big[\frac{\underline{\boldsymbol{X}}_i^{\otimes 2} u_i^2}{G(T_i \wedge \tau_G)}\Big] - q(1-q)\boldsymbol{L}\boldsymbol{L}^T E(\breve{u}_i)^2 \\ \doteq& \Sigma_{\boldsymbol{\beta},u} - \Sigma_{\boldsymbol{\beta},\breve{z}},\end{aligned} \tag{S1.18}$$

and, by (S1.16),

$$\begin{aligned}\widetilde{E}[\boldsymbol{V}_{i,1}\boldsymbol{V}_{i,2}^T] =&\widetilde{E}[\boldsymbol{V}_{i,1}\frac{\delta_i \underline{\boldsymbol{X}}_i^T u_i}{G(Y_i)}] + \widetilde{E}\left[\boldsymbol{V}_{i,1}\breve{u}_i(I_i - q)\boldsymbol{L}^T\right] \\ =&\widetilde{E}\left[\int_0^\infty \frac{\boldsymbol{B}_1(u)(1-\delta_i)}{\pi(u)} dI\{Y_i \leq u\} \frac{\delta_i \underline{\boldsymbol{X}}_i^T u_i}{G(Y_i)}\right] \\ &- \widetilde{E}\left[\int_0^\infty \frac{\boldsymbol{B}_1(u)}{\pi(u)} I\{Y_i \geq u\} d\Lambda_G(u) \frac{\delta_i \underline{\boldsymbol{X}}_i^T u_i}{G(Y_i)}\right] + \boldsymbol{0} \\ =&-\widetilde{E}\left[\int_0^\infty \frac{\boldsymbol{B}_1(u)}{\pi(u)} I\{Y_i \geq u\} \frac{\delta_i \underline{\boldsymbol{X}}_i^T u_i}{G(Y_i)} d\Lambda_G(u)\right] \\ =&-\int_0^\infty \frac{\boldsymbol{B}_1^{\otimes 2}(u)}{\pi(u)} d\Lambda_G(u) = -\Sigma_{\boldsymbol{\beta},G}.\end{aligned}$$

Hence, by Lemma 1,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{V}_{i,1}+\boldsymbol{V}_{i,2})^{\otimes 2}=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{V}_{i,2}^{\otimes 2}+\boldsymbol{V}_{i,1}^{\otimes 2}-2\boldsymbol{V}_{i,1}\boldsymbol{V}_{i,2}^{T})$$

$$=\Sigma_{\boldsymbol{\beta},u}-\Sigma_{\boldsymbol{\beta},\check{z}}-\Sigma_{\boldsymbol{\beta},G}\doteq\Sigma_{\boldsymbol{\beta}}\ \text{ in probability.} \tag{S1.19}$$

Also,

$$\frac{1}{n}E\left[\max_{i\le n}\|\boldsymbol{V}_{i,1}+\boldsymbol{V}_{i,2}\|^2\right]\le\frac{1}{nq(1-q)}\widetilde{E}\left[\max_{i\le n}\|\boldsymbol{V}_{i,1}+\boldsymbol{V}_{i,2}\|^2\right]$$

$$\le\frac{1}{\sqrt{n}q(1-q)}+\frac{1}{q(1-q)}\widetilde{E}\left[\|\boldsymbol{V}_{1,1}+\boldsymbol{V}_{1,2}\|^2 I\{\|\boldsymbol{V}_{1,1}+\boldsymbol{V}_{1,2}\|^2\ge\sqrt{n}\}\right]\to 0.$$

