Biometrika (2022), **107**, 1, *p.* 1
Printed in Great Britain

Supplementary Material for "Covariate-Adjusted Log-Rank Test: Guaranteed Efficiency Gain and Universal Applicability"

BY TING YE
Department of Biostatistics, University of Washington, Seattle, Washington 98195, U.S.A.

JUN SHAO
School of Statistics, East China Normal University, Shanghai 200241, China
Department of Statistics, University of Wisconsin, Madison, Wisconsin 53706, U.S.A.

YANYAO YI
Global Statistical Sciences, Eli Lilly and Company, Indianapolis, Indiana 46285, U.S.A.

S1. ADDITIONAL SIMULATIONS

S1.1. Additional simulations with $n = 500$ under Case I-IV

Based on 10,000 simulations, power curves of four tests for θ ranging from 0 to 0.6 under simple randomization and minimization are in Figure S1.

S1.2. Simulations with $n = 200$ under Case I-IV

The simulation setting is the same as the simulations in the main article, except that $n = 200$ and Z is the 2-dimensional vector whose first component is a two-level discretized first component of W and second component is a two-level discretized second component of W. Thus, the average sample size in each treatment and Z-level combination is $200/(2 \times 4) = 25$. Type I error rates for four tests under four cases and three randomization schemes are shown in Table S1. Power curves under three randomization schemes are in Figure S2. From Table S1, we see that with a smaller sample size, the type I error rates of the two covariate-adjusted tests $\mathcal{T}_{\rm CL}$ and $\mathcal{T}_{\rm CSL}$ can be slightly inflated but the inflation is not too severe. Otherwise, the results with $n = 200$ are similar to the results with $n = 500$.

S1.3. Simulations under violations of Assumption C

This simulation setting is the same as Case III, except that C_j follows a Cox model with conditional hazard $\log(1.1)\exp(-\psi j + \eta_C^\top W)$ for $j = 0, 1$, ψ ranges from 0 to 1 for different extent of assumption violation, and $\eta_C = (0.2, 0.2, 0.2)^\top$. Type I error rates for four tests under three randomization schemes are shown in Table S2. It can be seen that type I error is inflated as ψ becomes larger.

tingye1@uw.edu
shao@stat.wisc.edu
yi_yanyao@lilly.com

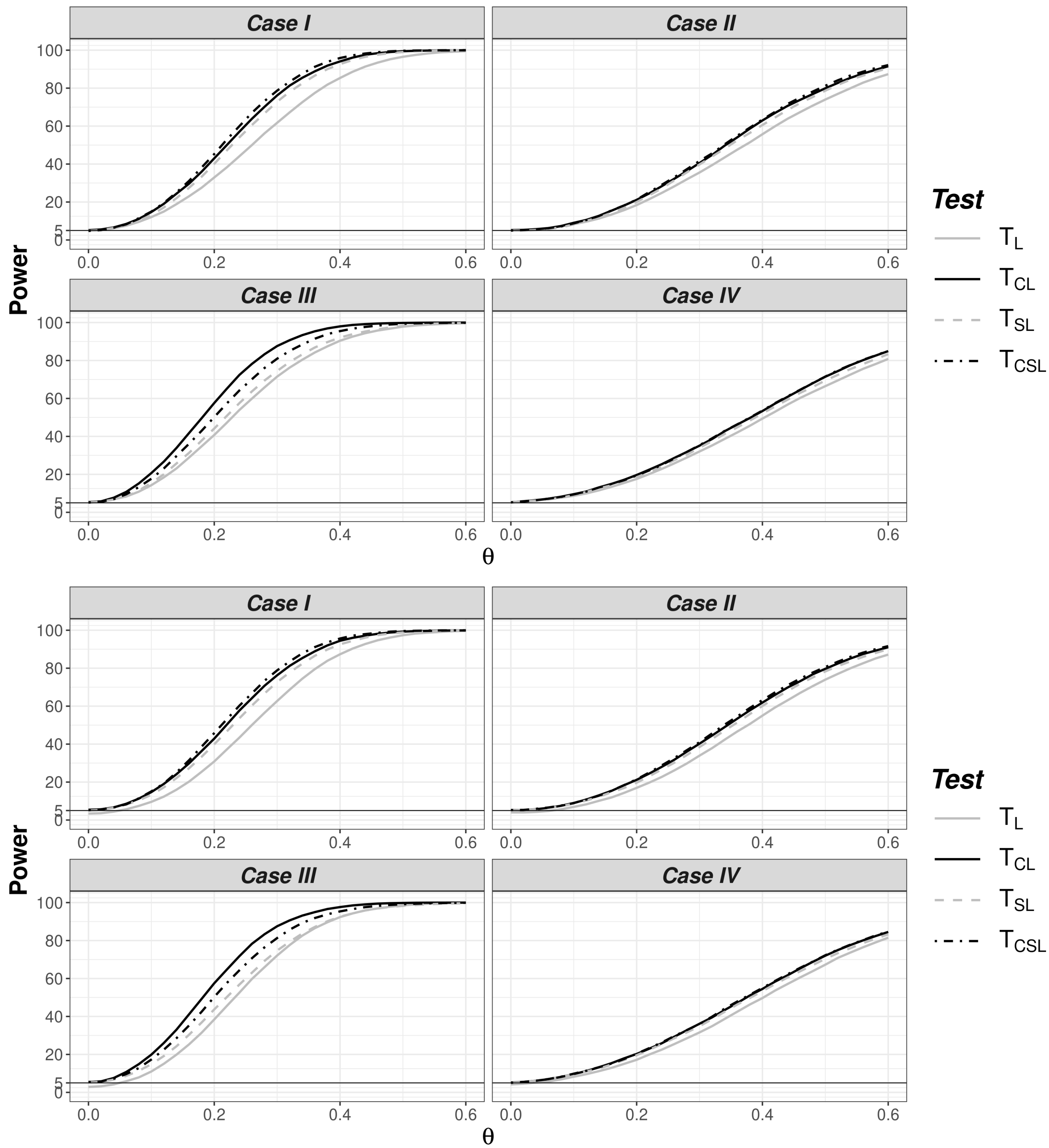
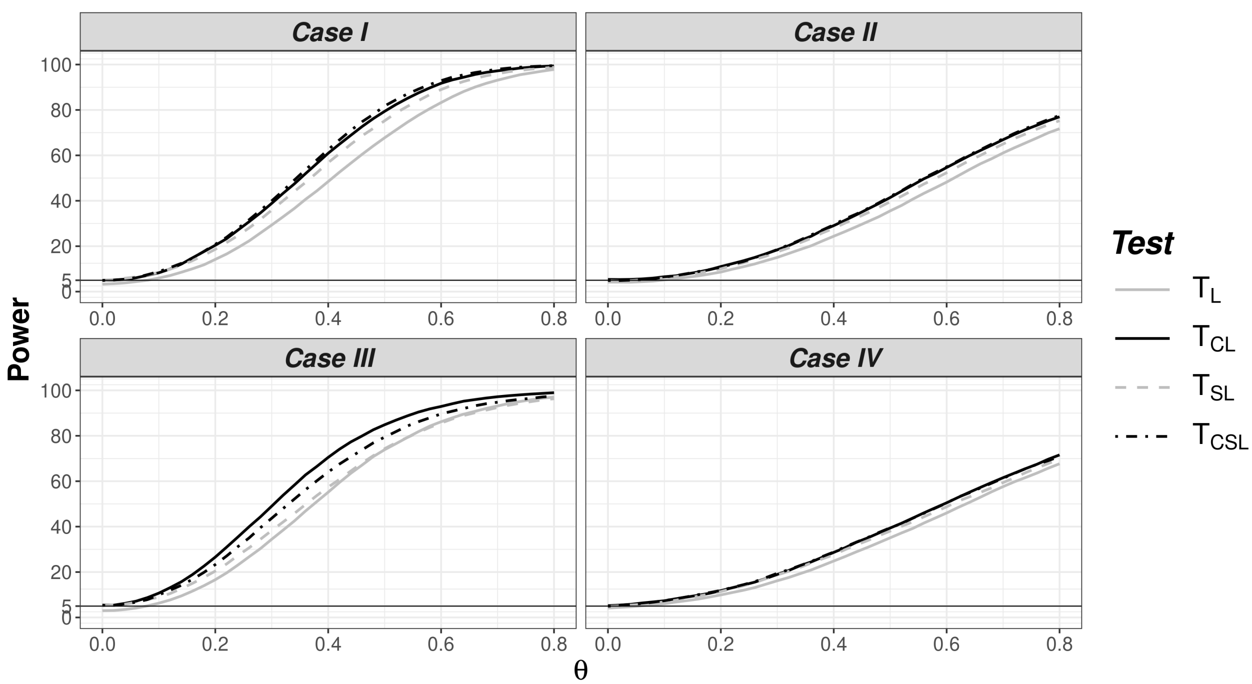


Fig. S1: Power curves based on 10,000 simulations with $n = 500$ under simple randomization (top) and minimization (bottom).

Table S1: Type I errors (in %) based on 10,000 simulations with $n = 200$

Case	Randomization	$\mathcal{T}_{\mathrm{L}}$	$\mathcal{T}_{\mathrm{CL}}$	$\mathcal{T}_{\mathrm{SL}}$	$\mathcal{T}_{\mathrm{CSL}}$
Case I	simple	4.90	5.34	4.94	5.01
	permuted block	3.29	5.00	5.07	4.92
	minimization	3.36	5.03	5.01	5.24
Case II	simple	5.05	5.59	5.16	5.28
	permuted block	4.19	5.37	4.86	4.92
	minimization	4.04	5.42	4.92	5.27
Case III	simple	4.88	5.53	5.12	5.18
	permuted block	2.98	5.44	5.28	5.32
	minimization	3.28	5.60	5.46	5.52
Case IV	simple	4.85	5.36	5.19	5.39
	permuted block	4.16	5.19	4.87	4.96
	minimization	4.64	5.64	5.35	5.38



