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

Hongjian Zhu\textsuperscript{1}, Li-Xin Zhang\textsuperscript{2*}, Jing Ning\textsuperscript{3} and Lu Wang\textsuperscript{4}

\textsuperscript{1}AbbVie Inc.

\textsuperscript{2*}Zhejiang Gongshang University and Zhejiang University

\textsuperscript{3}University of Texas MD Anderson Cancer Center

\textsuperscript{4}University of Texas Health Science Center at Houston

Supplementary Material

S1 Proof

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

\[
\log T_i^{(1)} = \mu_1 + \beta_1 X_{i,1} + \cdots + \beta_{p1} X_{i,p1} + \gamma_1 Z_{i,1} + \cdots + \gamma_{p2} Z_{i,p2} + \varepsilon_i,
\]

\[
\log T_i^{(0)} = \mu_2 + \beta_1 X_{i,1} + \cdots + \beta_{p1} X_{i,p1} + \gamma_1 Z_{i,1} + \cdots + \gamma_{p2} Z_{i,p2} + \varepsilon_i.
\]

Let \( \delta_i^{(j)} \), \( Y_i^{(j)} \) be the values of \( \delta_i \), \( Y_i \) when \( I_i = j \), \( j = 1, 0 \). Define

\[
\tilde{E}h_i = qEh_i^{(1)} + (1 - q)Eh_i^{(0)}
\]
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 
\[
E|h_i| \leq \tilde{E}|h_i|/(q(1-q)).
\]
For a vector \( \mathbf{v} \), we denote \( \mathbf{v}^{\odot 2} = \mathbf{v}\mathbf{v}^T \).

To obtain the asymptotic properties, we let 
\[
\pi(s) = G(s)\tilde{P}(T_1 \geq s), \quad B_1(s) = \tilde{E}[I\{T_1 \geq s\}X_1(Z_1^T\gamma + \varepsilon_1)], \quad \text{and} \quad B_2(s) = \tilde{E}[I\{T_1 \geq s\}X_1^{\odot 2}].
\]
Moreover, let \( \Lambda_c(u) \) be the cumulative hazard function of the censoring times \( C \). We need the following regularity conditions:

(Ra) \( \epsilon_i \) and \( C_i \) have continuous distributions with \( \sup\{t : P(T_i^{(j)} > t) > 0\} \geq \sup\{t : P(C > t) > 0\} \hat{=} \tau_G, j = 1, 0; \)

(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 \( \epsilon_i, t = 1, \ldots, p_1, s = 1, \ldots, p_2; \)

(Rc) \( \det\left( \int_0^\infty B_2(u)d\Lambda_G(u) \right) < \infty, \det\left( \int_0^\infty B_1^{\odot 2}(u)/\pi(u)d\Lambda_G(u) \right) < \infty. \)

Suppose that the covariance matrix \( Var(X_1) \) is nonsingular. Write 
\[
u_i = \gamma_1Z_{i,1} + \ldots + \gamma_{p_2}Z_{i,p_2} + \epsilon_i, \quad \bar{\nu}_i = E[u_i|W_i] - E[u_i].
\]
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) \( Cov(X_{i,k}, u_i) =, k = 1, \ldots, p_1; \)
\( (B) \sum_{i=1}^{n} (I_i - q) \tilde{u}_i = o_P(\sqrt{n}); \)

\( (C) \) 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 \( H_0 : \mu_1 - \mu_2 = 0 \),

\[ \mathcal{T}(n) \overset{D}{\to} N(0, \tau^2), \quad \text{with} \quad \tau^2 = \frac{\sigma_{\delta,G}^2}{\sigma_{z,G}^2} \]  

(S1.1)

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

(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

\[ \mathcal{T}(n) \overset{D}{\to} N(\Delta, \tau^2), \quad \text{with} \quad \Delta = \frac{\delta \sqrt{q(1-q)}}{\sigma_{z,G}}. \]  

(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 \( X \). Then Assumptions (A) and (B) are satisfied. In this case, \( E[\hat{u}_i^2] = \sum_{j=1}^{p_2} \gamma_j^2 Var(E[Z_{ij}|\tilde{Z}_{ij}]). \)
(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 $r_j$s, and that $Z$ is independent of $X$. 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 $\tilde{u}_i = E[u_i - E[u_i]|W_i] = g(W_i)$ is a function of $W_i$. If the order $o(n)$ in Assumption (C) is strengthened to $o(\sqrt{n})$, then

$$\sum_{i=1}^n (I_i - q)\tilde{u}_i = \sum_{i=1}^n (I_i - q)g(W_i)$$

$$= \sum_{t_1, \ldots, t_{p_1}, r_1, \ldots, r_{p_2}} (I_i - q)I\{W_i = (x_{t_1}^{t_1}, \ldots, x_{t_{p_1}}^{t_{p_1}}, z_r^{r_1}, \ldots, z_{r_{p_2}}^{r_{p_2}})\}$$

$$\cdot g(x_{t_1}^{t_1}, \ldots, x_{t_{p_1}}^{t_{p_1}}, z_r^{r_1}, \ldots, z_{r_{p_2}}^{r_{p_2}})$$

$$= \sum_{t_1, \ldots, t_{p_1}, r_1, \ldots, r_{p_2}} D_n^{(q)}(t_1, \ldots, t_{p_1}, r_1, \ldots, r_{p_2})g(x_{t_1}^{t_1}, \ldots, x_{t_{p_1}}^{t_{p_1}}, z_r^{r_1}, \ldots, z_{r_{p_2}}^{r_{p_2}})$$

$$= \sum_{t_1, \ldots, t_{p_1}, r_1, \ldots, r_{p_2}} o_P(\sqrt{n})g(x_{t_1}^{t_1}, \ldots, x_{t_{p_1}}^{t_{p_1}}, z_r^{r_1}, \ldots, z_{r_{p_2}}^{r_{p_2}}) = o_P(\sqrt{n}). \quad (S1.3)$$

Assumption (B) is satisfied.

When $Z$ is independent of $X$, then $\tilde{u}_i = E[u_i - E[u_i]|W_i] = E[u_i - E[u_i]|\tilde{Z}_i] = g(\tilde{Z}_i)$ is a function of $\tilde{Z}_i = (\tilde{Z}_{i,1}, \ldots, \tilde{Z}_{i,p_2})$. Then similar to
\[(S1.3)\]
\[
\sum_{i=1}^{n} (I_i - q) \tilde{u}_i = \sum_{i=1}^{n} (I_i - q) g(\tilde{Z}_i) = o_p(\sqrt{n})
\]
if the within-stratum $q$-imbalances for $Z_1, \ldots, Z_{p_2}$ are of order $o(\sqrt{n})$ in probability.