By the central limit theorem for martingale differences (see Theorem 3.2 of Hall and Heyde (1980)),

$$\sqrt{n}\boldsymbol{A}_n\xrightarrow{D}N(0,\Sigma_{\boldsymbol{\beta}}).$$

It follows that

$$\begin{aligned}\sqrt{n}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})=&(\widehat{\Gamma}_{\boldsymbol{\beta}}^{-1}-\Gamma_{\boldsymbol{\beta}}^{-1})\sqrt{n}\boldsymbol{A}_n+\Gamma_{\boldsymbol{\beta}}^{-1}\sqrt{n}\boldsymbol{A}_n\\ =&\Gamma_{\boldsymbol{\beta}}^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(\boldsymbol{V}_{i,1}+\boldsymbol{V}_{i,2})+o_P(1)\\ &\xrightarrow{D}N\big(0,\Gamma_{\boldsymbol{\beta}}^{-1}\Sigma_{\boldsymbol{\beta}}\Gamma_{\boldsymbol{\beta}}^{-1}\big).\end{aligned}$$

The proof of (S1.13) is complete.

Recall that (S1.17), (S1.9), and (S1.7) hold. A consistent estimator of $\Sigma_{\boldsymbol{\beta},G}$ is

$$\widehat{\Sigma}_{\boldsymbol{\beta},G}=\int_0^\infty\frac{\widehat{\boldsymbol{B}}_1^{\otimes 2}(s)}{\widehat{\pi}(s)}d\widehat{\Lambda}_G(s), \tag{S1.20}$$

where $\widehat{\boldsymbol{B}}_1(s) = \frac{1}{n}\sum_{i=1}^n \delta_i I\{Y_i \geq s\}\underline{\boldsymbol{X}}_i(\log Y_i - \underline{\boldsymbol{X}}_i\widehat{\boldsymbol{\beta}})/\widehat{G}(Y_i)$, $\widehat{\pi}(s) = \frac{1}{n}\sum_{i=1}^n I\{Y_i \geq s\}$, and $\widehat{\Lambda}_G(s)$ is the Nelson estimate for the cumulative hazard function $\Lambda_G(s)$ of C.

As for $\Sigma_{\boldsymbol{\beta},u} - \Sigma_{\boldsymbol{\beta},\check{z}}$, by (S1.19), $\Sigma_{\boldsymbol{\beta},u} - \Sigma_{\boldsymbol{\beta},\check{z}}$ is a limit of $\frac{1}{n}\sum_{i=1}^n \boldsymbol{V}_{i,2}^{\otimes 2}$ with $\boldsymbol{V}_{i,2} = \frac{\delta_i \underline{\boldsymbol{X}}_i u_i}{G(Y_i)} - \boldsymbol{L}(I_i - q)\check{u}_i$. To obtain an estimator, we should replace the unobservable terms u_i and $G(\cdot)$ by their estimators $\widehat{u}_i = \log Y_i - \underline{X}_i\widehat{\boldsymbol{\beta}}$ and $\widehat{G}(\cdot)$. However, the term $\check{u}_i = E[\boldsymbol{Z}_i\boldsymbol{\gamma} + \epsilon_i|W_i] - E[\boldsymbol{Z}_i\boldsymbol{\gamma} + \epsilon_i]$ is neither observable nor estimable under the working AFT model (2.2). When W_i is independent of $\boldsymbol{Z}_i$, $\check{u}_i = 0$. In general, the $\check{u}_i$s have zero means and $\sum_{i=1}^n (I_i - 1/2)\check{u}_i = o_P(\sqrt{n})$ by the assumption (B). It is reasonable to replace $\check{u}_i$ by zero and obtain the estimator of $\Sigma_{\boldsymbol{\beta},u} - \Sigma_{\boldsymbol{\beta},\check{z}}$ as

$$\widehat{\Sigma}_{\boldsymbol{\beta},WLS} = \frac{1}{n}\sum_{i=1}^n \frac{\delta_i}{\widehat{G}^2(Y_i)}\underline{\boldsymbol{X}}_i\underline{\boldsymbol{X}}_i^T(\log Y_i - \underline{\boldsymbol{X}}_i\widehat{\boldsymbol{\beta}})^2. \tag{S1.21}$$

The estimator of $\Sigma_{\boldsymbol{\beta}}$ is now given by

$$\widehat{\Sigma}_{\boldsymbol{\beta}} = \widehat{\Sigma}_{\boldsymbol{\beta},WLS} - \widehat{\Sigma}_{\boldsymbol{\beta},G}. \qquad \square \tag{S1.22}$$