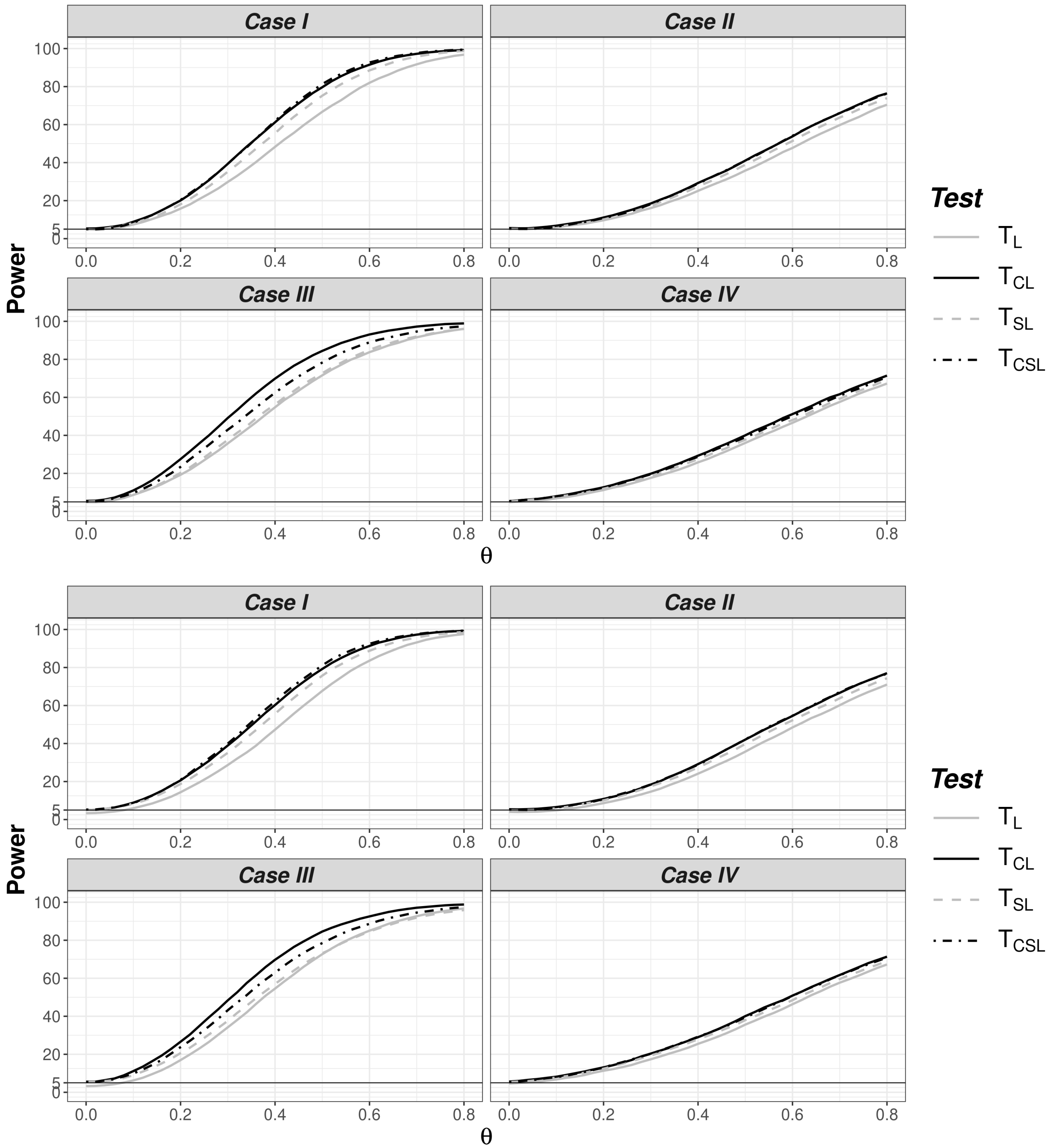


Fig. S2: Power curves based on 10,000 simulations with $n = 200$ under permuted block randomization (previous page), simple randomization (top this page), and minimization (bottom this page).

Table S2: Type I errors (in %) based on 10,000 simulations with $n = 500$ when Assumption C is violated.

Randomization	ψ	$\mathcal{T}_{\mathrm{L}}$	$\mathcal{T}_{\mathrm{CL}}$	$\mathcal{T}_{\mathrm{SL}}$	$\mathcal{T}_{\mathrm{CSL}}$
simple	0.0	4.89	5.14	5.12	4.80
	0.1	4.84	5.21	5.22	4.93
	0.2	5.00	5.23	5.15	5.12
	0.3	5.27	5.59	5.29	5.03
	0.4	5.35	5.96	5.41	5.30
	0.5	5.84	6.51	5.34	5.40
	0.6	6.34	7.25	5.35	5.75
	0.7	6.79	7.99	5.76	5.82
	0.8	7.85	8.95	5.94	5.93
	0.9	8.63	10.07	6.12	6.43
	1.0	9.70	11.60	6.52	6.94
permuted block	0.0	3.41	5.36	5.40	4.86
	0.1	3.38	5.49	5.43	5.04
	0.2	3.37	5.36	5.36	4.97
	0.3	3.63	5.79	5.41	5.20
	0.4	3.87	6.04	5.44	5.20
	0.5	4.27	6.76	5.60	5.30
	0.6	4.62	7.08	5.69	5.43
	0.7	5.28	7.68	5.84	5.68
	0.8	5.78	8.84	6.06	5.95
	0.9	6.53	9.90	6.35	6.39
	1.0	7.52	11.34	6.71	6.81
minimization	0.0	3.13	5.26	5.15	5.10
	0.1	3.11	5.27	5.10	5.08
	0.2	3.28	5.54	5.26	5.14
	0.3	3.42	5.72	5.18	4.97
	0.4	3.63	6.19	5.08	5.10
	0.5	4.10	6.87	5.17	5.27
	0.6	4.45	7.46	5.55	5.38
	0.7	5.14	8.10	5.55	5.65
	0.8	5.74	9.02	5.69	6.06
	0.9	6.48	10.15	5.92	6.25
	1.0	7.39	11.15	6.50	6.57

S2. LEMMAS AND ADDITIONAL THEORETICAL RESULTS

S2.1. Asymptotic optimality

Consider a general class of log-rank score functions

$$\widehat{U}_{\mathrm{CL}}(b_0, b_1) = \frac{1}{n}\sum_{i=1}^{n}\Big[I_i\{O_{i1} - (X_i - \bar{X})^\top b_1\} - (1 - I_i)\{O_{i0} - (X_i - \bar{X})^\top b_0\}\Big] + n^{-1/2}o_p(1),$$

where b_1 and b_0 are any fixed constants. The following theorem derives the asymptotic distribution of $\widehat{U}_{\mathrm{CL}}(b_0, b_1)$, and shows that $\widehat{U}_{\mathrm{CL}}(\beta_0, \beta_1)$ has the smallest variance.

THEOREM S1. *Assume (D), (D†), and that all levels of Z used in covariate-adaptive randomization are included in X_i as a sub-vector. Then the following results hold.*

$$\sqrt{n}\left(\widehat{U}_{\mathrm{CL}}(b_0, b_1) - \frac{n_1\theta_1 - n_0\theta_0}{n}\right) \xrightarrow{d} N\left(0, \sigma^2_{\mathrm{CL}}(b_0, b_1)\right),$$

where

$$\begin{aligned}\sigma^2_{\mathrm{CL}}(b_0, b_1) =& \pi E\left\{\mathrm{var}(O_{i1} - X_i^\top b_1 \mid Z_i)\right\} + (1-\pi)E\left\{\mathrm{var}(O_{i0} - X_i^\top b_0 \mid Z_i)\right\}\\ &+ \nu \mathrm{var}\left\{E(O_{i1} - X_i^\top b_1 \mid Z_i) + E(O_{i0} - X_i^\top b_0 \mid Z_i)\right\}\\ &+ \mathrm{var}\left\{\pi E(O_{i1} - X_i^\top b_1 \mid Z_i) - (1-\pi)E(O_{i0} - X_i^\top b_0 \mid Z_i)\right\}\\ &+ \{2\pi\beta_1 - 2(1-\pi)\beta_0 - \pi b_1 + (1-\pi)b_0\}^\top \mathrm{var}(X_i)\{\pi b_1 - (1-\pi)b_0\}.\end{aligned}$$

Furthermore,

$$\begin{aligned}&\sigma^2_{\mathrm{CL}}(b_0, b_1) - \sigma^2_{\mathrm{CL}}(\beta_0, \beta_1)\\ &= \pi(1-\pi)(\beta_1 - b_1 + \beta_0 - b_0)^\top\left[E\{\mathrm{var}(X_i \mid Z_i)\} + \frac{\nu}{\pi(1-\pi)}\mathrm{var}\{E(X_i \mid Z_i)\}\right](\beta_1 - b_1 + \beta_0 - b_0)\end{aligned}$$

which is ≥ 0 with strict inequality holding unless $\beta_1 + \beta_0 = b_0 + b_1$ or $\mathrm{var}(X_i \mid Z_i) = 0$ and $\nu = 0$ almost surely.

To end this section we emphasize that the use of $\widehat{O}_{ij}$'s in (3) as derived outcomes to reduce variability of $\widehat{U}_{\mathrm{L}}$ is a key to our results. The use of other derived outcomes may not achieve guaranteed efficiency gain. For example, Tangen & Koch (1999) and Jiang et al. (2008) considered log-rank scores $\widetilde{O}_{ij} = \int_0^\tau \frac{1}{2}\{dN_{ij}(t) - Y_{ij}(t)d\bar{N}(t)/\bar{Y}(t)\}$ as derived outcomes; however, using $\widetilde{O}_{ij}$ to replace $\widehat{O}_{ij}$ in (4) and (5) produces an adjusted score that is not necessarily more efficient than the unadjusted $\widehat{U}_{\mathrm{L}}$ and is always less efficient than $\widehat{U}_{\mathrm{CL}}$ in (4) (as shown in Theorem S1), due to the reason that using $\widetilde{O}_{ij}$ instead of $\widehat{O}_{ij}$ in (5) does not correctly capture the true correlation between O_{ij} and X_i. Furthermore, using $\widetilde{O}_{ij}$ may not produce a valid test under covariate-adaptive randomization, even with (C)-(D) and all joint levels of Z_i included in X_i.

S2.2. Covariate adjustment in hazard ratio estimation under Cox model

After testing the null hypothesis of no treatment effect using the covariate-adjusted log-rank test $\mathcal{T}_{\mathrm{CL}}$, it is of interest to also report an effect size estimate and confidence interval. One common parameter is the hazard ratio e^θ under the Cox proportional hazards model

$$\lambda_1(t) = \lambda_0(t)e^\theta.$$

Without using covariates, the score equation from the partial likelihood is

$$\widehat{U}_{\mathrm{L}}(\vartheta) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau}\left\{I_i - \frac{S^{(1)}(\vartheta,t)}{S^{(0)}(\vartheta,t)}\right\} dN_i(t),$$

where $S^{(1)}(\vartheta,t) = n^{-1}\sum_{i=1}^{n} I_i Y_i(t) e^{\vartheta I_i} = e^{\vartheta}\bar{Y}_1(t)$ and $S^{(0)}(\vartheta,t) = n^{-1}\sum_{i=1}^{n} Y_i(t) e^{\vartheta I_i} = e^{\vartheta}\bar{Y}_1(t) + \bar{Y}_0(t)$. The log-rank test uses $\widehat{U}_{\mathrm{L}}(0)$.

Let

$$O_{i1}(\theta) = \int_0^{\tau}\left\{1 - \frac{s^{(1)}(\theta,t)}{s^{(0)}(\theta,t)}\right\}\left\{dN_{i1}(t) - Y_{i1}(t)e^{\theta}\frac{E(dN_i(t))}{s^{(0)}(\theta,t)}\right\}$$

$$O_{i0}(\theta) = \int_0^{\tau}\frac{s^{(1)}(\theta,t)}{s^{(0)}(\theta,t)}\left\{dN_{i0}(t) - Y_{i0}(t)\frac{E(dN_i(t))}{s^{(0)}(\theta,t)}\right\},$$

where $s^{(1)}(\theta,t) = e^{\theta}\mu(t)E\{Y_i(t)\}$ and $s^{(0)}(\theta,t) = \{e^{\theta}\mu(t) + 1 - \mu(t)\}E\{Y_i(t)\}$.