When $X, Z_1, \ldots, Z_{p_2}$ are independent, $\tilde{u}_i = E[u_i - E[u_i]|W_i] = E[u_i - E[u_i]|\tilde{Z}_i] = \sum_{i=1}^{p_2} \gamma_j E[Z_{i,j} - E[Z_{i,j}]|\tilde{Z}_{i,j}]$, 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
\[
\sum_{i=1}^{n} (I_i - q) \tilde{u}_i = \sum_{j=1}^{q} \gamma_j \sum_{i=1}^{n} (I_i - q) g_j(\tilde{Z}_{i,j})
\]
\[
= \sum_{j=1}^{q} \gamma_j \sum_{r_j=1}^{^s_j} D_n^{(q)}(Z, j; r_j) g_j(z_{r_j}^j) = o_p(\sqrt{n}).
\]
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, $E\tilde{u}_i^2 = Var(\sum_{i=1}^{p_2} \gamma_j E[Z_{i,j} - E[Z_{i,j}]|\tilde{Z}_{i,j}]) = \sum_{j=1}^{p_2} \gamma_j^2 Var(E[Z_{i,j} - E[Z_{i,j}]|\tilde{Z}_{i,j}])$. □.

To prove Theorem 1, we need some lemmas.

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

In particular, we have the following results:

\[(1)\]
\[
\frac{1}{n} \sum_{i=1}^{n} I_i \xrightarrow{P} q, \quad \frac{1}{n} \sum_{i=1}^{n} (1 - I_i) \xrightarrow{P} 1 - q.
\]
\[ \begin{align*}
\sum_{i=1}^{n} I_i X_{i,k} & \xrightarrow{P} qEX_k, \\
\frac{1}{n} \sum_{i=1}^{n} (1 - I_i) X_{i,k} & \xrightarrow{P} (1-q)EX_k, \\
\frac{1}{n} \sum_{i=1}^{n} I_i Z_{i,j} & \xrightarrow{P} qEZ_j, \text{ and} \\
\frac{1}{n} \sum_{i=1}^{n} (1 - I_i) Z_{i,j} & \xrightarrow{P} (1-q)EZ_j, \quad k = 1, \ldots, p_1, \quad j = 1, \ldots, p_2.
\end{align*} \]

\[ \begin{align*}
\sum_{i=1}^{n} \frac{1}{n} (I_i - q) g(W_i) & \xrightarrow{P} qEg(W_i). \quad \text{(S1.4)}
\end{align*} \]

**Proof.** Let \( \mathcal{F}_{i-1} = \sigma(I_{i-1}, X_{i-1}, Z_{i-1}, C_{i-1}, T_{i-1}) \) 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

\[ \frac{1}{n} \sum_{i=1}^{n} I_i g(W_i) \xrightarrow{P} qEg(W_i). \]

By assumption (C) of Theorem 1, we have

\[ \begin{align*}
\sum_{i=1}^{n} (I_i - q) g(W_i) \\
= \sum_{t_1, \ldots, t_{p_1}, r_1, \ldots, r_{p_2}} (I_i - q) I\{W_i = (x_{t_1}^{t_{p_1}}, x_{p_{1}}^{t_{p_1}}, z_{r_1}^{r_{p_1}}, \ldots, z_{r_{p_2}}^{r_{p_2}})\} \\
\quad \cdot g(x_{t_1}^{t_{p_1}}, x_{p_{1}}^{t_{p_1}}, z_{r_1}^{r_{p_1}}, \ldots, z_{r_{p_2}}^{r_{p_2}}) \\
= \sum_{t_1, \ldots, t_{p_1}, r_1, \ldots, r_{p_2}} D_n(t_1, \ldots, t_{p_1}, r_1, \ldots, r_{p_2}) g(x_{t_1}^{t_{p_1}}, x_{p_{1}}^{t_{p_1}}, z_{r_1}^{r_{p_1}}, \ldots, z_{r_{p_2}}^{r_{p_2}}) \\
= \sum_{t_1, \ldots, t_{p_1}, r_1, \ldots, r_{p_2}} o_p(n) g(x_{t_1}^{t_{p_1}}, x_{p_{1}}^{t_{p_1}}, z_{r_1}^{r_{p_1}}, \ldots, z_{r_{p_2}}^{r_{p_2}}) = o_p(n).
\end{align*} \]

On the other hand, by the law of large numbers for i.i.d. random variables,
we have \( \sum_{i=1}^{n} g(W_i)/n \rightarrow^P E_g(W_i) \). It follows that
\[
\frac{1}{n} \sum_{i=1}^{n} I_i g(W_i) = \frac{1}{2n} \sum_{i=1}^{n} (I_i - q)g(W_i) + q \frac{1}{n} \sum_{i=1}^{n} g(W_i)
\]
\[\rightarrow^P qE_g(W_i).\]

The proof of (S1.4) is complete.

Now, note that \( I_i h_i = I_i h_i^{(1)} \). Since we have
\[
\frac{1}{n} \left| \sum_{i=1}^{n} I_i h_i^{(1)} I \{|h_i^{(1)}| \geq M\} \right| \leq \frac{1}{n} \sum_{i=1}^{n} |h_i^{(1)}| I \{|h_i^{(1)}| \geq M\}
\]
\[\rightarrow^P E[|h_i^{(1)}| I \{|h_i^{(1)}| \geq M\}] \rightarrow 0 \quad \text{as } M \rightarrow \infty\]
by the law of large numbers for i.i.d. random variables, we can assume without loss of generality that the \( h_i \)s are bounded. Let \( g_i = E[h_i^{(1)}|W_i] = g(W_i) \). Then
\[
E[I_i(h_i^{(1)} - g_i)|\mathcal{F}_{i-1}, W_i] = E[I_i|\mathcal{F}_{i-1}, W_i] E[h_i^{(1)} - g_i|W_i] = 0.
\]
Thus, \( E[I_i(h_i^{(1)} - g_i)|\mathcal{F}_{i-1}] = 0 \). 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,
\[
\frac{1}{n} \sum_{i=1}^{n} I_i(h_i^{(1)} - g_i) \rightarrow^P 0.
\]
On the other hand, by (S1.4),
\[
\frac{1}{n} \sum_{i=1}^{n} I_i g_i \rightarrow^P qE[g(W_i)] = qE[E[h_i^{(1)}|W_i]] = qEh_i^{(1)}.
\]
It follows that $\frac{1}{n} \sum_{i=1}^{n} I_i h_i \overset{P}{\to} q E h_i^{(1)}$. The proof of $\sum_{i=1}^{n} (1 - I_i) h_i \overset{P}{\to} (1 - q) E h_i^{(0)}$ is similar.