Proof of Theorem 1. Note that $\boldsymbol{L}^T\Gamma_\beta^{-1} = (q^{-1}, -(1-q)^{-1}, 0, \ldots, 0)$ and $\boldsymbol{L}^T\Gamma_\beta^{-1}\underline{\boldsymbol{X}_i} = (I_i - q)/(q(1-q))$. For the test statistic the nominator part is equal to

$$
\begin{aligned}
&\sqrt{n}\big[(\widehat{\mu_1}-\widehat{\mu_2})-(\mu_1-\mu_2)\big]=\boldsymbol{L}^T\sqrt{n}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{L}^T\Gamma_{\boldsymbol{\beta}}^{-1}\left(\boldsymbol{V}_{i,1}+\boldsymbol{V}_{i,2}\right)+o_P\left(1\right)\\
=&\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{\boldsymbol{L}\Gamma_{\boldsymbol{\beta}}^{-1}\boldsymbol{B}_1(u)}{\pi(u)}dM_i^c(u)+\frac{I_i-q}{q(1-q)}\Big(\frac{\delta_i u_i}{G(Y_i)}-\breve{u}_i\Big)\right]+o_P(1).
\end{aligned}
\tag{S1.23}
$$

$$
\xrightarrow{D} N\Big(0,\boldsymbol{L}^T\Gamma_{\boldsymbol{\beta}}^{-1}\Sigma_{\boldsymbol{\beta}}\Gamma_{\boldsymbol{\beta}}^{-1}\boldsymbol{L}\Big).
$$

It can easily be seen that

$$
\begin{aligned}
&\boldsymbol{L}^T\Gamma_{\boldsymbol{\beta}}^{-1}\Sigma_{\boldsymbol{\beta}}\Gamma_{\boldsymbol{\beta}}^{-1}\boldsymbol{L}\\
=&\boldsymbol{L}^T\Gamma_{\beta}^{-1}\Sigma_{\boldsymbol{\beta},u}\Gamma_{\beta}^{-1}\boldsymbol{L}-\boldsymbol{L}^T\Gamma_{\beta}^{-1}\Sigma_{\boldsymbol{\beta},G}\Gamma_{\beta}^{-1}\boldsymbol{L}\\
=&-\int_0^\infty\frac{(\boldsymbol{L}^T\Gamma_{\beta}^{-1}\boldsymbol{B}_1(u))^2}{\pi(u)}d\Lambda_G(u)\\
&+\frac{1}{q^2(1-q)^2}\widetilde{E}\left[\frac{(I_1-q)^2\delta_1u_1^2}{G^2(Y_1)}\right]-\frac{1}{q(1-q)}E(\breve{u}_1)^2.
\end{aligned}
$$

Since

$$
\begin{aligned}
\boldsymbol{L}^T\Gamma_{\beta}^{-1}\boldsymbol{B}_1(s)&=\widetilde{E}\left[I\{T_1\ge s\}\boldsymbol{L}^T\Gamma_{\beta}^{-1}\underline{\boldsymbol{X}}_1u_1\right]\\
&=\widetilde{E}\left[I\{T_1\ge s\}(I_1-q)u_1\right]/(q(1-q))\\
&=E\left[\Big(I\{T_1^{(1)}\ge s\}-I\{T_1^{(0)}\ge s\}\Big)u_1\right]\\
&=E\left[\Big(I\{T_1^{(1)}\ge s\}-I\{T_1^{(1)}e^{\mu_2-\mu_1}\ge s\}\Big)u_1\right],
\end{aligned}
$$

we have $\boldsymbol{L}^T\Gamma_{\beta}^{-1}\boldsymbol{B}_1(s)=0$ and $\widetilde{E}\big[(I_1-q)^2\delta_1u_1^2/G^2(Y_1)\big]=q(1-q)E\big[\delta_1u_1^2/G^2(Y_1)|H_0\big]$

under the null hypothesis $\mu_1 = \mu_2$. It follows that

$$\boldsymbol{L}^T \Gamma_{\boldsymbol{\beta}} \Sigma_{\boldsymbol{\beta}} \Gamma_{\boldsymbol{\beta}}^{-1} \boldsymbol{L} = \sigma_{\delta,G}^2 / (q(1-q)).$$