THEOREM S2. *Assume (C) and (D), and all joint levels of Z_i used in covariate-adaptive randomization are included in X_i as a sub-vector. Also assume the Cox proportional hazards model $\lambda_1(t) = \lambda_0(t)e^{\theta}$. Then, the following results hold regardless of which covariate-adaptive randomization scheme is applied.*

$$\sqrt{n}(\widehat{\theta}_{\mathrm{CL}} - \theta) \xrightarrow{d} N\left(0, \frac{\sigma^2_{\mathrm{CL}}(\theta)}{\sigma^4_{\mathrm{L}}(\theta)}\right),$$

where $\sigma^2_{\mathrm{CL}}(\theta) = \sigma^2_{\mathrm{L}}(\theta) - \pi(1-\pi)(\beta_1(\theta) + \beta_0(\theta))^{\top}\Sigma_X(\beta_1(\theta) + \beta_0(\theta))$,

$$\begin{aligned}\sigma^2_{\mathrm{CL}}(\theta) &= \pi \mathrm{var}(O_{i1}(\theta) - X_i^{\top}\beta_1(\theta)) + (1-\pi)\mathrm{var}(O_{i0}(\theta) - X_i^{\top}\beta_0(\theta)) \\ &\quad + (\pi\beta_1(\theta) - (1-\pi)\beta_0(\theta))^{\top}\Sigma_X(\pi\beta_1(\theta) - (1-\pi)\beta_0(\theta)), \\ \sigma^2_{\mathrm{L}}(\theta) &= \pi \mathrm{var}(O_{i1}(\theta)) + (1-\pi)\mathrm{var}(O_{i0}(\theta)),\end{aligned}$$

and $\beta_j(\theta) = \Sigma_X^{-1}\mathrm{cov}(X_i, O_{ij}(\theta))$ *for* $j = 0, 1$.

Inference can be made based on Theorem S2 and estimated variance $\widehat{\sigma}^2_{\mathrm{CL}}(\widehat{\theta}_{\mathrm{CL}})/\widehat{\sigma}^4_{\mathrm{L}}(\widehat{\theta}_{\mathrm{CL}})$, with

$$\widehat{\sigma}^2_{\mathrm{L}}(\widehat{\theta}_{\mathrm{CL}}) = -\partial\widehat{U}_{\mathrm{L}}(\vartheta)/\partial\vartheta \mid_{\vartheta=\widehat{\theta}_{\mathrm{CL}}} = n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\frac{e^{\widehat{\theta}_{\mathrm{CL}}}\bar{Y}_1(t)\bar{Y}_0(t)}{\{e^{\widehat{\theta}_{\mathrm{CL}}}\bar{Y}_1(t) + \bar{Y}_0(t)\}^2}dN_i(t)$$

and $\widehat{\sigma}^2_{\mathrm{CL}}(\widehat{\theta}_{\mathrm{CL}}) = \widehat{\sigma}^2_{\mathrm{L}}(\widehat{\theta}_{\mathrm{CL}}) - \pi(1-\pi)(\widehat{\beta}_1(\widehat{\theta}_{\mathrm{L}})) + \widehat{\beta}_0(\widehat{\theta}_{\mathrm{L}}))^{\top}\widehat{\Sigma}_X(\widehat{\beta}_1(\widehat{\theta}_{\mathrm{L}})) + \widehat{\beta}_0(\widehat{\theta}_{\mathrm{L}}))$.

S2.3. Covariate adjustment in hazard ratio estimation under stratified Cox model

In this section, we assume the stratified Cox proportional hazards model:

$$\lambda_{z1}(t) = \lambda_{z0}(t)e^{\theta}.$$

Without using covariates, the score equation from the partial likelihood is

$$\widehat{U}_{\mathrm{SL}}(\vartheta) = n^{-1}\sum_{z}\sum_{i:Z_i=z}\int_0^{\tau}\left\{I_i - \frac{S_z^{(1)}(\vartheta,t)}{S_z^{(0)}(\vartheta,t)}\right\}dN_i(t).$$

where $S_z^{(1)}(\vartheta,t) = n^{-1}\sum_{i:Z_i=z} I_i Y_i(t)e^{\vartheta I_i} = e^{\vartheta}\bar{Y}_{z1}(t)$ and $S_z^{(0)}(\vartheta,t) = n^{-1}\sum_{i:Z_i=z} Y_i(t)e^{\vartheta I_i} = e^{\vartheta}\bar{Y}_{z1}(t) + \bar{Y}_{z0}(t)$. The log-rank test uses $\widehat{U}_{\mathrm{SL}}(0)$. The maximum partial likelihood estimator $\widehat{\theta}_{\mathrm{SL}}$ of θ is a solution to $\widehat{U}_{\mathrm{SL}}(\vartheta) = 0$. Our covariate-adjusted score is

$$\widehat{U}_{\mathrm{CSL}}(\vartheta) = \widehat{U}_{\mathrm{SL}}(\vartheta) - \frac{1}{n}\sum_{z}\sum_{i:Z_i=z}\left\{I_i(X_i - \bar{X}_z)^\top\widehat{\gamma}_1(\widehat{\theta}_{\mathrm{SL}}) - (1-I_i)(X_i - \bar{X}_z)^\top\widehat{\gamma}_0(\widehat{\theta}_{\mathrm{SL}})\right\},$$

where, for $j = 0, 1$, $\widehat{\gamma}_j(\widehat{\theta}_{\mathrm{SL}})$ is equal to $\widehat{\gamma}_j$ in the main article with $\widehat{O}_{zij}$ replaced by

$$\widehat{O}_{zij}(\widehat{\theta}_{\mathrm{SL}}) = \int_0^\tau \frac{\{e^{\widehat{\theta}_{\mathrm{SL}}}\bar{Y}_{z1}(t)\}^{(1-j)}\{\bar{Y}_{z0}(t)\}^j}{e^{\widehat{\theta}_{\mathrm{SL}}}\bar{Y}_{z1}(t) + \bar{Y}_{z0}(t)}\left\{dN_{ij}(t) - \frac{Y_{ij}(t)e^{j\widehat{\theta}_{\mathrm{SL}}}d\bar{N}_z(t)}{e^{\widehat{\theta}_{\mathrm{SL}}}\bar{Y}_{z1}(t) + \bar{Y}_{z0}(t)}\right\}.$$

We obtain $\widehat{\theta}_{\mathrm{CSL}}$ from solving $\widehat{U}_{\mathrm{CSL}}(\vartheta) = 0$.

Let

$$O_{zi1}(\theta) = \int_0^\tau \left\{1 - \frac{s_z^{(1)}(\theta,t)}{s_z^{(0)}(\theta,t)}\right\}\left\{dN_{i1}(t) - Y_{i1}(t)e^\theta\frac{E(dN_i(t) \mid Z_i = z)}{s_z^{(0)}(\theta,t)}\right\}$$

$$O_{zi0}(\theta) = \int_0^\tau \frac{s_z^{(1)}(\theta,t)}{s_z^{(0)}(\theta,t)}\left\{dN_{i0}(t) - Y_{i0}(t)\frac{E(dN_i(t) \mid Z_i = z)}{s_z^{(0)}(\theta,t)}\right\},$$

where $s_z^{(1)}(\theta,t) = e^\theta\mu_z(t)E\{Y_i(t)\}$ and $s_z^{(0)}(\theta,t) = \{e^\theta\mu_z(t) + 1 - \mu_z(t)\}E\{Y_i(t)\}$.

THEOREM S3. *Under $C_I \perp T_I \mid I, Z$ and (D). Also assume the Cox proportional hazards model $\lambda_{z1}(t) = \lambda_{z0}(t)e^\theta$. Then, the following results hold regardless of which covariate-adaptive randomization scheme is applied.*

$$\sqrt{n}(\widehat{\theta}_{\mathrm{CSL}} - \theta) \xrightarrow{d} N\left(0, \frac{\sigma^2_{\mathrm{CSL}}(\theta)}{\sigma^4_{\mathrm{SL}}(\theta)}\right),$$

where $\sigma^2_{\mathrm{CSL}}(\theta) = \sigma^2_{\mathrm{SL}}(\theta) - \pi(1-\pi)(\gamma_1(\theta) + \gamma_0(\theta))^\top E\{\mathrm{var}(X_i \mid Z_i)\}(\gamma_1(\theta) + \gamma_0(\theta))$, $\sigma^2_{\mathrm{SL}}(\theta) = \sum_z \mathrm{pr}(Z_i = z)\{\pi\mathrm{var}(O_{zi1}(\theta) \mid Z_i = z) + (1-\pi)\mathrm{var}(O_{zi0}(\theta) \mid Z_i = z)\}$.

The proof of Theorem S3 is similar to the proof of Theorem S2 and thus is omitted. Based on Theorem S3, inference can be made based on the estimated variance $\widehat{\sigma}^2_{\mathrm{CSL}}(\widehat{\theta}_{\mathrm{CSL}})/\widehat{\sigma}^4_{\mathrm{SL}}(\widehat{\theta}_{\mathrm{CSL}})$, with

$$\widehat{\sigma}^2_{\mathrm{SL}}(\widehat{\theta}_{\mathrm{CSL}}) = -\partial\widehat{U}_{\mathrm{SL}}(\vartheta)/\partial\vartheta\mid_{\vartheta=\widehat{\theta}_{\mathrm{CSL}}} = n^{-1}\sum_z\sum_{i:Z_i=z}\int_0^\tau \frac{e^{\widehat{\theta}_{\mathrm{CSL}}}\bar{Y}_{z1}(t)\bar{Y}_{z0}(t)}{\{e^{\widehat{\theta}_{\mathrm{CSL}}}\bar{Y}_{z1}(t) + \bar{Y}_{z0}(t)\}^2}dN_i(t)$$

and $\widehat{\sigma}^2_{\mathrm{CSL}}(\widehat{\theta}_{\mathrm{CSL}}) = \widehat{\sigma}^2_{\mathrm{SL}}(\widehat{\theta}_{\mathrm{CSL}}) - \pi(1-\pi)(\widehat{\gamma}_1(\widehat{\theta}_{\mathrm{SL}}) + \widehat{\gamma}_0(\widehat{\theta}_{\mathrm{SL}}))^\top\{\sum_z(n_z/n)\widehat{\Sigma}_{X|z}\}(\widehat{\gamma}_1(\widehat{\theta}_{\mathrm{SL}}) + \widehat{\gamma}_0(\widehat{\theta}_{\mathrm{SL}}))$.

S2.4. Lemmas

The following lemmas are useful for proving the main theorems.

LEMMA S1. *Under condition (D),*
(a) $\widehat{\beta}_1 = \beta_1 + o_p(1)$ and $\widehat{\beta}_0 = \beta_0 + o_p(1)$.
(b) $\widehat{\gamma}_1 = \gamma_1 + o_p(1)$ and $\widehat{\gamma}_0 = \gamma_0 + o_p(1)$.

LEMMA S2. *Assume (C) and (D). Let E_{H_0}, λ_{H_0}, $\mu_{H_0}(t)$, and $p_{H_0}(t)$ be the expectation, hazard, $\mu(t)$, and $p(t)$ under the null hypothesis $H_0 : \lambda_1(t) = \lambda_0(t)$ for all t. Then, for any t,*
(a) $E_{H_0}\{I_iY_i(t)\} = \pi E_{H_0}\{Y_{i1}(t)\} = \mu_{H_0}(t)E_{H_0}\{Y_i(t)\}$,
(b) $E_{H_0}\{(1-I_i)Y_i(t)\} = (1-\pi)E_{H_0}\{Y_{i0}(t)\} = \{1 - \mu_{H_0}(t)\}E_{H_0}\{Y_i(t)\}$,
(c) $E_{H_0}\{Y_{ij}(t)\lambda_{H_0}(t)\} = p_{H_0}(t)E_{H_0}\{Y_{ij}(t)\},\ j = 0, 1$.