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

$$E\left[\frac{\delta_i^{(j)}}{G(Y_i^{(j)})} | X_i, Z_i, T_i^{(j)}\right] = E\left[\frac{\delta_i^{(j)}}{G(T_i^{(j)})} | X_i, Z_i, T_i^{(j)}\right] = 1,$$

$j = 1, 0$, and

$$E\left[\frac{\delta_i}{G(Y_i)} | X_i, Z_i, I_i, T_i\right] = 1. \quad (S1.5)$$

It follows that if $f(X_i, Z_i, \epsilon_i)$ is a function of $(X_i, Z_i, \epsilon_i)$, then

$$E\left[\frac{\delta_i^{(j)}}{G(Y_i^{(j)})} f(X_i, Z_i, \epsilon_i)\right] = E[f(X_i, Z_i, \epsilon_i)], j = 1, 0.$$

Hence, (1.3) holds. □

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

$$\sqrt{n}\{\hat{G}(t) - G(t)\}/\hat{G}(t) = - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_{0}^{\infty} \frac{I\{s \leq t\} dM_i^c(s)}{\pi(s)} + o_P(1), \text{ uniformly in } t \leq \tau$$

for any $\tau < \tau_G$, where

$$\pi(s) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} I\{Y_i \geq s\} = \tilde{P}(Y_1 \geq s) \quad (S1.7)$$
in probability, and \( M_c^s(t) \) is a continuous-time martingale with predictable variation process

\[
\langle M_c^s, M_c^s \rangle(t) = \int_0^t I\{Y_i \geq u\} \, d\Lambda_c(u).
\]

**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 \( \mathcal{D} = \sigma(X_i, Z_i, I_i; i = 1, 2, \ldots) \), 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 \( \Lambda_G \). We consider \( C_i \) to be the survival time and \( T_i \) the censoring time. It follows that for given \( \mathcal{D} \), \( M(s) = \sum_{i=1}^n M_c^s(s) \) \( (s \geq 0) \) is a martingale with predictable variation process

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

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

\[
\frac{G(t) - \hat{G}(t)}{G(t)} = \frac{\hat{F}_G(t) - F_G(t)}{1 - F_G(t)} = \int_0^t \frac{\hat{G}(u-)}{G(u)} \frac{J(u)}{Y(u)} \, dM(u) - I\{\zeta < t\} \frac{\hat{G}(\zeta)(G(\zeta) - G(t))}{G(t)G(\zeta)},
\]

(S1.8)

where \( \zeta = \zeta_n = \inf\{t > 0 : Y(t) = 0\} \) is a stopping time and \( J(s) = \)
For the second term of (S1.8), note that
\[
P(\zeta < \tau | D) = P(Y_i < \tau, i = 1, \ldots, n| D) = \prod_{i=1}^{n} P(Y_i < \tau | D) \\
\leq \left( 1 - \min_{j=0,1} P(T_1^{(j)} \geq \tau) G(\tau) \right)^n = (1 - a)^n.
\]
Hence, \( P(I \{\xi_n < \tau\} \neq 0, \text{i.o.}) = 0 \) since \( \sum_n (1 - a)^n < \infty \). Given \( D \), the first term of (S1.8), denoted by \( \tilde{M}(t) \), is a martingale with predictable variation process
\[
\langle \tilde{M}, \tilde{M} \rangle(t) = \int_0^t \frac{\hat{G}^2(u-)}{G^2(u)} \frac{J(u)}{Y(u)} d\Lambda_G(u) \leq \frac{\Lambda_G(\tau) J(\tau)}{G^2(\tau) Y(\tau)}.
\]
Note that \( Y(\tau)/n \xrightarrow{P} \pi(\tau) = \hat{P}(T_1 \geq \tau) G(\tau) > 0 \), which implies that \( \langle \tilde{M}, \tilde{M} \rangle(\tau) \xrightarrow{P} 0 \). By Lenglart’s inequality ([Lenglart, 1977]), for any \( \delta > 0 \) and \( \epsilon > 0 \),
\[
P \left( \sup_{t \leq \tau} |\tilde{M}(t)| > \epsilon | D \right) \leq \frac{\eta}{\epsilon^2} + P \left( \langle \tilde{M}, \tilde{M} \rangle(\tau) > \eta | D \right).
\]
Hence,
\[
P \left( \sup_{t \leq \tau} |\tilde{M}(t)| > \epsilon \right) \leq \frac{\eta}{\epsilon^2} + P \left( \langle \tilde{M}, \tilde{M} \rangle(\tau) > \eta \right) \to 0
\]
as \( n \to \infty \) and then \( \eta \to 0 \). That is, \( \sup_{t \leq \tau} |\tilde{M}(t)| \xrightarrow{P} 0 \). It follows that
\[
\sup_{t \leq \tau} \frac{|G(t) - \hat{G}(t)|}{G(t)} \xrightarrow{P} 0.
\]
Now, let
\[
\overline{M}(t) = \int_0^t \frac{1}{\sqrt{n}} \left( \frac{\hat{G}(u-)}{G(u)} \frac{nJ(u)}{Y(u)} - \frac{1}{\pi(u)} \right) dM(u).
\]
Then given $\mathcal{D}$, $\mathcal{M}$ is also a martingale with predictable variation process
$$\langle \mathcal{M}, \mathcal{M} \rangle(t) = \int_0^t \left( \frac{\hat{G}(u)nJ(u)}{G(u)Y(u)} - \frac{1}{\pi(u)} \right)^2 \frac{Y(u)}{n} d\Lambda_G(u).$$

Note that $\frac{\hat{G}(u) - G(u)}{G(u)} \xrightarrow{P} 1$ for almost all $u$ and $\frac{Y(u)}{n} \xrightarrow{P} \pi(u)$, which implies that $\langle \mathcal{M}, \mathcal{M} \rangle(\tau) \xrightarrow{P} 0$. By applying Lenglart’s inequality again we have
$$\sup_{t \leq \tau} |\mathcal{M}(t)| \xrightarrow{P} 0.$$ We conclude that
$$\sqrt{n}(G(t) - \hat{G}(t)) = \frac{1}{\sqrt{n}} \int_0^t \frac{dM(u)}{\pi(u)} + M(t) + o_P(1)$$
$$= \frac{1}{\sqrt{n}} \int_0^t \frac{dM(u)}{\pi(u)} + o_P(1) \text{ uniformly in } t \leq \tau.$$