We conclude that, under H_0, the first term in (S1.23) is zero and

$$\begin{aligned} &\sqrt{n}\{(\widehat{\mu}_1 - \widehat{\mu}_2) - (\mu_1 - \mu_2)\} \\ =&\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{I_i - q}{q(1-q)} \Big(\frac{\delta_i u_i}{G(Y_i)} - \breve{u}_i \Big) + o_P(1) \qquad \text{(S1.24)} \\ &\xrightarrow{D} N\left(0, \sigma_{\delta,G}^2 / (q(1-q))\right). \qquad \text{(S1.25)} \end{aligned}$$

On the other side we show that the estimator $\widehat{\mathrm{Var}}(\boldsymbol{L}^T \widehat{\boldsymbol{\beta}})$ we use for the variance of $\boldsymbol{L}^T \widehat{\boldsymbol{\beta}}$ is inflated. Recall that

$$n\widehat{\mathrm{Var}}(\boldsymbol{L}^T \widehat{\boldsymbol{\beta}}) = \boldsymbol{L}^T \widehat{\Gamma}_{\boldsymbol{\beta}}^{-1} \Big(\widehat{\Sigma}_{\boldsymbol{\beta},u} - \widehat{\Sigma}_{\boldsymbol{\beta},G} \Big) \widehat{\Gamma}_{\boldsymbol{\beta}}^{-1} \boldsymbol{L},$$

$$\boldsymbol{L}^T \widehat{\Gamma}_{\boldsymbol{\beta}}^{-1} \xrightarrow{P} \boldsymbol{L}^T \Gamma_{\boldsymbol{\beta}}^{-1},$$

and $\widehat{\Sigma}_{\boldsymbol{\beta},G}$ is a consistent estimator of $\Sigma_{\boldsymbol{\beta},G}$. Therefore,

$$\begin{aligned} &\boldsymbol{L}^T \widehat{\Gamma}_{\boldsymbol{\beta}}^{-1} \widehat{\Sigma}_{\boldsymbol{\beta},G} \widehat{\Gamma}_{\boldsymbol{\beta}}^{-1} \boldsymbol{L} \\ \xrightarrow{P} &\int_0^\infty \frac{(\boldsymbol{L}^T \Gamma_{\boldsymbol{\beta}}^{-1} \boldsymbol{B}_1(u))^2}{\pi(u)} d\Lambda_G(u) = 0 \quad \text{under } H_0. \end{aligned}$$

It follows that

$$n\widehat{\mathrm{Var}}(\boldsymbol{L}^T \widehat{\boldsymbol{\beta}}) = n\widehat{\mathrm{Var}}_{WLS}(\boldsymbol{L}^T \widehat{\boldsymbol{\beta}}) + o_P(1).$$

For $\widehat{\Sigma}_{\boldsymbol{\beta},WLS}$, we have

$$\widehat{\Sigma}_{\boldsymbol{\beta},WLS} = \frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i^2\underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}^2(Y_i)}u_i^2 + \frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i^2\underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}^2(Y_i)}\left\{\underline{\boldsymbol{X}}_i^T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})\right\}^2 - \frac{2}{n}\sum_{i=1}^{n}\frac{\delta_i^2\underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}^2(Y_i)}u_i\underline{\boldsymbol{X}}_i^T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}).$$

Note that

$$\widetilde{E}\left[\frac{\delta_i^2(1+X_{i,t}^4+Z_{i,t}^4+\epsilon_i^4)}{G^2(Y_i)}\right] = \widetilde{E}\left[\frac{1+X_{i,t}^4+Z_{i,t}^4+\epsilon_i^4}{G(T_i\wedge\tau_G)}\right] < \infty.$$