LEMMA S3. *Let* $\widetilde{\sigma}_{\mathrm{L}}^2 = \int_0^\tau \mu(t)\{1-\mu(t)\}E\{dN_i(t)\}$. *Assume (C) and (D),* $\widehat{\sigma}_{\mathrm{L}}^2 \xrightarrow{p} \widetilde{\sigma}_{\mathrm{L}}^2$ *and*

$$\widetilde{\sigma}_{\mathrm{L}}^2 = \int_0^\tau E\left[\mu(t)\{1-\mu(t)\}\{\pi Y_{i1}(t)\lambda_1(t) + (1-\pi)Y_{i0}(t)\lambda_0(t)\}\right] dt.$$

LEMMA S4. *Assume the conditions in Theorem 1 and the local alternative hypothesis specified in Theorem 1(c). Then,* $E(O_{ij}) \to 0$, $j = 0, 1$, *and both* σ_{L}^2 *and* $\widetilde{\sigma}_{\mathrm{L}}^2 \to \pi \mathrm{var}_{H_0}(O_{i1}) + (1-\pi)\mathrm{var}_{H_0}(O_{i0})$, *where* var_{H_0} *denotes the variance under* H_0.

S3. TECHNICAL PROOFS

S3.1. Proofs of Lemmas

PROOF OF LEMMA S1.
(a) We show the proof for $\widehat{\beta}_1$ and β_1. Note first that

$$\frac{1}{n_1}\sum_{i=1}^n I_i(X_i - \bar{X}_1)(X_i - \bar{X}_1)^\top \xrightarrow{p} \Sigma_X$$

from the proof of Lemma 3 in Ye et al. (2022). From Lemma 3 of Ye & Shao (2020), we have that $\bar{Y}_0(t) \xrightarrow{p} (1-\pi)E\{Y_{i0}(t)\}, \bar{Y}_1(t) \xrightarrow{p} \pi E\{Y_{i1}(t)\}, \bar{Y}(t) \xrightarrow{p} E\{Y_i(t)\}$. Similarly, we can show that $n_1^{-1}\sum_{i=1}^n I_i X_i dN_{i1}(t) \xrightarrow{p} E\{X_i dN_{i1}(t)\}$, $n_1^{-1}\sum_{i=1}^n I_i X_i Y_{i1}(t) \xrightarrow{p} E\{X_i Y_{i1}(t)\}$, and $\bar{N}(t) \xrightarrow{p} E\{N_i(t)\}$. Hence,

$$\frac{1}{n_1}\sum_{i=1}^n I_i(X_i - \bar{X}_1)\widehat{O}_{i1} \xrightarrow{p} \mathrm{cov}(X_i, O_{i1}),$$

concluding the proof that $\widehat{\beta}_1 = \beta_1 + o_p(1)$. The result for $\widehat{\beta}_0$ can be shown in the same way.
(b) The proof for $\widehat{\gamma}_1, \widehat{\gamma}_0$ and γ_1, γ_0 are similar and can be established from showing

$$\frac{1}{n_1}\sum_z \sum_{i:Z_i=z} I_i(X_i - \bar{X}_{z1})(X_i - \bar{X}_{z1})^\top \xrightarrow{p} E\{\mathrm{var}(X_i \mid Z_i)\}$$

$$\begin{aligned}\frac{1}{n_1}\sum_z \sum_{i:Z_i=z} I_i(X_i - \bar{X}_{z1})\widehat{O}_{zi1} &\xrightarrow{p} \sum_z P(Z=z)\mathrm{cov}\left(X_i, O_{zi1} \mid Z_i = z\right) \\ &= \mathrm{cov}\left(X_i, \sum_z I(Z_i = z)(O_{zi1} - \theta_{z1})\right),\end{aligned}$$

where $\gamma_j = E\{\mathrm{var}(X_i \mid Z_i)\}^{-1}\mathrm{cov}\left(X_i, \sum_z I(Z_i = z)(O_{zij} - \theta_{zj})\right)$.

PROOF OF LEMMA S2.
For simplicity, we remove the subscript H_0 and assume all calculations are under H_0.
(a) Note that

$$\begin{aligned}&\pi E\{Y_{i1}(t)\} = E\left[E(I_i \mid Z_1, \ldots, Z_n)E\{Y_{i1}(t) \mid Z_1, \ldots, Z_n\}\right] \\ &= E\left[I_i Y_i(t)\right] = E\left[E\{I_i \mid Y_i(t) = 1\}Y_i(t)\right] \\ &= \mu(t)E\left\{Y_i(t)\right\},\end{aligned}$$

where the first equality is because of $E(I_i \mid Z_1, \ldots, Z_n) = \pi$, the second equality is because of the conditional independence $I_i \perp Y_{i1}(t) \mid Z_1, \ldots, Z_n$ in (D).

(b) The proof is the same as that for (a).

(c) The result is straightforward from showing that

$$
\begin{aligned}
E\{Y_{i1}(t)\lambda(t)\} &= \pi^{-1}E\{I_iY_i(t)\lambda(t)\} = \pi^{-1}E\left[E\{I_i \mid Y_i(t)=1\}Y_i(t)\lambda(t)\right] \\
&= \pi^{-1}\mu(t)E\left[Y_i(t)\lambda(t)\right] = \pi^{-1}\mu(t)E\left[Y_i(t)\right]p(t) \\
E\{Y_{i1}(t)\} &= \pi^{-1}E\{I_iY_i(t)\} = \pi^{-1}E\left[E\{I_i \mid Y_i(t)=1\}Y_i(t)\right] \\
&= \pi^{-1}\mu(t)E\left[Y_i(t)\right],
\end{aligned}
$$

as well as the counterparts for $E\{Y_{i0}(t)\lambda(t)\}$ and $E\{Y_{i0}(t)\}$.

PROOF OF LEMMA S3.

The first result $\widehat{\sigma}^2_{\mathrm{L}} \xrightarrow{p} \widetilde{\sigma}^2_{\mathrm{L}}$ is because $\bar{Y}_1(t)\bar{Y}_0(t)/\bar{Y}(t)^2 \xrightarrow{p} \mu(t)\{1-\mu(t)\}$ from Lemma 3 of Ye & Shao (2020), and $d\bar{N}(t) \xrightarrow{p} E\{dN_i(t)\}$.

Then, under (C), from the theory of counting processes in survival analysis (Andersen & Gill, 1982), the process $N_{ij}(t)$ has random intensity process of the form $Y_{ij}(t)\lambda_j(t)$, $j=0,1$. Hence, for $i=1,...,n$, $j=0,1$, the process $N_{ij}(t)-\int_0^t Y_{ij}(s)\lambda_j(s)ds$ is a local square integrable martingale with respect to the filtration $\mathcal{F}_t=\sigma\{N_{ij}(u),(1-\delta_{ij})\mathcal{I}(X_{ij}\le u): 0\le u\le t\}$. From the fact that martingales have expectation zero, we conclude that $E\{dN_{ij}(t)\} = E\{Y_{ij}(t)\lambda_j(t)\}dt$. The second result follows from

$$
\begin{aligned}
E\{dN_i(t)\} &= E\{I_idN_{i1}(t)+(1-I_i)dN_{i0}(t)\} \\
&= \pi E\{Y_{i1}(t)\lambda_1(t)\}dt + (1-\pi)E\{Y_{i0}(t)\lambda_0(t)\}dt.
\end{aligned}
$$

PROOF OF LEMMA S4.

Note that

$$
E(O_{i1}) = \int_0^\tau E[\{1-\mu(t)\}Y_{i1}(t)\{\lambda_1(t)-p(t)\}]dt.
$$

Under the local alternative and from the dominated convergence theorem, for every t,

$$
\begin{aligned}
E\{Y_{i1}(t)\} &= e^{-\int_0^t \lambda_1(s)ds}\mathrm{pr}(C_1\ge t) \\
&\to e^{-\int_0^t \lambda_{H_0}(s)ds}\mathrm{pr}(C_1\ge t) \\
&= E_{H_0}\{Y_{i1}(t)\},
\end{aligned} \tag{S1}
$$

where λ_{H_0} and E_{H_0} denote the hazard and expectation under H_0, respectively. Similarly, we can show that

$$
E\left\{Y_{i1}(t)\lambda_1(t)\right\} \to E_{H_0}\left\{Y_{i1}(t)\lambda_{H_0}(t)\right\}. \tag{S2}
$$

These imply that $\mu(t)\to\mu_{H_0}(t)$ and $p(t)\to p_{H_0}(t)$, where $\mu_{H_0}(t)$ and $p_{H_0}(t)$ are $\mu(t)$ and $p(t)$ under H_0, respectively. Hence, again from the dominant convergence theorem,

$$
\begin{aligned}
E(O_{i1}) &= \int_0^\tau \{1-\mu(t)\}E[Y_{i1}(t)\{\lambda_1(t)-p(t)\}]dt \\
&\to \int_0^\tau \{1-\mu_{H_0}(t)\}E_{H_0}[Y_{i1}(t)\{\lambda_{H_0}(t)-p_{H_0}(t)\}]dt \\
&= E_{H_0}(O_{i1}) \\
&= 0
\end{aligned}
$$

where the last equality follows from Lemma S2(c). The proof for $E(O_{i0})\to 0$ is the same.

Let var_{H_0} denote the variance under H_0. Theorem S1 with $b_0 = b_1 = 0$ implies that $\sigma_{\mathrm{L}}^2 = \pi \mathrm{var}(O_{i1}) + (1-\pi)\mathrm{var}(O_{i0})$. Next, we show that under the local alternative, $\sigma_{\mathrm{L}}^2 \to \pi \mathrm{var}_{H_0}(O_{i1}) + (1-\pi)\mathrm{var}_{H_0}(O_{i0})$. For $\mathrm{var}(O_{ij})$, since $\mathrm{var}(O_{ij}) = E(O_{ij}^2) - \{E(O_{ij})\}^2$ and $E(O_{ij}) \to 0$, it suffices to show that $E(O_{ij}^2) \to \mathrm{var}_{H_0}(O_{ij}^2)$. Note that

$$
\begin{aligned}
E(O_{i1}^2) = E\Bigg[&\left\{\int_0^\tau \{1-\mu(t)\} dN_{i1}(t)\right\}^2 \\
&- 2\int_0^\tau \int_0^\tau \{1-\mu(t)\} Y_{i1}(t)\lambda_1(t)\{1-\mu(s)\} Y_{i1}(s) p(s) dt ds \\
&+ \int_0^\tau \int_0^\tau \{1-\mu(t)\}\{1-\mu(s)\} Y_{i1}(s) Y_{i1}(t) p(s) p(t) dt ds\Bigg] \\
= &\int_0^\tau \{1-\mu(t)\}^2 E\{dN_{i1}(t)\} \\
&- 2\int_0^\tau \int_{t \geq s} \{1-\mu(t)\}\{1-\mu(s)\} p(s) E\{Y_{i1}(t)\lambda_1(t) - Y_{i1}(t)p(t)\} dt ds,
\end{aligned}
$$

where the second equality is because, when $t > s$, $Y_{i1}(t) dN_{i1}(s) = 0$, $E\{Y_{i1}(s) dN_{i1}(t) \mid \mathcal{F}_{t-}\} = Y_{i1}(s) Y_{i1}(t) \lambda_1(t) dt$ for $t \geq s$, and $Y_{i1}(t) Y_{i1}(s) = Y_{i1}\{\max(t,s)\}$. These techniques will be used frequently in the following proofs, and will not be further elaborated. From (S1)-(S2), we have that $E\{Y_{i1}(t)\lambda_1(t) - Y_{i1}(t)p(t)\} \to 0$ for every t, and consequently

$$
E(O_{i1}^2) \to \int_0^\tau \{1-\mu_{H_0}(t)\}^2 E_{H_0}\{Y_{i1}(t)\lambda_{H_0}(t)\} dt,
$$

which is equal to $\mathrm{var}_{H_0}(O_{i1})$ by the same argument and the fact that $E_{H_0}(O_{i1}) = 0$. Similarly, we can show that $E(O_{i0}^2) \to \int_0^\tau \mu_{H_0}(t)^2 E_{H_0}\{Y_{i0}(t)\lambda_{H_0}(t)\} dt = \mathrm{var}_{H_0}(O_{i0})$. This concludes the proof that $\sigma_{\mathrm{L}}^2 \to \pi \mathrm{var}_{H_0}(O_{i1}) + (1-\pi)\mathrm{var}_{H_0}(O_{i0})$ under the local alternative.