(S1.6) is proved. □

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

**Proof.** Recall that $u_i = Z_i^T \gamma + \epsilon_i$ and
$$\hat{\beta} = \beta + \left( \sum_{i=1}^n \frac{\delta_iX_i^{\otimes 2}}{G(Y_i)} \right)^{-1} \sum_{i=1}^n \frac{\delta_iX_iu_i}{G(Y_i)}.$$
. Then
$$\hat{\beta} = \beta^* + \left( \sum_{i=1}^n \frac{\delta_iX_i^{\otimes 2}}{G(Y_i)} \right)^{-1} \sum_{i=1}^n \frac{\delta_iX_i(u_i - E[u_i])}{G(Y_i)}.$$
Without loss of generality, we can assume that $E[u_i] = 0$. Otherwise, we can replace $u_i$ by $u_i - Eu_i$, $\mu_1$ and $\mu_2$ by $\mu_1 + Eu_i$ and $\mu_2 + Eu_i$ respectively.
We can write
\[
\sum_{i=1}^{n} \frac{\delta_i X_i u_i}{G(Y_i)} = \sum_{i=1}^{n} \left\{ \frac{1}{\tilde{G}(Y_i)} - \frac{1}{G(Y_i)} \right\} \delta_i X_i u_i + \sum_{i=1}^{n} \frac{\delta_i X_i u_i}{G(Y_i)}.
\]

From the martingale integral representation (S1.6) for \((\hat{G} - G) / \hat{G}\), we have
\[
\frac{1}{n} \sum_{i=1}^{n} \frac{G(Y_i) - \hat{G}(Y_i)}{\hat{G}(Y_i) G(Y_i)} \delta_i X_i u_i
\]
\[
= \frac{1}{n} \sqrt{n} \sum_{i=1}^{n} \left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \int_{0}^{\infty} I\{s \leq Y_i\} dM_c^i(s) / \pi(s) \right\} \delta_i X_i u_i / G(Y_i) + o_P(n^{-1/2}).
\]

Also, by Lemma 1,
\[
\frac{1}{n} \sum_{i=1}^{n} \frac{I\{Y_i \geq s\} \delta_i X_i u_i}{G(Y_i)} \to \tilde{E}\left[ \frac{I\{Y_1 \geq s\} \delta_1 X_1 u_1}{G(Y_1)} \right] = B_1(s), \tag{S1.9}
\]

where \(B_1(s) = \tilde{E}[I\{T_1 \geq s\} X_1 u_1]\). Hence,
\[
\frac{1}{n} \sum_{i=1}^{n} \frac{G(Y_i) - \hat{G}(Y_i)}{G(Y_i) G(Y_i)} \delta_i X_i u_i = \frac{1}{n} \sum_{i=1}^{n} \int_{0}^{\infty} \frac{B_1(s)}{\pi(s)} dM_c^i(s) + o_P(n^{-1/2}).
\]

Note that \(\int_{0}^{\infty} \frac{B_1(s)}{\pi(s)} dM_c^i(s)\) is a function of \(Y_i\) and \(\delta_i\) with mean zero. By Lemma 1
\[
\frac{1}{n} \sum_{i=1}^{n} \int_{0}^{\infty} \frac{B_1(s)}{\pi(s)} dM_c^i(s) \to 0.
\]
By Lemma 1 again,

\[
\frac{1}{n} \sum_{i=1}^{n} \delta_{i} X_{i} u_{i} / G(Y_{i}) = \frac{1}{n} \left[ \sum_{i=1}^{n} \delta_{i} I_{i} u_{i} / G(Y_{i}) - \sum_{i=1}^{n} \delta_{i} (1 - I_{i}) u_{i} / G(Y_{i}) + \sum_{i=1}^{n} \delta_{i} X_{i} u_{i} / G(Y_{i}) \right]
\]

\[
\xrightarrow{P} \begin{bmatrix}
qE[u_1] \\
(1-q)E[u_1] \\
E[X_1 u_1]
\end{bmatrix} = 0
\]

by the assumption that \( Eu_i = 0 \) and \( Cov(X_i, u_i) = 0 \). Combining these we have

\[
\frac{1}{n} \sum_{i=1}^{n} \delta_{i} X_{i} u_{i} G(Y_{i}) \xrightarrow{P} 0.
\]

(S1.10)

Next, we consider

\[
\sum_{i=1}^{n} \frac{\delta_{i} X \otimes^2 i}{G(Y_{i})} = \sum_{i=1}^{n} \left\{ \frac{1}{G(Y_{i})} - \frac{1}{G(Y_{i})} \right\} \delta_{i} X \otimes^2 i + \sum_{i=1}^{n} \frac{\delta_{i} X \otimes^2 i}{G(Y_{i})}.
\]

Similarly, by noting that

\[
\frac{1}{n} \sum_{i=1}^{n} \frac{I\{Y_i \geq s\} \delta_{i} X \otimes^2 i}{G(Y_{i})} \xrightarrow{P} \tilde{E} \left[ \frac{I\{Y_1 \geq s\} \delta_{i} X \otimes^2 i}{G(Y_{i})} \right] = B_2(s),
\]

where \( B_2(s) = \tilde{E} \left[ I\{T_1 \geq s\} X \otimes^2 1 \right] \), we have

\[
\frac{1}{n} \sum_{i=1}^{n} \frac{G(Y_{i}) - \hat{G}(Y_{i})}{G(Y_{i}) G(Y_{i})} \delta_{i} X \otimes^2 i = \frac{1}{n} \sum_{i=1}^{n} \int_{0}^{\infty} \frac{B_2(s)}{\pi(s)} dM_i^c(s) + o_P(1)
\]
and

\[
\frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i X_i \otimes^2}{G(Y_i)} = \frac{1}{n} \frac{\delta_i}{G(Y_i)} \begin{bmatrix}
\text{diag}(I_i, 1-I_i) & (I_i, 1-I_i)^T X_i \\
(I_i, 1-I_i) X_i^T & X_i \otimes^2
\end{bmatrix} \xrightarrow{P} \begin{bmatrix}
\text{diag}(q, 1-q) & (q, 1-q)^T E X_1^T \\
E X_1(q, 1-q) & E X_1^\otimes^2
\end{bmatrix} \doteq \Gamma_\beta. \quad (S1.11)
\]

The inverse matrix of $\Gamma_\beta$ is

\[
\Gamma_\beta^{-1} = \begin{bmatrix}
\Gamma_{11} & \Gamma_{12} \\
\Gamma_{12}^T & (\text{Var}(X_1))^{-1}
\end{bmatrix},
\]

$\Gamma_{11} = \text{diag}(\frac{1}{q}, \frac{1}{1-q})+(1, 1)^T E X_1^T (\text{Var}(X_1))^{-1} E X_1$, $\Gamma_{12} = -(1, 1)^T E X_1^T (\text{Var}(X_1))^{-1}$.