Using arguments similar to those used to show (S1.12) we have

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i^2\underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}(Y_i)^2}u_i^2 \xrightarrow{P} \widetilde{E}\left[\frac{\delta_1\underline{\boldsymbol{X}}_1^{\otimes 2}}{G^2(Y_1)}u_1^2\right] = \widetilde{E}\left[\frac{\underline{\boldsymbol{X}}_1^{\otimes 2}}{G(T_1\wedge\tau_G)}u_1^2\right],$$

and

$$\left\|\frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i^2\underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}(Y_i)^2}\left\{\underline{\boldsymbol{X}}_i^T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})\right\}^2\right\| \le \frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i^2(1+\|\boldsymbol{X}_i\|^2)^2}{\widehat{G}^2(Y_i)}\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\|^2 = O_P(1)\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\|^2 \xrightarrow{P} 0,$$

$$\left\|\frac{2}{n}\sum_{i=1}^{n}\frac{\delta_i^2\underline{\boldsymbol{X}}_i^{\otimes 2}}{\widehat{G}^2(Y_i)}u_i\boldsymbol{X}_i^T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})\right| = O_P(1)\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\| \xrightarrow{P} 0,$$

since $\widehat{\boldsymbol{\beta}} \xrightarrow{P} \boldsymbol{\beta}$. It follows that

$$\widehat{\Sigma}_{\boldsymbol{\beta},WLS} \xrightarrow{P} \widetilde{E}\left[\frac{\delta_1\underline{\boldsymbol{X}}_1^{\otimes 2}}{G^2(Y_1)}u_1^2\right] = \Sigma_{\boldsymbol{\beta},u}. \tag{S1.26}$$

Hence,

$$
\begin{aligned}
& n\widehat{\mathrm{Var}}(\boldsymbol{L}^T\widehat{\boldsymbol{\beta}}) = n\widehat{\mathrm{Var}}_{WLS}(\boldsymbol{L}^T\widehat{\boldsymbol{\beta}}) + o_P(1) \qquad \text{(S1.27)} \\
=& \boldsymbol{L}^T(\widehat{\Gamma}_{\boldsymbol{\beta}}^{-1} - \Gamma_{\boldsymbol{\beta}}^{-1})\widehat{\Sigma}_{\boldsymbol{\beta},u}(\widehat{\Gamma}_{\boldsymbol{\beta}}^{-1} - \Gamma_{\boldsymbol{\beta}}^{-1})\boldsymbol{L} \\
& + 2\boldsymbol{L}^T(\widehat{\Gamma}^{-1} - \Gamma_{\boldsymbol{\beta}}^{-1})\widehat{\Sigma}_{\boldsymbol{\beta},u}\Gamma_{\boldsymbol{\beta}}^{-1}\boldsymbol{L} + \boldsymbol{L}^T\Gamma_{\boldsymbol{\beta}}^{-1}\widehat{\Sigma}_{\boldsymbol{\beta},u}\Gamma_{\boldsymbol{\beta}}^{-1}\boldsymbol{L} + o_P(1) \\
=& \frac{1}{n(q^2(1-q)^2}\sum_{i=1}^n \frac{\delta_i(I_i-q)^2}{\widehat{G}^2(Y_i)}(\log Y_i - \underline{X}_i\widehat{\boldsymbol{\beta}})^2 + o_P(1) \\
=& \frac{1}{nq^2(1-q)^2}\sum_{i=1}^n \frac{\delta_i(I_i-q)^2}{\widehat{G}^2(Y_i)}(\log Y_i - \underline{X}_i\boldsymbol{\beta})^2 + o_P(1) \\
\xrightarrow{P}& \frac{1}{q^2(1-q)^2}\widetilde{E}\left[\frac{(I_1-q)^2\delta_1 u_1^2}{G^2(Y_1)}\right] = \frac{1}{q(1-q)}\sigma_{z,G}^2 \quad \text{under } H_0.
\end{aligned}
$$