For $\widetilde{\sigma}_{\mathrm{L}}^2$, under the local alternative, from Lemma S2, (S1)-(S2), and a similar argument as above,

$$
\begin{aligned}
\widetilde{\sigma}_{\mathrm{L}}^2 &= \int_0^\tau \mu(t)\{1-\mu(t)\}\left[\pi E\{Y_{i1}(t)\lambda_1(t)\} + (1-\pi) E\{Y_{i0}(t)\lambda_0(t)\}\right] dt \\
&\to \int_0^\tau \mu_{H_0}(t)\{1-\mu_{H_0}(t)\}\left[\pi E_{H_0}\{Y_{i1}(t)\lambda_{H_0}(t)\} + (1-\pi) E_{H_0}\{Y_{i0}(t)\lambda_{H_0}(t)\}\right] dt \\
&= \pi \mathrm{var}_{H_0}(O_{i1}) + (1-\pi)\mathrm{var}_{H_0}(O_{i0}).
\end{aligned}
$$

S3.2. Proofs of Theorems

PROOF OF THEOREM 1.

(a) Following a Taylor expansion as in the Appendix of Lin & Wei (1989), we obtain that under either the null or alternative hypothesis,

$$
\begin{aligned}
\widehat{U}_{\mathrm{L}} &= n^{-1} \sum_{i=1}^n \int_0^\tau \{I_i - \mu(t)\}\{dN_i(t) - Y_i(t) p(t) dt\} + o_p(n^{-1/2}) \\
&= n^{-1} \sum_{i=1}^n \{I_i O_{i1} - (1-I_i) O_{i0}\} + o_p(n^{-1/2}),
\end{aligned}
$$

where $(O_{i1}, O_{i0}), i = 1, \ldots, n$ are i.i.d.. Then from $n^{-1}\sum_{i=1}^{n} I_i(X_i - \bar{X}) = n^{-1}n_1(\bar{X}_1 - \bar{X}) = O_p(n^{-1/2})$, and Lemma S1, we have that

$$\widehat{U}_{\mathrm{CL}} = n^{-1}\sum_{i=1}^{n}\Big[I_i\{O_{i1} - (X_i - \bar{X})^\top\beta_1\} - (1 - I_i)\{O_{i0} - (X_i - \bar{X})^\top\beta_0\}\Big] + o_p(n^{-1/2}).$$

The rest of the proof is similar to the proof of Theorem 2 in Ye et al. (2022). Define $\mathcal{I} = \{I_1, \ldots, I_n\}$ and $\mathcal{S} = \{Z_1, \ldots, Z_n\}$, then

$$\begin{aligned}
&\widehat{U}_{\mathrm{CL}} - \Big(\frac{n_1}{n}\theta_1 - \frac{n_0}{n}\theta_0\Big)\\
&= \frac{1}{n}\sum_{i=1}^{n} I_i\{O_{i1} - \theta_1 - (X_i - \bar{X})^\top\beta_1\} - (1 - I_i)\{O_{i0} - \theta_0 - (X_i - \bar{X})^\top\beta_0\} + o_p(n^{-1/2})\\
&= \frac{1}{n}\sum_{i=1}^{n} I_i\{O_{i1} - \theta_1 - (X_i - \mu_X)^\top\beta_1\} - (1 - I_i)\{O_{i0} - \theta_0 - (X_i - \mu_X)^\top\beta_0\}\\
&\quad + \frac{n_1}{n}(\bar{X} - \mu_X)^\top\beta_1 - \frac{n_0}{n}(\bar{X} - \mu_X)^\top\beta_0 + o_p(n^{-1/2})\\
&= \frac{1}{n}\sum_{i=1}^{n} I_i\{O_{i1} - \theta_1 - (X_i - \mu_X)^\top\beta_1\} - \frac{1}{n}\sum_{i=1}^{n}(1 - I_i)\{O_{i0} - \theta_0 - (X_i - \mu_X)^\top\beta_0\}\\
&\quad + \pi(\bar{X} - \mu_X)^\top\beta_1 - (1 - \pi)(\bar{X} - \mu_X)^\top\beta_0 + o_p(n^{-1/2})\\
&= \underbrace{\frac{1}{n}\sum_{i=1}^{n} I_i\{O_{i1} - \theta_1 - (X_i - \mu_X)^\top\beta_1\}}_{M_1} - \underbrace{\frac{1}{n}\sum_{i=1}^{n}(1 - I_i)\{O_{i0} - \theta_0 - (X_i - \mu_X)^\top\beta_0\}}_{M_2}\\
&\quad + \underbrace{(\bar{X} - E(\bar{X} \mid \mathcal{I}, \mathcal{S}))^\top(\pi\beta_1 - (1 - \pi)\beta_0)}_{M_3} + \underbrace{(E(\bar{X} \mid \mathcal{I}, \mathcal{S}) - \mu_X)^\top(\pi\beta_1 - (1 - \pi)\beta_0)}_{M_4}\\
&\quad + o_p(n^{-1/2})\\
&:= M_1 - M_2 + M_3 + M_4 + o_p(n^{-1/2}),
\end{aligned}$$

By using the definition $\beta_j = \Sigma_X^{-1}cov(X_i, O_{ij})$, we have

$$E\big[X_i^\top\{O_{ij} - \theta_j - (X_i - \mu_X)^\top\beta_j\}\big] = cov(X_i, O_{ij}) - cov(X_i, O_{ij}) = 0$$

Because Z_i is discrete and X_i contains all joint levels of Z_i as a sub-vector, according to the estimation equations from the least squares, we have that

$$E\left[I(Z_i = z)\{O_{ij} - \theta_j - (X_i - \mu_X)^\top\beta_j\}\right] = 0,\ \forall z \in \mathcal{Z},$$

and thus,

$$E\left[O_{ij} - \theta_j - (X_i - \mu_X)^\top\beta_j \mid Z_i\right] = 0,\ \text{a.s..} \tag{S3}$$

Next, we show that $\sqrt{n}(M_1 - M_2 + M_3)$ is asymptotically normal. Consider the random vector

$$\sqrt{n}\begin{pmatrix} E_n\left[I_i(O_{i1} - \theta_1 - (X_i - \mu_X)^\top\beta_1)\right] \\ E_n\left[(1 - I_i)(O_{i0} - \theta_0 - (X_i - \mu_X)^\top\beta_0)\right] \\ E_n\left[(X_i - E(X_i \mid Z_i))\right] \end{pmatrix}, \tag{S4}$$

where $E_n[K_i] := \frac{1}{n}\sum_{i=1}^n K_i$. Conditional on $\mathcal{I}, \mathcal{S}$, every component in (S4) is an average of independent terms. Similar to the proof of Theorem 2 in Ye et al. (2022), the Lindeberg's Central Limit Theorem justifies that (S4) is asymptotically normal with mean 0 conditional on $\mathcal{I}, \mathcal{S}$, as $n \to \infty$. This implies that $\sqrt{n}(M_1 - M_2 + M_3)$ is asymptotically normal with mean 0 conditional on $\mathcal{I}, \mathcal{S}$. Then, we calculate its variance. Note that

$$
\begin{aligned}
\operatorname{var}\left(\sqrt{n}(M_1 - M_2) \mid \mathcal{I}, \mathcal{S}\right) &= \frac{1}{n}\sum_{i=1}^n I_i \operatorname{var}(O_{i1} - \theta_1 - (X_i - \mu_X)^\top \beta_1 \mid Z_i) \\
&\quad + \frac{1}{n}\sum_{i=1}^n (1 - I_i) \operatorname{var}(O_{i0} - \theta_0 - (X_i - \mu_X)^\top \beta_0 \mid Z_i) \\
&= \frac{1}{n}\sum_z \sum_{i:Z_i = z} I_i \operatorname{var}(O_{i1} - \theta_1 - (X_i - \mu_X)^\top \beta_1 \mid Z_i = z) \\
&\quad + (1 - I_i) \operatorname{var}(O_{i0} - \theta_0 - (X_i - \mu_X)^\top \beta_0 \mid Z_i = z) \\
&= \sum_z \frac{n_1(z)}{n} \operatorname{var}(O_{i1} - \theta_1 - (X_i - \mu_X)^\top \beta_1 \mid Z_i = z) \\
&\quad + \frac{n_0(z)}{n} \operatorname{var}(O_{i0} - \theta_0 - (X_i - \mu_X)^\top \beta_0 \mid Z_i = z) \\
&= \sum_z \frac{n_1(z)}{n(z)} \frac{n(z)}{n} \operatorname{var}(O_{i1} - \theta_1 - (X_i - \mu_X)^\top \beta_1 \mid Z_i = z) \\
&\quad + \frac{n_0(z)}{n(z)} \frac{n(z)}{n} \operatorname{var}(O_{i0} - \theta_0 - (X_i - \mu_X)^\top \beta_0 \mid Z_i = z) \\
&= \pi \sum_z P(Z = z) \operatorname{var}(O_{i1} - \theta_1 - (X_i - \mu_X)^\top \beta_1 \mid Z_i = z) \\
&\quad + (1 - \pi) \sum_z P(Z = z) \operatorname{var}(O_{i0} - \theta_0 - (X_i - \mu_X)^\top \beta_0 \mid Z_i = z) + o_p(1) \\
&= \pi E\big\{\operatorname{var}(O_{i1} - \theta_1 - (X_i - \mu_X)^\top \beta_1 \mid Z_i)\big\} \\
&\quad + (1 - \pi) E\big\{\operatorname{var}(O_{i0} - \theta_0 - (X_i - \mu_X)^\top \beta_0 \mid Z_i)\big\} + o_p(1) \\
&= \pi \operatorname{var}(O_{i1} - \theta_1 - (X_i - \mu_X)^\top \beta_1) \\
&\quad + (1 - \pi) \operatorname{var}(O_{i0} - \theta_0 - (X_i - \mu_X)^\top \beta_0) + o_p(1) \\
&= \pi \operatorname{var}(O_{i1} - X_i^\top \beta_1) + (1 - \pi) \operatorname{var}(O_{i0} - X_i^\top \beta_0) + o_p(1),
\end{aligned}
$$

$$
\operatorname{var}\left(\sqrt{n}\bar{X} \mid \mathcal{I}, \mathcal{S}\right) = \frac{1}{n}\sum_{i=1}^n \operatorname{var}(X_i \mid Z_i) = E\{\operatorname{var}(X_i \mid Z_i)\} + o_p(1),
$$