Note that $\int_0^\infty \frac{B_2(s)}{\pi(s)} dM_i^c(s)$ is a function of $Y_i$ and $\delta_i$, with mean zero. By Lemma [1]

\[
\frac{1}{n} \sum_{i=1}^{n} \int_0^\infty \frac{B_2(s)}{\pi(s)} dM_i^c(s) \xrightarrow{P} 0.
\]

By Slutsky’s theorem

\[
\hat{\beta} = \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i X_i \otimes^2}{G(Y_i)} \xrightarrow{P} \Gamma_\beta. \quad (S1.12)
\]

Hence,

\[
\left\{ \sum_{i=1}^{n} \frac{\delta_i X_i \otimes^2}{G(Y_i)} \right\}^{-1} \left( \sum_{i=1}^{n} \frac{\delta_i X_i u_i}{G(Y_i)} \right) \xrightarrow{P} 0.
\]

It follows that

\[
\hat{\beta} - \beta \xrightarrow{P} 0.
\]
The proof of Lemma 3 is complete. □

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

\[ \sqrt{n}(\hat{\beta} - \beta^*) \overset{D}{\to} N\left(0, \Gamma^{-1}_\beta \Sigma^{-1}_\beta \Gamma^{-1}_\beta\right), \]  

(S1.13)

where \( \Gamma_\beta \) and \( \Sigma_\beta \) are defined as in (S1.11) and

\[ \Sigma_\beta = \mathbb{E}\left[ \frac{X^i G(Y_i)}{G(T_i, \tau_G)} \right] - q(1 - q)LL^T E(\hat{u}_i)^2 - \int_0^\infty \frac{B^i_1(u)}{\pi(u)} d\Lambda_c(u), \]  

(S1.14)

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

Proof. Without loss of generality, we assume \( \mathbb{E}[u_i] = 0 \). Note that

\[ \hat{\beta} - \beta = \left\{ \frac{1}{n} \sum_{i=1}^n \frac{\delta_i X^i G(Y_i)}{\hat{G}(Y_i)} \right\}^{-1} \frac{1}{n} \sum_{i=1}^n \frac{\delta_i X^i u_i}{\hat{G}(Y_i)}. \]

We have shown in the proof of Lemma 3 that

\[ \hat{\Gamma}_\beta = \frac{1}{n} \sum_{i=1}^n \frac{\delta_i X^i}{\hat{G}(Y_i)} \overset{p}{\to} \Gamma_\beta. \]  

(S1.15)

Let

\[ A_n = \frac{1}{n} \sum_{i=1}^n \frac{\delta_i X^i u_i}{\hat{G}(Y_i)}, \quad B_n = \left\{ \Gamma^{-1}_\beta - \Gamma^{-1}_\beta \right\} A_n. \]

Then

\[ \sqrt{n}A_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left( \frac{1}{\hat{G}(Y_i)} - \frac{1}{G(Y_i)} \right) \delta_i X^i u_i + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i X^i u_i}{\hat{G}(Y_i)} \]

\[ = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[ \int_0^\infty \frac{B^i_1(s)}{\pi(s)} dM^c(s) + \frac{\delta_i X^i u_i}{\hat{G}(Y_i)} \right] + o_P(1). \]
The first term in the above bracket is a sequence of martingale differences.

However, the second term needs to be modified. Let

\[ V_{i,1} = \int_0^\infty \frac{B_1(s)}{\pi(s)} dM^c_i(s), \quad V_{i,2} = \frac{\delta_i X_i u_i}{G(Y_i)} - (I_i - q) L \dot{u}_i. \]

By assumption (B),

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

It follows that

\[ \sqrt{n} A_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n (V_{i,1} + V_{i,2}) + o_P(1). \]

Since for given \( \mathcal{D} = \sigma(I_j, X_j, Z_j, j = 1, \ldots), T_1, C_1, T_2, C_2, \ldots \) are independent and \( M^c_i(t) \ (t \geq 0) \) is a martingale, it is easy to show that

\[ E[V_{i,1}|\mathcal{F}_{i-1}, X_i, Z_i, I_i] = E[E[V_{i,1}|\mathcal{D}]|\mathcal{F}_{i-1}, X_i, Z_i, I_i] = 0. \quad (S1.16) \]

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

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

\[ = E[I_i|\mathcal{F}_{i-1}, W_i](E[u_i - \bar{u}_i|W_i]) + q \bar{u}_i = q E[u_i|W_i]. \]

Similarly,

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

\[ = (1 - q) \bar{u}_i = (1 - q) E[u_i|W_i]. \]
It follows that

$$E[V_{i,2} | \mathcal{F}_{i-1}] = E[(q, 1 - q, X_i^T u_i)] = (q E u_i, (1 - q) E u_i, E[X_i^T u_i]) = 0$$

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

It can be checked that

$$\tilde{E}[V_{i,1}^{\otimes 2}] = \tilde{E} \left[ \int_{0}^{\infty} \frac{B_1^{\otimes 2}(u)}{\pi(u)} I\{Y_i \geq u\} d\Lambda_{G}(u) \right] \tag{S1.17}$$

$$= \int_{0}^{\infty} \frac{B_1^{\otimes 2}(u)}{\pi(u)} d\Lambda_{G}(u) = \Sigma_{\beta, G},$$

$$\tilde{E}[V_{i,2}^{\otimes 2}] = \tilde{E} \left[ \frac{\delta_i^2 X_i^{\otimes 2} u_i^2}{G^2(Y_i)} \right] - q(1 - q) LL^T E(\tilde{u}_i)^2 \tag{S1.18}$$

$$= \tilde{E} \left[ \frac{X_i^{\otimes 2} u_i^2}{G(T_i \wedge \tau_G)} \right] - q(1 - q) LL^T E(\tilde{u}_i)^2$$