By combining (S1.25) and (S1.27), we obtain

$$
\frac{\boldsymbol{L}^T\widehat{\boldsymbol{\beta}} - (\mu_1 - \mu_2)}{\{\widehat{\mathrm{Var}}(\boldsymbol{L}^T\widehat{\boldsymbol{\beta}})\}^{1/2}} \xrightarrow{D} N(0,\tau^2), \quad \tau^2 = \frac{\sigma_{\delta,G}^2}{\sigma_{z,G}^2} \quad \text{under } H_0. \qquad \text{(S1.28)}
$$

The proof of Theorem 1 (i) is complete.

Now, suppose that $\mu_2 = \mu_1 - \delta/\sqrt{n}$. Let $h_i = h(T_i, I_i, C_i, \boldsymbol{X}_i, \boldsymbol{Z}_i)$ be a Borel function. Note that

$$
\begin{aligned}
& \frac{1}{n}\sum_{i=1}^n (1-I_i)\left|h_i^{(0)} - h(T_i^{(1)}, 0, C_i, \boldsymbol{X}_i, \boldsymbol{Z}_i)\right| \\
\le& \frac{1}{n}\sum_{i=1}^n \sup_{|x|\le\epsilon}\left|h(T_i^{(1)}e^x, 0, C_i, \boldsymbol{X}_i, \boldsymbol{Z}_i) - h(T_i^{(1)}, 0, C_i, \boldsymbol{X}_i, \boldsymbol{Z}_i)\right| \\
\xrightarrow{P}& E\left[\sup_{|x|\le\epsilon}\left|h(T_1^{(1)}e^x, 0, C_1, \boldsymbol{X}_1, \boldsymbol{Z}_1) - h(T_1^{(1)}, 0, C_1, \boldsymbol{X}_1, \boldsymbol{Z}_1)\right|\right] \\
& \to 0 \text{ as } n \to \infty \text{ and then } \epsilon \to 0.
\end{aligned}
$$

It follows that

$$\frac{1}{n}\sum_{i=1}^{n}(1-I_i)h_i = \frac{1}{n}\sum_{i=1}^{n}(1-I_i)h_i^{(0)}$$
$$=\frac{1}{n}\sum_{i=1}^{n}(1-I_i)h(T_i^{(1)},0,C_i,\boldsymbol{X}_i,\boldsymbol{Z}_i)+o_P(1)$$
$$\xrightarrow{P}(1-q)Eh(T_1^{(1)},0,C_1,\boldsymbol{X}_1,\boldsymbol{Z}_1)=(1-q)E[h_1^{(0)}|H_0]$$

and

$$\frac{1}{n}\sum_{i=1}^{n}I_ih_i = \frac{1}{n}\sum_{i=1}^{n}I_ih_i^{(1)} \xrightarrow{P} qE[h_1^{(1)}|H_0],$$
$$\frac{1}{n}\sum_{i=1}^{n}h_i \xrightarrow{P}\widetilde{E}[h_1|H_0].$$

In particular,

$$\frac{1}{n}\sum_{i=1}^{n}\frac{I\{Y_i\ge s\}\delta_i\underline{\boldsymbol{X}}_iu_i}{G\left(Y_i\right)} \xrightarrow{P} \widetilde{E}\left[\frac{I\left\{Y_1\ge s\right\}\delta_1\underline{\boldsymbol{X}}_1u_1}{G\left(Y_1\right)}\Big|H_0\right]=\boldsymbol{B}_1(s),$$

with $\boldsymbol{L}^T\Gamma_{\boldsymbol{\beta}}^{-1}\boldsymbol{B}_1(s) = \widetilde{E}\left[I\{T_1\ge s\}\boldsymbol{L}^T\Gamma_{\boldsymbol{\beta}}^{-1}\underline{\boldsymbol{X}}_1u_1|H_0\right] = 0$. The arguments used to show (S1.24), (S1.25) and (S1.27) remain valid with $\sigma_{z,G}^2 = E\left[\frac{\delta_1u_1^2}{G^2(Y_1)}|H_0\right]$ and $\sigma_\delta^2=\sigma_{u,G}^2-E(\breve{u}_1)^2$. Hence,