and

$$
\begin{aligned}
n\operatorname{cov}(M_1, \bar{X} \mid \mathcal{I}, \mathcal{S}) &= n\operatorname{cov}\left(\frac{1}{n}\sum_{i=1}^n I_i\{O_{i1} - \theta_1 - (X_i - \mu_X)^\top \beta_1\}, \frac{1}{n}\sum_{i=1}^n X_i \mid \mathcal{I}, \mathcal{S}\right) \\
&= \frac{1}{n}\sum_{i=1}^n I_i \operatorname{cov}\Big(O_{i1} - \theta_1 - (X_i - \mu_X)^\top \beta_1, X_i \mid Z_i\Big)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_z \frac{n_1(z)}{n(z)}\frac{n(z)}{n}\text{cov}\left(O_{i1}-\theta_1-(X_i-\mu_X)^\top\beta_1, X_i \mid Z_i=z\right)\\
&= \pi\sum_z P(Z=z)\text{cov}\left(O_{i1}-\theta_1-(X_i-\mu_X)^\top\beta_1, X_i \mid Z_i=z\right)+o_p(1)\\
&= \pi E\left\{\text{cov}\left(O_{i1}-\theta_1-(X_i-\mu_X)^\top\beta_1, X_i \mid Z_i\right)\right\}+o_p(1)\\
&= o_p(1),
\end{aligned}
$$

where the last equality holds because $E(O_{i1}-X_i^\top\beta_1 \mid Z_i)=\theta_1-\mu_X^\top\beta_1$ and, thus, $\text{cov}\{E(O_{i1}-X_i^\top\beta_1 \mid Z_i), E(X_i \mid Z_i)\}=0$ and $E\big\{\text{cov}(O_{i1}-\theta_1-(X_i-\mu_X)^\top\beta_1, X_i \mid Z_i)\big\}=\text{cov}(O_{i1}-\theta_1-(X_i-\mu_X)^\top\beta_1, X_i)=0$ according to the definition of β_1. Similarly, we can show that $n\text{cov}(M_2,\bar{X} \mid \mathcal{I},\mathcal{S})=o_p(1)$.

Combining the above derivations and from the Slutsky's theorem, we have shown that

$$
\begin{aligned}
\sqrt{n}(M_1-M_2+M_3) \mid \mathcal{I},\mathcal{S} \xrightarrow{d} N\Big(0, &\pi\text{var}(O_{i1}-X_i^\top\beta_1)+(1-\pi)\text{var}(O_{i0}-X_i^\top\beta_0)\\
&+(\pi\beta_1-(1-\pi)\beta_0)^\top E\{\text{var}(X_i \mid Z_i)\}(\pi\beta_1-(1-\pi)\beta_0)\Big).
\end{aligned}
$$

From the bounded convergence theorem, this result also holds unconditionally, i.e.,

$$
\begin{aligned}
\sqrt{n}(M_1-M_2+M_3) \xrightarrow{d} N\Big(0, &\pi\text{var}(O_{i1}-X_i^\top\beta_1)+(1-\pi)\text{var}(O_{i0}-X_i^\top\beta_0)\\
&+(\pi\beta_1-(1-\pi)\beta_0)^\top E\{\text{var}(X_i \mid Z_i)\}(\pi\beta_1-(1-\pi)\beta_0)\Big).
\end{aligned}
$$

Moreover, since M_4 is an average of i.i.d. terms, by the central limit theorem,

$$
\sqrt{n}(E(\bar{X} \mid \mathcal{I},\mathcal{S})-\mu_X)=n^{-1/2}\sum_{i=1}^n\{E(X_i \mid Z_i)-\mu_X\} \xrightarrow{d} N(0,\text{var}(E(X_i \mid Z_i))),
$$

and

$$
\sqrt{n}M_4 \xrightarrow{d} N\big(0,(\pi\beta_1-(1-\pi)\beta_0)^\top\text{var}\{E(X_i \mid Z_i)\}(\pi\beta_1-(1-\pi)\beta_0)\big)
$$

Next, we show that $(\sqrt{n}(M_1-M_2+M_3),\sqrt{n}M_4) \xrightarrow{d} (\xi_1,\xi_2)$, where (ξ_1,ξ_2) are mutually independent. This can be seen from

$$
\begin{aligned}
&P(\sqrt{n}(M_1-M_2+M_3)\le t_1,\sqrt{n}M_4\le t_2)\\
&=E\{I(\sqrt{n}(M_1-M_2+M_3)\le t_1)I(\sqrt{n}M_4\le t_2)\}\\
&=E\{P(\sqrt{n}(M_1-M_2+M_3)\le t_1 \mid \mathcal{I},\mathcal{S})I(\sqrt{n}M_4\le t_2)\}\\
&=E\big[\{P(\sqrt{n}(M_1-M_2+M_3)\le t_1 \mid \mathcal{I},\mathcal{S})-P(\xi_1\le t_1)\}I(\sqrt{n}M_4\le t_2)\big]\\
&\quad+P(\xi_1\le t_1)P(\sqrt{n}M_4\le t_2)\\
&\to P(\xi_1\le t_1)P(\xi_2\le t_2),
\end{aligned}
$$

where the last step follows from the bounded convergence theorem. Finally, using the definitions of β_0,β_1, it is easy to show that

$$
\begin{aligned}
&\pi\text{var}(O_{i1}-X_i^\top\beta_1)+(1-\pi)\text{var}(O_{i0}-X_i^\top\beta_0)+(\pi\beta_1-(1-\pi)\beta_0)^\top\Sigma_X(\pi\beta_1-(1-\pi)\beta_0)\\
&=\pi\text{var}(O_{i1})+(1-\pi)\text{var}(O_{i0})-\pi(1-\pi)(\beta_1+\beta_0)^\top\Sigma_X(\beta_1+\beta_0),
\end{aligned}
$$

concluding the proof that

$$\sqrt{n}\left\{\widehat{U}_{\mathrm{CL}}-\left(\frac{n_1}{n}\theta_1-\frac{n_0}{n}\theta_0\right)\right\}\xrightarrow{d}N(0,\sigma^2_{\mathrm{CL}}).$$

(b) Since

$$\begin{aligned}\theta_1=E(O_{i1})&=\int_0^\tau\{1-\mu(t)\}\left[E\{dN_{i1}(t)\}-E\{Y_{i1}(t)\}p(t)dt\right]\\&=\int_0^\tau\{1-\mu(t)\}\left[E\{Y_{i1}(t)\lambda_1(t)\}dt-E\{Y_{i1}(t)\}p(t)dt\right],\end{aligned}$$

Thus, the fact that $\theta_1=0$ under H_0 follows from Lemma S2(c). Similarly, $\theta_0=0$ under H_0.

Next, $\widehat{\sigma}^2_{\mathrm{CL}}\xrightarrow{p}\sigma^2_{\mathrm{CL}}$ under H_0 is from $\widehat{\sigma}^2_{\mathrm{L}}\xrightarrow{p}\sigma^2_{\mathrm{L}}$ as shown in the Proof of (17) in Ye & Shao (2020), and $\widehat{\beta}_j=\beta_j+o_p(1), j=0,1$ and $\widehat{\Sigma}_X=\Sigma_X+o_p(1)$ from Lemma S1. The result that $\mathcal{T}_{\mathrm{CL}}\xrightarrow{d}N(0,1)$ follows from Slutsky's theorem.

(c) Under the local alternative, from Lemma S4, we have that $\widehat{\sigma}^2_{\mathrm{CL}}=\sigma^2_{\mathrm{CL}}+o_p(1)$, and thus

$$\begin{aligned}\mathcal{T}_{\mathrm{CL}}-\frac{\pi c_1-(1-\pi)c_0}{\sigma_{\mathrm{CL}}}&=\frac{\sqrt{n}\widehat{U}_{\mathrm{CL}}}{\widehat{\sigma}_{\mathrm{CL}}}-\frac{\pi c_1-(1-\pi)c_0}{\sigma_{\mathrm{CL}}}\\&=\frac{\sqrt{n}\left(\widehat{U}_{\mathrm{CL}}-\frac{n_1\theta_1-n_0\theta_0}{n}\right)}{\widehat{\sigma}_{\mathrm{CL}}}+\frac{n_1c_1-n_0c_0}{n\widehat{\sigma}_{\mathrm{CL}}}-\frac{\pi c_1-(1-\pi)c_0}{\sigma_{\mathrm{CL}}}\\&=\frac{\sqrt{n}\left(\widehat{U}_{\mathrm{CL}}-\frac{n_1\theta_1-n_0\theta_0}{n}\right)}{\sigma_{\mathrm{CL}}}+o_p(1)\\&\xrightarrow{d}N(0,1).\end{aligned}$$

PROOF OF THEOREM S1.

Similar to the Taylor expansion in the proof of Theorem 1(a),

$$\begin{aligned}&\widehat{U}_{\mathrm{CL}}(b_0,b_1)-\frac{n_1\theta_1-n_0\theta_0}{n}\\&=\frac{1}{n}\sum_{i=1}^n\Big[I_i\{O_{i1}-\theta_1-(X_i-\mu_X)^\top b_1\}-(1-I_i)\{O_{i0}-\theta_0-(X_i-\mu_X)^\top b_0\}\Big]\\&\quad+\frac{n_1}{n}(\bar{X}-\mu_X)^\top b_1-\frac{n_0}{n}(\bar{X}-\mu_X)^\top b_0+o_p(n^{-1/2})\\&=\underbrace{\frac{1}{n}\sum_{i=1}^n I_i\Big[\{O_{i1}-\theta_1-(X_i-\mu_X)^\top b_1\}-E\{O_{i1}-\theta_1-(X_i-\mu_X)^\top b_1\mid Z_i\}\Big]}_{M_1}\\&\quad\underbrace{-\frac{1}{n}\sum_{i=1}^n(1-I_i)\Big[\{O_{i0}-\theta_0-(X_i-\mu_X)^\top b_0\}-E\{O_{i0}-\theta_0-(X_i-\mu_X)^\top b_0\mid Z_i\}\Big]}_{M_2}\\&\quad\underbrace{+\left(\bar{X}-\frac{1}{n}\sum_{i=1}^n E(X_i\mid Z_i)\right)^\top(\pi b_1-(1-\pi)b_0)}_{M_3}\end{aligned}$$

$$+\underbrace{\frac{1}{n}\sum_{i=1}^{n}(I_i-\pi)\Big[E\{O_{i1}-\theta_1-(X_i-\mu_X)^\top b_1\mid Z_i\}+E\{O_{i0}-\theta_0-(X_i-\mu_X)^\top b_0\mid Z_i\}\Big]}_{M_4}$$

$$+\underbrace{\frac{1}{n}\sum_{i=1}^{n}\Big[\pi E\{O_{i1}-\theta_1-(X_i-\mu_X)^\top b_1\mid Z_i\}-(1-\pi)E\{O_{i0}-\theta_0-(X_i-\mu_X)^\top b_0\mid Z_i\}\Big]}_{M_5}$$

$$+\underbrace{\left(\frac{1}{n}\sum_{i=1}^{n}E(X_i\mid Z_i)-\mu_X\right)^\top(\pi b_1-(1-\pi)b_0)}_{M_6}+o_p(n^{-1/2}).$$