$$= -\Sigma_{\beta, u} - \Sigma_{\beta, z},$$

and, by (S1.16),

$$\tilde{E}[V_{i,1} V_{i,2}^T] = \tilde{E}[V_{i,1} \frac{\delta_i X_i^T u_i}{G(Y_i)}] + \tilde{E} \left[ V_{i,1} \tilde{u}_i (I_i - q) L \right]$$

$$= \tilde{E} \left[ \int_0^\infty \frac{B_1(u)(1 - \delta_i)}{\pi(u)} I\{Y_i \leq u\} d\Lambda_{G}(u) \frac{\delta_i X_i^T u_i}{G(Y_i)} \right]$$

$$- \tilde{E} \left[ \int_0^\infty \frac{B_1(u)}{\pi(u)} I\{Y_i \geq u\} d\Lambda_{G}(u) \frac{\delta_i X_i^T u_i}{G(Y_i)} \right] + 0$$

$$= - \tilde{E} \left[ \int_0^\infty \frac{B_1(u)}{\pi(u)} I\{Y_i \geq u\} \delta_i X_i^T u_i d\Lambda_{G}(u) \right]$$

$$- \tilde{E} \left[ \int_0^\infty \frac{B_1^{\otimes 2}(u)}{\pi(u)} d\Lambda_{G}(u) \right] = -\Sigma_{\beta, G}.$$
Hence, by Lemma 1,
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} (V_{i,1} + V_{i,2})^2 = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} (V_{i,2}^2 + V_{i,1}^2 - 2V_{i,1}V_{i,2}) = \Sigma_{\beta,u} - \Sigma_{\beta,\varepsilon} - \Sigma_{\beta,G} \hat{=} \Sigma_{\beta} \text{ in probability.} \quad \text{ (S1.19)}
\]

Also,
\[
\frac{1}{n} E \left[ \max_{i \leq n} \|V_{i,1} + V_{i,2}\|^2 \right] \leq \frac{1}{nq(1-q)} \tilde{E} \left[ \max_{i \leq n} \|V_{i,1} + V_{i,2}\|^2 \right] \\
\leq \frac{1}{\sqrt{nq(1-q)}} + \frac{1}{q(1-q)} \tilde{E} \left[ \|V_{1,1} + V_{1,2}\|^2 I\{\|V_{1,1} + V_{1,2}\|^2 \geq \sqrt{n}\} \right] \to 0.
\]

By the central limit theorem for martingale differences (see Theorem 3.2 of Hall and Heyde (1980)),
\[
\sqrt{n} A_n \xrightarrow{D} N(0, \Sigma_{\beta}).
\]

It follows that
\[
\sqrt{n} (\hat{\beta} - \beta) = (\hat{\Gamma}_\beta^{-1} - \Gamma_\beta^{-1}) \sqrt{n} A_n + \Gamma_\beta^{-1} \sqrt{n} A_n \\
= \Gamma_\beta^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (V_{i,1} + V_{i,2}) + o_P(1) \\
\xrightarrow{D} N(0, \Gamma_\beta^{-1} \Sigma_{\beta} \Gamma_\beta^{-1}).
\]

The proof of (S1.13) is complete.

Recall that (S1.17), (S1.9), and (S1.7) hold. A consistent estimator of \( \Sigma_{\beta,G} \) is
\[
\hat{\Sigma}_{\beta,G} = \int_{0}^{\infty} \frac{\hat{B}_1^2(s)}{\hat{\pi}(s)} d\hat{\Lambda}_G(s), \quad \text{(S1.20)}
\]
where $\hat{B}_1(s) = \frac{1}{n} \sum_{i=1}^{n} \delta_i I\{Y_i \geq s\} \mathbf{X}_i (\log Y_i - \mathbf{X}_i \hat{\beta}) / \hat{G}(Y_i)$, $\hat{\pi}(s) = \frac{1}{n} \sum_{i=1}^{n} I\{Y_i \geq s\}$, and $\hat{\Lambda}_G(s)$ is the Nelson estimate for the cumulative hazard function $\Lambda_G(s)$ of $C$.

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

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

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

$$\hat{\Sigma}_{\beta} = \hat{\Sigma}_{\beta,WLS} - \hat{\Sigma}_{\beta,G}. \quad (S1.22)$$

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

\[
\sqrt{n} \left[ (\hat{\mu}_1 - \hat{\mu}_2) - (\mu_1 - \mu_2) \right] = L^T \sqrt{n}(\hat{\beta} - \beta)
\]

\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^n L^T \Gamma^{-1}_\beta (V_{i,1} + V_{i,2}) + o_P(1)
\]

\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left\{ \frac{L \Gamma^{-1}_\beta B_1(u_i)}{\pi(u_i)} d\Lambda_G(u_i) + \frac{I_i - q}{q(1-q)} \left( \frac{\delta_i u_i}{G(Y_i)} - \hat{u}_i \right) \right\} + o_P(1).
\]

(S1.23)

\[
\overset{D}{\rightarrow} N \left( 0, L^T \Gamma^{-1}_\beta \Sigma \Gamma^{-1}_\beta L \right).
\]

It can easily be seen that

\[
L^T \Gamma^{-1}_\beta \Sigma \Gamma^{-1}_\beta L
\]

\[
= L^T \Gamma^{-1}_\beta \Sigma_{\beta,a} \Gamma^{-1}_\beta L - L^T \Gamma^{-1}_\beta \Sigma_{\beta,G} \Gamma^{-1}_\beta L
\]

\[
= - \int_0^\infty \frac{(L^T \Gamma^{-1}_\beta B_1(u))^2}{\pi(u)} d\Lambda_G(u)
\]

\[
+ \frac{1}{q^2(1-q)^2} \tilde{E} \left[ \frac{(I_1 - q)^2 \delta_1 u_1^2}{G^2(Y_1)} \right] - \frac{1}{q(1-q)} E(\hat{u}_1)^2.
\]

Since

\[
L^T \Gamma^{-1}_\beta B_1(s) = \tilde{E} \left[ I\{T_1 \geq s\} L^T \Gamma^{-1}_\beta X_1 u_1 \right]
\]

\[
= \tilde{E} \left[ I\{T_1 \geq s\} (I_1 - q) u_1 \right] / (q(1-q))
\]

\[
= E \left[ \left( I\{T_1^{(1)} \geq s\} - I\{T_1^{(0)} \geq s\} \right) u_1 \right]
\]

\[
= E \left[ \left( I\{T_1^{(1)} \geq s\} - I\{T_1^{(1)} e^{\mu_2 - \mu_1} \geq s\} \right) u_1 \right],
\]

we have \(L^T \Gamma^{-1}_\beta B_1(s) = 0\) and \(\tilde{E}[(I_1 - q)^2 \delta_1 u_1^2/G^2(Y_1)] = q(1-q)E[\delta_1 u_1^2/G^2(Y_1)|H_0]\)
under the null hypothesis $\mu_1 = \mu_2$. It follows that