$$\frac{\boldsymbol{L}\widehat{\boldsymbol{\beta}}-(\mu_1-\mu_2)}{\{\widehat{\mathrm{Var}}(\boldsymbol{L}\widehat{\boldsymbol{\beta}})\}^{1/2}} \xrightarrow{D} N(0,\tau^2),\quad \tau^2=\frac{\sigma_{\delta,G}^2}{\sigma_{z,G}^2}\quad \text{under } H_A.$$

Note that $(\mu_1-\mu_2)/\{\widehat{\mathrm{Var}}(\boldsymbol{L}\widehat{\boldsymbol{\beta}})\}^{1/2}\to\delta\sqrt{q(1-q)}/\sigma_{u,G}$. We conclude that

$$\frac{\boldsymbol{L}\widehat{\boldsymbol{\beta}}}{\{\widehat{\mathrm{Var}}(\boldsymbol{L}\widehat{\boldsymbol{\beta}})\}^{1/2}} \xrightarrow{D} N(\Delta,\tau^2),\quad \tau^2=\frac{\sigma_{\delta,G}^2}{\sigma_{z,G}^2},\Delta=\frac{\delta\sqrt{q(1-q)}}{\sigma_{u,G}}\quad \text{under } H_A.$$

The proof of Theorem 1 is complete. □

Proof of Theorem 2

From (S1.28) in Theorem 1, if $\gamma_j = 0, j = 1, 2, \ldots, p_2$, the terms involving $\boldsymbol{Z}$ will disappear, so $u_i = \epsilon_i$, $\breve{u}_i = 0$ and $\tau^2 = 1$. That is, when the correct model is used in the analysis, the hypothesis test of (2.4) can achieve the correct Type I error. If $E[\boldsymbol{Z}_1^T + \epsilon_1 \boldsymbol{\gamma} | W_1] \not\equiv Const$, we have $\tau < 1$. Then the hypothesis test of (2.4) is conservative. □

Proof of Theorem 5

By Lemma 4,

$$\sqrt{n}\mathcal{P}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{D} N\big(0, \mathcal{P}\Gamma_{\boldsymbol{\beta}}^{-1}\Sigma_{\boldsymbol{\beta}}\Gamma_{\boldsymbol{\beta}}^{-1}\mathcal{P}^T\big).$$

On the other hand,

$$\widehat{\boldsymbol{M}} \doteq n\widehat{\mathrm{Var}}(\widehat{\boldsymbol{\beta}}) = \widehat{\Gamma}_{\boldsymbol{\beta}}^{-1}\Big(\widehat{\Sigma}_{\boldsymbol{\beta},WLS} - \widehat{\Sigma}_{\boldsymbol{\beta},G}\Big)\widehat{\Gamma}_{\boldsymbol{\beta}}^{-1} \to \boldsymbol{M},$$

by (S1.26), where $\boldsymbol{M} = \Gamma_{\boldsymbol{\beta}}^{-1}(\Sigma_{\boldsymbol{\beta},u} - \Sigma_{\boldsymbol{\beta},G})\Gamma_{\boldsymbol{\beta}}^{-1} = \Gamma_{\boldsymbol{\beta}}^{-1}\big[\Sigma_{\boldsymbol{\beta}} + q(1-q)\boldsymbol{L}\boldsymbol{L}^T E(\breve{u}_1)^2\big]\Gamma_{\boldsymbol{\beta}}^{-1}$. Note that the first two columns of $\mathcal{P}$ are all zeros. It follows that $\mathcal{P}\Gamma_{\boldsymbol{\beta}}^{-1}\boldsymbol{L} = \mathcal{P}\big(\frac{1}{q}, -\frac{1}{1-q}, 0, \ldots, 0\big)^T = \boldsymbol{0}$. Hence, $\mathcal{P}\boldsymbol{M}\mathcal{P}^T = \mathcal{P}\Gamma_{\boldsymbol{\beta}}^{-1}\Sigma_{\boldsymbol{\beta}}\Gamma_{\boldsymbol{\beta}}^{-1}\mathcal{P}^T$. It follows that

$$\sqrt{n}\mathcal{P}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{D} N\big(0, \mathcal{P}\boldsymbol{M}\mathcal{P}^T\big).$$