In what follows, we will analyze these terms separately. Note that

$$\begin{aligned}
n\mathrm{var}(M_1-M_2\mid\mathcal{I},\mathcal{S}) &= \pi E\left\{\mathrm{var}(O_{i1}-X_i^\top b_1\mid Z_i)\right\}\\
&\quad+(1-\pi)E\left\{\mathrm{var}(O_{i0}-X_i^\top b_0\mid Z_i)\right\}+o_p(1)\\
n\mathrm{var}(M_3\mid\mathcal{I},\mathcal{S}) &= (\pi b_1-(1-\pi)b_0)^\top E\{\mathrm{var}(X_i\mid Z_i)\}(\pi b_1-(1-\pi)b_0)\\
n\mathrm{cov}(M_1-M_2,M_3\mid\mathcal{I},\mathcal{S}) &= \{\pi(\beta_1-b_1)-(1-\pi)(\beta_0-b_0)\}^\top E\{\mathrm{var}(X_i\mid Z_i)\}(\pi b_1-(1-\pi)b_0).
\end{aligned}$$

Similar to the proof of Theorem 3 in Ye et al. (2022), the Lindeberg's Central Limit Theorem and Slutsky theorem justify that $\sqrt{n}(M_1-M_2+M_3)$ is asymptotically normal with mean 0 conditional on $\mathcal{I},\mathcal{S}$. Namely,

$$\begin{aligned}
\sqrt{n}(M_1-M_2+M_3)\mid\mathcal{I},\mathcal{S}\xrightarrow{d}N\Big(&0,\pi E\{\mathrm{var}(O_{i1}-X_i^\top b_1\mid Z_i)\}\\
&+(1-\pi)E\{\mathrm{var}(O_{i0}-X_i^\top b_0\mid Z_i)\}+\{2\pi\beta_1-2(1-\pi)\beta_0\\
&-\pi b_1+(1-\pi)b_0\}^\top E\{\mathrm{var}(X_i\mid Z_i)\}(\pi b_1-(1-\pi)b_0)\Big).
\end{aligned}$$

Moreover, from (C4), $\sqrt{n}M_4$ is asymptotically normal conditional on $\mathcal{S}$, i.e.,

$$\sqrt{n}M_4\mid\mathcal{S}\xrightarrow{d}N\left(0,\nu\mathrm{var}\{E(O_{i1}-X_i^\top b_1\mid Z_i)+E(O_{i0}-X_i^\top b_0\mid Z_i)\}\right).$$

Because M_5,M_6 only involve sums of identically and independently distributed terms, $E(M_5+M_6)=0$, and

$$\begin{aligned}
n\mathrm{var}(M_5) &= \mathrm{var}\left\{\pi E(O_{i1}-X_i^\top b_1\mid Z_i)-(1-\pi)E(O_{i0}-X_i^\top b_0\mid Z_i)\right\}\\
n\mathrm{var}(M_6) &= (\pi b_1-(1-\pi)b_0)^\top\mathrm{var}\{E(X_i\mid Z_i)\}(\pi b_1-(1-\pi)b_0)\\
n\mathrm{cov}(M_5,M_6) &= \{\pi(\beta_1-b_1)-(1-\pi)(\beta_0-b_0)\}^\top\mathrm{var}\{E(X_i\mid Z_i)\}(\pi b_1-(1-\pi)b_0),
\end{aligned}$$

we therefore have

$$\begin{aligned}
\sqrt{n}(M_5+M_6)\xrightarrow{d}N\Big(&0,\mathrm{var}\left\{\pi E(O_{i1}-X_i^\top b_1\mid Z_i)-(1-\pi)E(O_{i0}-X_i^\top b_0\mid Z_i)\right\}\\
&\{2\pi\beta_1-2(1-\pi)\beta_0-\pi b_1+(1-\pi)b_0\}^\top\mathrm{var}\{E(X_i\mid Z_i)\}(\pi b_1-(1-\pi)b_0)\Big).
\end{aligned}$$

Combining all the above derivations and similarly to the proof of Theorem 1, we can show that $(\sqrt{n}(M_1 - M_2 + M_3), \sqrt{n}M_4, \sqrt{n}(M_5 + M_6)) \xrightarrow{d} (\xi_1, \xi_2, \xi_3)$, where (ξ_1, ξ_2, ξ_3) are mutually independent. Therefore, $\sqrt{n}\left(\widehat{U}_{\mathrm{CL}}(b_0, b_1) - \frac{n_1\theta_1 - n_0\theta_0}{n}\right)$ is asymptotically normal with mean 0 and variance

$$\begin{aligned}\sigma^2_{\mathrm{L}}(b_0, b_1) =& \pi E\left\{\mathrm{var}(O_{i1} - X_i^\top b_1 \mid Z_i)\right\} + (1-\pi)E\left\{\mathrm{var}(O_{i0} - X_i^\top b_0 \mid Z_i)\right\} \\ &+ \nu \mathrm{var}\left\{E(O_{i1} - X_i^\top b_1 \mid Z_i) + E(O_{i0} - X_i^\top b_0 \mid Z_i)\right\} \\ &+ \mathrm{var}\left\{\pi E(O_{i1} - X_i^\top b_1 \mid Z_i) - (1-\pi)E(O_{i0} - X_i^\top b_0 \mid Z_i)\right\} \\ &+ \{2\pi\beta_1 - 2(1-\pi)\beta_0 - \pi b_1 + (1-\pi)b_0\}^\top \mathrm{var}(X_i)(\pi b_1 - (1-\pi)b_0).\end{aligned}$$

Let $\sigma^2_{\mathrm{CL,SR}}(\beta_0, \beta_1)$ be the asymptotic variance under simple randomization. From the fact that $\sigma^2_{\mathrm{CL}}(\beta_0, \beta_1) = \sigma^2_{\mathrm{CL,SR}}(\beta_0, \beta_1)$, we have

$$\sigma^2_{\mathrm{CL}}(b_0, b_1) - \sigma^2_{\mathrm{CL}}(\beta_0, \beta_1) = \sigma^2_{\mathrm{CL}}(b_0, b_1) - \sigma^2_{\mathrm{CL},SR}(b_0, b_1) + \sigma^2_{\mathrm{CL,SR}}(b_0, b_1) - \sigma^2_{\mathrm{CL,SR}}(\beta_0, \beta_1).$$

Note first that

$$\begin{aligned}&\sigma^2_{\mathrm{CL}}(b_0, b_1) - \sigma^2_{\mathrm{CL,SR}}(b_0, b_1) \\ &= -\{\pi(1-\pi) - \nu\}\mathrm{var}\left\{E(O_{i1} - X_i^\top b_1 \mid Z_i) + E(O_{i0} - X_i^\top b_0 \mid Z_i)\right\} \\ &= -\{\pi(1-\pi) - \nu\}(\beta_1 - b_1 + \beta_0 - b_0)^\top \mathrm{var}\{E(X_i \mid Z_i)\}(\beta_1 - b_1 + \beta_0 - b_0),\end{aligned}$$

where the third line is because X_i includes all joint levels of Z_i and thus $E(O_{i1} - X_i^\top \beta_1 \mid Z_i) = \theta_1 - \mu_X^\top \beta_1$.

To calculate $\sigma^2_{\mathrm{CL,SR}}(b_0, b_1) - \sigma^2_{\mathrm{CL,SR}}(\beta_0, \beta_1)$, we first note that

$$\begin{aligned}\sigma^2_{\mathrm{CL,SR}}(b_0, b_1) &= \mathrm{var}\left[I_i\{O_{i1} - \theta_1 - (X_i - \mu_X)^\top b_1\} - (1 - I_i)\{O_{i0} - \theta_0 - (X_i - \mu_X)^\top b_0\}\right] \\ &\quad + \{2\pi\beta_1 - 2(1-\pi)\beta_0 - \pi b_1 + (1-\pi)b_0\}^\top \Sigma_X\{\pi b_1 - (1-\pi)b_0\} \\ &= \mathrm{var}\Bigg[I_i\{O_{i1} - \theta_1 - (X_i - \mu_X)^\top \beta_1 + (X_i - \mu_X)^\top(\beta_1 - b_1)\} \\ &\quad - (1 - I_i)\{O_{i0} - \theta_0 - (X_i - \mu_X)^\top \beta_0 + (X_i - \mu_X)^\top(\beta_0 - b_0)\}\Bigg] \\ &\quad + \{2\pi\beta_1 - 2(1-\pi)\beta_0 - \pi b_1 + (1-\pi)b_0\}^\top \Sigma_X\{\pi b_1 - (1-\pi)b_0\} \\ &= \mathrm{var}\left[I_i\{O_{i1} - \theta_1 - (X_i - \mu_X)^\top \beta_1\} - (1 - I_i)\{O_{i0} - \theta_0 - (X_i - \mu_X)^\top \beta_0\}\right] \\ &\quad + \mathrm{var}\left[I_i(X_i - \mu_X)^\top(\beta_1 - b_1) - (1 - I_i)(X_i - \mu_X)^\top(\beta_0 - b_0)\right] \\ &\quad + \{2\pi\beta_1 - 2(1-\pi)\beta_0 - \pi b_1 + (1-\pi)b_0\}^\top \Sigma_X\{\pi b_1 - (1-\pi)b_0\}.\end{aligned}$$

Then, we can calculate that

$$\begin{aligned}&\sigma^2_{\mathrm{CL,SR}}(b_0, b_1) - \sigma^2_{\mathrm{CL,SR}}(\beta_0, \beta_1) \\ &= \mathrm{var}\left\{I_i(X_i - \mu_X)^\top(\beta_1 - b_1) - (1 - I_i)(X_i - \mu_X)^\top(\beta_0 - b_0)\right\} \\ &\quad + \{2\pi\beta_1 - 2(1-\pi)\beta_0 - \pi b_1 + (1-\pi)b_0\}^\top \Sigma_X\{\pi b_1 - (1-\pi)b_0\} \\ &\quad - \{\pi\beta_1 - (1-\pi)\beta_0\}^\top \Sigma_X\{\pi\beta_1 - (1-\pi)\beta_0\}\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{var}\left\{I_i(X_i-\mu_X)^\top(\beta_1-b_1)\right\} + \mathrm{var}\left\{(1-I_i)(X_i-\mu_X)^\top(\beta_0-b_0)\right\} \\
&\quad + \{2\pi\beta_1 - 2(1-\pi)\beta_0 - \pi b_1 + (1-\pi)b_0\}^\top \Sigma_X \{\pi b_1 - (1-\pi)b_0\} \\
&\quad - \{\pi\beta_1 - (1-\pi)\beta_0\}^\top \Sigma_X \{\pi\beta_1 - (1-\pi)\beta_0\} \\
&= \pi(\beta_1-b_1)^\top\Sigma_X(\beta_1-b_1) + (1-\pi)(\beta_0-b_0)^\top\Sigma_X(\beta_0-b_0) \\
&\quad + \{\pi\beta_1 - (1-\pi)\beta_0\}^\top \Sigma_X \{\pi b_1 - (1-\pi)b_0\} \\
&\quad + \{\pi(\beta_1-b_1) - (1-\pi)(\beta_0-b_0)\}^\top \Sigma_X \{\pi b_1 - (1-\pi)b_0\} \\
&\quad - \{\pi\beta_1 - (1-\pi)\beta_0\}^\top \Sigma_X \{\pi\beta_1 - (1-\pi\beta_0)\} \\
&= \pi(1-\pi)(\beta_1 - b_1 + \beta_0 - b_0)^\top \Sigma_X (\beta_1 - b_1 + \beta_0 - b_0).
\end{aligned}
$$

Combining aforementioned results, we conclude that

$$
\begin{aligned}
&\sigma^2_{\mathrm{CL}}(b_0,b_1) - \sigma^2_{\mathrm{CL}}(\beta_0,\beta_1) \\
&\quad = \pi(1-\pi)(\beta_1-b_1+\beta_0-b_0)^\top E\{\mathrm{var}(X_i \mid Z_i)\}(\beta_1-b_1+\beta_0-b_0) \\
&\qquad + \nu(\beta_1-b_1+\beta_0-b_0)^\top \mathrm{var}\{E(X_i \mid Z_i)\}(\beta_1-b_1+\beta_0-b_0),
\end{aligned}
$$

which is greater or equal to zero because $E\{\mathrm{var}(X_i \mid Z_i)\}$ and $\mathrm{var}\{E(X_i \mid Z_i)\}$ are positive definite.