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

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

$$\sqrt{n}\{(\hat{\mu}_1 - \hat{\mu}_2) - (\mu_1 - \mu_2)\}$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{I_i - q - \delta_i u_i}{G(Y_i)} - \hat{\mu}_i) + o_P(1) \quad (S1.24)$$

$$\overset{D}{\rightarrow} N\left(0, \sigma^2_{\delta,G}/(q(1 - q))\right). \quad (S1.25)$$

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

$$n \hat{\text{Var}}(L^T \hat{\beta}) = L^T \hat{\Gamma}_\beta^{-1} \left(\hat{\Sigma}_{\beta,u} - \hat{\Sigma}_{\beta,G}\right) \hat{\Gamma}_\beta^{-1} L,$$

$$L^T \hat{\Gamma}_\beta^{-1} \overset{P}{\rightarrow} L^T \Gamma_\beta^{-1},$$

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

$$L^T \hat{\Gamma}_\beta^{-1} \hat{\Sigma}_{\beta,G} \hat{\Gamma}_\beta^{-1} L$$

$$\overset{P}{\rightarrow} \int_{0}^{\infty} \left(\frac{L^T \Gamma_\beta^{-1} B_1(u) \Sigma_{\beta}}{\pi(u)}\right) d\Lambda_G(u) = 0 \quad \text{under } H_0.$$

It follows that

$$n \hat{\text{Var}}(L^T \hat{\beta}) = n \hat{\text{Var}}_{WLS}(L^T \hat{\beta}) + o_P(1).$$
For $\hat{\Sigma}_{\beta,WLS}$, we have

\[
\hat{\Sigma}_{\beta,WLS} = \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i^2 X_i^{\otimes 2}}{G^2(Y_i)} u_i^2 + \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i^2 X_i^{\otimes 2}}{G^2(Y_i)} \left\{ X_i^T (\hat{\beta} - \beta) \right\}^2 - \frac{2}{n} \sum_{i=1}^{n} \frac{\delta_i^2 X_i^{\otimes 2}}{G^2(Y_i)} u_i X_i^T (\hat{\beta} - \beta).
\]

Note that

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

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

\[
\frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i^2 X_i^{\otimes 2}}{G(Y_i)^2} u_i^2 \xrightarrow{P} \tilde{E} \left[ \frac{\delta_1 X_1^{\otimes 2}}{G(Y_1)} u_1^2 \right] = \tilde{E} \left[ \frac{X_1^{\otimes 2}}{G(T_1 \wedge \tau_G)} u_1^2 \right],
\]

and

\[
\left\| \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i^2 X_i^{\otimes 2}}{G(Y_i)^2} \left\{ X_i^T (\hat{\beta} - \beta) \right\}^2 \right\| \\
\leq \frac{1}{n} \sum_{i=1}^{n} \frac{\delta_i^2 (1 + \|X_i\|^2)^2}{G^2(Y_i)} \| \hat{\beta} - \beta \|^2 = O_P(1) \| \hat{\beta} - \beta \|^2 \xrightarrow{P} 0,
\]

\[
\left\| \frac{2}{n} \sum_{i=1}^{n} \frac{\delta_i^2 X_i^{\otimes 2}}{G(Y_i)} u_i X_i^T (\hat{\beta} - \beta) \right\| = O_P(1) \| \hat{\beta} - \beta \| \xrightarrow{P} 0,
\]

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

\[
\hat{\Sigma}_{\beta,WLS} \xrightarrow{P} \tilde{E} \left[ \frac{\delta_1 X_1^{\otimes 2}}{G^2(Y_1)} u_1^2 \right] = \Sigma_{\beta,u}.
\]
Hence,

\[
n\hat{\text{Var}}(L^T \hat{\beta}) = n\hat{\text{Var}}_{WLS}(L^T \hat{\beta}) + o_P(1) \tag{S1.27}
\]

\[
= L^T(\hat{\Gamma}^{-1} - \Gamma^{-1})\hat{\Sigma}_{\beta,u}(\hat{\Gamma}^{-1} - \Gamma^{-1})L + 2L^T(\hat{\Gamma}^{-1} - \Gamma^{-1})\hat{\Sigma}_{\beta,u}\Gamma^{-1}L + o_P(1)
\]

\[
= \frac{1}{n(q^2(1-q))^2} \sum_{i=1}^{n} \frac{\delta_i(I_i - q)^2}{G^2(Y_i)} (\log Y_i - X_i\hat{\beta})^2 + o_P(1)
\]

\[
= \frac{1}{nq^2(1-q)^2} \sum_{i=1}^{n} \frac{\delta_i(I_i - q)^2}{G^2(Y_i)} (\log Y_i - X_i\beta)^2 + o_P(1)
\]

\[
\overset{p}{\to} \frac{1}{q^2(1-q)^2} \hat{E} \left[ \frac{(I_1 - q)^2 \delta_1 u_1^2}{G^2(Y_1)} \right] = \frac{1}{q(1-q)} \sigma_{z,G}^2 \text{ under } H_0.
\]

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

\[
\frac{L^T \hat{\beta} - (\mu_1 - \mu_2)}{\{\text{Var}(L^T \hat{\beta})\}^{1/2}} \overset{D}{\to} N(0, \tau^2), \quad \tau^2 = \frac{\sigma_{z,G}^2}{\sigma_{z,G}^2} \text{ under } H_0. \tag{S1.28}
\]

The proof of Theorem 1 (i) is complete.