Hence

$$\big(\mathcal{P}\widehat{\boldsymbol{M}}\mathcal{P}^T\big)^{-1/2}\sqrt{n}\mathcal{P}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{D} N\big(0, \boldsymbol{I}_{m\times m}\big).$$

Under the null hypothesis $H_0' : \mathcal{P}\boldsymbol{\beta} = \boldsymbol{\xi}_0$,

$$\big(\mathcal{P}\widehat{\boldsymbol{M}}\mathcal{P}^T\big)^{-1/2}\sqrt{n}(\mathcal{P}\widehat{\boldsymbol{\beta}} - \boldsymbol{\xi}_0) \xrightarrow{D} N\big(0, \boldsymbol{I}_{m\times m}\big).$$

Hence,

$$\mathcal{T}_\beta = n(\mathcal{P}\widehat{\boldsymbol{\beta}} - \boldsymbol{\xi}_0)^T\big(\mathcal{P}\widehat{\boldsymbol{M}}\mathcal{P}^T\big)^{-1}(\mathcal{P}\widehat{\boldsymbol{\beta}} - \boldsymbol{\xi}_0)^T \xrightarrow{D} \chi^2_{(m)}.$$

Under the local alternative $H_A' : \mathcal{P}\boldsymbol{\beta} = \boldsymbol{\xi}_0 + \boldsymbol{\eta}/\sqrt{n}$,

$$\big(\mathcal{P}\widehat{\boldsymbol{M}}\mathcal{P}^T\big)^{-1/2}\sqrt{n}(\mathcal{P}\widehat{\boldsymbol{\beta}} - \boldsymbol{\xi}_0) \xrightarrow{D} N\big(\boldsymbol{\delta}, \boldsymbol{I}_{m\times m}\big),$$

with $\boldsymbol{\delta} = \big(\mathcal{P}\boldsymbol{M}\mathcal{P}^T\big)^{-1/2}\boldsymbol{\eta}$. Thus,

$$\mathcal{T}_\beta = n(\mathcal{P}\widehat{\boldsymbol{\beta}} - \boldsymbol{\xi}_0)^T\big(\mathcal{P}\widehat{\boldsymbol{M}}\mathcal{P}^T\big)^{-1}(\mathcal{P}\widehat{\boldsymbol{\beta}} - \boldsymbol{\xi}_0)^T \xrightarrow{D} \chi^2_{(m)}(\lambda),$$

with $\lambda = \boldsymbol{\delta}^T\boldsymbol{\delta} = \boldsymbol{\eta}^T\big(\mathcal{P}\boldsymbol{M}\mathcal{P}^T\big)^{-1}\boldsymbol{\eta}$. The proof is complete. □

Bibliography

Gill, R. D. (1980). *Censoring and stochastic integrals,* Mathematical Centre Tracts, Volume 124. Mathematisch Centrum, Amsterdam.

Hall, P. and C. C. Heyde (1980). *Martingale Limit Theory and its Application.* Academic Press.

Lenglart, É. (1977). Relation de domination entre deux processus. In *Annales de l'IHP Probabilités et Statistiques*, Volume 13, pp. 171–179.

Shen, Y., J. Ning, and J. Qin (2009). Analyzing length-biased data with semiparametric transformation and accelerated failure time models. *Journal of the American Statistical Association 104*(487), 1192–1202.

Stute, W. (1993). Consistent estimation under random censorship when covariables are present. *Journal of Multivariate Analysis 45*(1), 89–103.