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

\[
\frac{1}{n} \sum_{i=1}^{n} (1 - I_i)|h_i^{(0)} - h(T_i^{(1)}, 0, C_i, X_i, Z_i)|
\]

\[
\leq \frac{1}{n} \sum_{i=1}^{n} \sup_{|x| \leq \epsilon} |h(T_i^{(1)} e^x, 0, C_i, X_i, Z_i) - h(T_i^{(1)}, 0, C_i, X_i, Z_i)|
\]

\[
\overset{p}{\to} E \left[ \sup_{|x| \leq \epsilon} |h(T_1^{(1)} e^x, 0, C_1, X_1, Z_1) - h(T_1^{(1)}, 0, C_1, X_1, Z_1)| \right]
\]

\[
\to 0 \text{ as } n \to \infty \text{ and then } \epsilon \to 0.
\]
It follows that
\[
\frac{1}{n} \sum_{i=1}^{n} (1 - I_i) h_i = \frac{1}{n} \sum_{i=1}^{n} (1 - I_i) h_i^{(0)}
\]
\[
= \frac{1}{n} \sum_{i=1}^{n} (1 - I_i) h(T_i^{(1)}, 0, C_i, X_i, Z_i) + o_P(1)
\]
\[
P(1 - q) Eh(T_1^{(1)}, 0, C_1, X_1, Z_1) = (1 - q) E[h_1^{(0)}|H_0]
\]
and
\[
\frac{1}{n} \sum_{i=1}^{n} I_i h_i = \frac{1}{n} \sum_{i=1}^{n} I_i h_i^{(1)} \xrightarrow{P} q E[h_1^{(1)}|H_0],
\]
\[
\frac{1}{n} \sum_{i=1}^{n} h_i \xrightarrow{P} E[h_1|H_0] + 1
\]
In particular,
\[
\frac{1}{n} \sum_{i=1}^{n} I\{Y_i \geq s\} \delta_i \frac{X_i u_i}{G(Y_i)} \xrightarrow{P} E\left[I\{Y_1 \geq s\} \delta_1 \frac{X_1 u_1}{G(Y_1)} \big| H_0\right] = B_1(s),
\]
with \(L^T \Gamma^{-1}_\beta B_1(s) = E\left[I\{T_1 \geq s\} L^T \Gamma^{-1}_\beta X_1 u_1 \big| H_0\right] = 0\). The arguments used to show (S1.24), (S1.25) and (S1.27) remain valid with \(\sigma^2_{z,G} = E\left[\frac{\delta_1 u_1^2}{G^2(Y_1)} \big| H_0\right]\) and \(\sigma^2_\delta = \sigma^2_{u,G} - E(\tilde{u}_1)^2\). Hence,
\[
\frac{L\hat{\beta} - (\mu_1 - \mu_2)}{\{\text{Var}(L\hat{\beta})\}^{1/2}} \xrightarrow{D} N(0, \tau^2), \quad \tau^2 = \frac{\sigma^2_\delta}{\sigma^2_{z,G}} \quad \text{under } H_A.
\]
Note that \((\mu_1 - \mu_2)/\{\text{Var}(L\hat{\beta})\}^{1/2} \xrightarrow{D} \delta \sqrt{q(1 - q)}/\sigma_{u,G}\). We conclude that
\[
\frac{L\hat{\beta}}{\{\text{Var}(L\hat{\beta})\}^{1/2}} \xrightarrow{D} N(\Delta, \tau^2), \quad \tau^2 = \frac{\sigma^2_\delta}{\sigma^2_{z,G}}, \quad \Delta = \frac{\delta \sqrt{q(1 - q)}}{\sigma_{u,G}} \quad \text{under } H_A.
\]
The proof of Theorem 1 is complete. \(\square\)
Proof of Theorem 2

From (S1.28) in Theorem 1, if $\gamma_j = 0, j = 1, 2, \ldots, p_2$, the terms involving $Z$ will disappear, so $u_i = \epsilon_i, \hat{u}_i = 0$ 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 $E[Z_1^T + \epsilon_1\gamma|W_1] \neq \text{Const}$, we have $\tau < 1$. Then the hypothesis test of (2.4) is conservative. □

Proof of Theorem 5

By Lemma [4]

$$\sqrt{n}P(\hat{\beta} - \beta) \overset{D}{\rightarrow} N\left(0, P\Gamma^{-1}_\beta \Sigma_\beta \Gamma^{-1}_\beta P^T\right).$$

On the other hand,

$$\hat{M} \doteq n\text{Var}(\hat{\beta}) = \hat{\Gamma}^{-1}_\beta \left(\hat{\Sigma}_{\beta,WLS} - \hat{\Sigma}_{\beta,G}\right)\hat{\Gamma}^{-1}_\beta \rightarrow M,$$

by (S1.26), where $M = \Gamma^{-1}_\beta (\Sigma_{\beta,u} - \Sigma_{\beta,G})\Gamma^{-1}_\beta = \Gamma^{-1}_\beta [\Sigma_\beta + q(1-q)LL^T E(\hat{u}_1)^2] \Gamma^{-1}_\beta$.

Note that the first two columns of $P$ are all zeros. It follows that $P\Gamma^{-1}_\beta L = P\left(\frac{1}{q} - \frac{1}{1-q}, 0, \ldots, 0\right)^T = 0$. Hence, $PM^T = P\Gamma^{-1}_\beta \Sigma_\beta \Gamma^{-1}_\beta P^T$. It follows that

$$\sqrt{n}P(\hat{\beta} - \beta) \overset{D}{\rightarrow} N\left(0, PM^P\right).$$

Hence

$$(PM^P)^{-1/2} \sqrt{n}P(\hat{\beta} - \beta) \overset{D}{\rightarrow} N\left(0, I_{m \times m}\right).$$
Under the null hypothesis $H'_0 : \mathcal{P}\beta = \xi_0$,

$$
(\mathcal{P}\mathcal{M}\mathcal{P}^T)^{-1/2} \sqrt{n}(\hat{\beta} - \xi_0) \overset{D}{\to} N(0, I_{m \times m}).
$$

Hence,

$$
T_\beta = n(\mathcal{P}\hat{\beta} - \xi_0)^T(\mathcal{P}\mathcal{M}\mathcal{P}^T)^{-1}(\mathcal{P}\hat{\beta} - \xi_0)^T \overset{D}{\to} \chi^2(m).
$$

Under the local alternative $H'_A : \mathcal{P}\beta = \xi_0 + \eta/\sqrt{n}$,

$$
(\mathcal{P}\mathcal{M}\mathcal{P}^T)^{-1/2} \sqrt{n}(\hat{\beta} - \xi_0) \overset{D}{\to} N(\delta, I_{m \times m}),
$$

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

$$
T_\beta = n(\mathcal{P}\hat{\beta} - \xi_0)^T(\mathcal{P}\mathcal{M}\mathcal{P}^T)^{-1}(\mathcal{P}\hat{\beta} - \xi_0)^T \overset{D}{\to} \chi^2(m)(\lambda),
$$

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

**Bibliography**