PROOF OF THEOREM S2.

Since $\widehat{\theta}_{\mathrm{CL}}$ solves $\widehat{U}_{\mathrm{CL}}(\vartheta)=0$, from the standard argument of M-estimation, we will show that under the Cox model $\lambda_1(t)=\lambda_0(t)e^\theta$,

$$
\sqrt{n}\widehat{U}_{\mathrm{CL}}(\theta) \xrightarrow{d} N(0, \sigma^2_{\mathrm{CL}}(\theta)) \tag{S5}
$$

$$
-\partial\widehat{U}_{\mathrm{CL}}(\vartheta)/\partial\vartheta \mid_{\vartheta=\bar\vartheta} = -\partial\widehat{U}_{\mathrm{L}}(\vartheta)/\partial\vartheta \mid_{\vartheta=\bar\vartheta} \xrightarrow{p} \sigma^2_{\mathrm{L}}(\theta) \tag{S6}
$$

where $\bar\vartheta$ lies between $\widehat{\theta}_{\mathrm{CL}}$ and θ. Therefore,

$$
\sqrt{n}(\widehat{\theta}_{\mathrm{CL}} - \theta) \xrightarrow{d} N\left(0, \frac{\sigma^2_{\mathrm{CL}}(\theta)}{\sigma^4_{\mathrm{L}}(\theta)}\right).
$$

We first consider (S5). Following the steps in the proof of Theorem 1, we can linearize $\widehat{U}_{\mathrm{L}}(\theta_0)$ and obtain

$$
\begin{aligned}
\widehat{U}_{\mathrm{L}}(\theta) &= \frac{1}{n}\sum_{i=1}^n \{I_i O_{i1}(\theta) - (1-I_i)O_{i0}(\theta)\} + n^{-1/2}o_p(1), \\
O_{i1}(\theta) &= \int_0^\tau \left\{1 - \frac{s^{(1)}(\theta,t)}{s^{(0)}(\theta,t)}\right\}\left\{dN_{i1}(t) - Y_{i1}(t)e^\theta \frac{E(dN_i(t))}{s^{(0)}(\theta,t)}\right\} \\
O_{i0}(\theta) &= \int_0^\tau \frac{s^{(1)}(\theta,t)}{s^{(0)}(\theta,t)}\left\{dN_{i0}(t) - Y_{i0}(t)\frac{E(dN_i(t))}{s^{(0)}(\theta,t)}\right\}
\end{aligned}
$$

where $s^{(1)}(\theta,t) = e^\theta\mu(t)E\{Y_i(t)\}$ and $s^{(0)}(\theta,t) = \{e^\theta\mu(t) + 1 - \mu(t)\}E\{Y_i(t)\}$. In addition, similar to the proof of Lemma S1, we can show that $\widehat{\beta}_j(\widehat{\theta}_{\mathrm{L}}) \xrightarrow{p} \beta_j(\theta)$ for $j=0,1$. Thus,

$$
\widehat{U}_{\mathrm{CL}}(\theta) = n^{-1}\sum_{i=1}^n \left[I_i\{O_{i1}(\theta) - (X_i-\bar{X})^\top\beta_1(\theta)\} - (1-I_i)\{O_{i0}(\theta) - (X_i-\bar{X})^\top\beta_0(\theta)\}\right]
$$

$$+ o_p(n^{-1/2}).$$

Then as $E\{O_{i1}(\theta)\} = E\{O_{i0}(\theta)\} = 0$, we have

$$\sqrt{n}\widehat{U}_{\mathrm{CL}}(\theta) \xrightarrow{d} N(0, \sigma^2_{\mathrm{CL}}(\theta)),$$

where

$$\begin{aligned}
\sigma^2_{\mathrm{CL}}(\theta) &= \pi \mathrm{var}(O_{i1}(\theta) - X_i^\top \beta_1(\theta)) + (1-\pi)\mathrm{var}(O_{i0}(\theta) - X_i^\top \beta_0(\theta)) \\
&\quad + (\pi\beta_1(\theta) - (1-\pi)\beta_0(\theta))^\top \Sigma_X (\pi\beta_1(\theta) - (1-\pi)\beta_0(\theta)) \\
&= \pi \mathrm{var}(O_{i1}(\theta)) + (1-\pi)\mathrm{var}(O_{i0}(\theta)) - \pi(1-\pi)(\beta_1(\theta) + \beta_0(\theta))^\top \Sigma_X (\beta_1(\theta) + \beta_0(\theta)).
\end{aligned}$$

For (S6), note that

$$\begin{aligned}
-\frac{\partial \widehat{U}_{\mathrm{CL}}(\vartheta)}{\partial \vartheta}\Big|_{\vartheta=\bar{\vartheta}} = -\frac{\partial \widehat{U}_{\mathrm{L}}(\vartheta)}{\partial \vartheta}\Big|_{\vartheta=\bar{\vartheta}} &= n^{-1}\sum_{i=1}^n \int_0^\tau \left\{\frac{S^{(1)}(\bar{\vartheta},t)}{S^{(0)}(\bar{\vartheta},t)}\right\} - \left\{\frac{S^{(1)}(\bar{\vartheta},t)}{S^{(0)}(\bar{\vartheta},t)}\right\}^2 dN_i(t) \\
&\xrightarrow{p} \int_0^\tau \left\{\frac{s^{(1)}(\theta,t)}{s^{(0)}(\theta,t)} - \frac{s^{(1)}(\theta,t)^2}{s^{(0)}(\theta,t)^2}\right\} s^{(0)}(\theta,t)dt.
\end{aligned}$$

It remains to verify that

$$\pi \mathrm{var}(O_{i1}(\theta)) + (1-\pi)\mathrm{var}(O_{i0}(\theta)) = \int_0^\tau \left\{\frac{s^{(1)}(\theta,t)}{s^{(0)}(\theta,t)} - \frac{s^{(1)}(\theta,t)^2}{s^{(0)}(\theta,t)^2}\right\} s^{(0)}(\theta,t)dt,$$

which is easy to show from $E(O_{i1}(\theta)) = 0$ and

$$\begin{aligned}
\mathrm{var}(O_{i1}(\theta)) = E\{O^2_{i1}(\theta)\} &= \int_0^\tau \left\{1 - \frac{s^{(1)}(\theta,t)}{s^{(0)}(\theta,t)}\right\}^2 E\{dN_{i1}(t)\}, \\
\mathrm{var}(O_{i0}(\theta)) = E\{O^2_{i0}(\theta)\} &= \int_0^\tau \left\{\frac{s^{(1)}(\theta,t)}{s^{(0)}(\theta,t)}\right\}^2 E\{dN_{i0}(t)\}.
\end{aligned}$$

PROOF OF THEOREM 2.
Part (a) is from Theorem S1(a) with $b_0 = b_1 = 0$. For part (b), $\theta_1 = \theta_0 = 0$ is proved in Theorem 1(b), $\widehat{\sigma}^2_{\mathrm{L}} \xrightarrow{p} \sigma^2_{\mathrm{L}}$ is proved in Ye & Shao (2020). For part (c), note that $\widehat{\sigma}^2_{\mathrm{L}} = \sigma^2_{\mathrm{L}} + o_p(1)$ under the local alternative (Lemmas S3-S4). Hence,

$$\begin{aligned}
\mathcal{T}_{\mathrm{L}} - \frac{\pi c_1 - (1-\pi)c_0}{\sigma_{\mathrm{L}}} &= \frac{\sqrt{n}\widehat{U}_{\mathrm{L}}}{\widehat{\sigma}_{\mathrm{L}}} - \frac{\pi c_1 - (1-\pi)c_0}{\sigma_{\mathrm{L}}} \\
&= \frac{\sqrt{n}\left(\widehat{U}_{\mathrm{L}} - \frac{n_1\theta_1 - n_0\theta_0}{n}\right)}{\widehat{\sigma}_{\mathrm{L}}} + \frac{n_1 c_1 - n_0 c_0}{n\widehat{\sigma}_{\mathrm{L}}} - \frac{\pi c_1 - (1-\pi)c_0}{\sigma_{\mathrm{L}}} \\
&= \frac{\sqrt{n}\left(\widehat{U}_{\mathrm{L}} - \frac{n_1\theta_1 - n_0\theta_0}{n}\right)}{\sigma_{\mathrm{L}}} + o_p(1) \\
&\xrightarrow{d} N\left(0, \frac{\sigma^2_{\mathrm{L}}(\nu)}{\sigma^2_{\mathrm{L}}}\right).
\end{aligned}$$

PROOF OF THEOREM 3.

(a) From linearizing $\widehat{U}_{SL}$ (Ye & Shao, 2020), we have that

$$\begin{aligned}&\widehat{U}_{\mathrm{SL}} - \sum_z \left(\frac{n_{z1}}{n}\theta_{z1} - \frac{n_{z0}}{n}\theta_{z0}\right)\\ &= \frac{1}{n}\sum_z \sum_{i:Z_i=z} I_i\left(O_{zi1}-\theta_{z1}\right) - (1-I_i)\left(O_{zi0}-\theta_{z0}\right) + o_p(n^{-1/2}).\end{aligned}$$

Following similar steps as in the proof of Theorem 1, we have that

$$\sqrt{n}\left\{\widehat{U}_{\mathrm{SL}} - \sum_z \left(\frac{n_{z1}}{n}\theta_{z1} - \frac{n_{z0}}{n}\theta_{z0}\right)\right\} \xrightarrow{d} N\left(0, \sigma^2_{\mathrm{SL}}\right).$$

For the calibrated stratified log-rank test $\widehat{U}_{CSL}$, from the linearization of $\widehat{U}_{\mathrm{SL}}$, $n^{-1}\sum_z\sum_{i:Z_i=z} I_i(X_i-\bar{X}_z) = O_p(n^{-1/2})$, and $\bar{X}_z - E(X_i \mid Z_i = z) = O_p(n^{-1/2})$, we have

$$\begin{aligned}&\widehat{U}_{\mathrm{CSL}} - \sum_z \left(\frac{n_{z1}}{n}\theta_{z1} - \frac{n_{z0}}{n}\theta_{z0}\right)\\ &= \sum_z \frac{1}{n}\sum_{i=1}^n I_i\mathcal{I}(Z_i=z)\Big(O_{zi1}-\theta_{z1}-(X_i-E(X_i\mid Z_i=z))^\top\gamma_1\Big)\\ &\quad - \sum_z \frac{1}{n}\sum_{i=1}^n (1-I_i)\mathcal{I}(Z_i=z)\Big(O_{zi0}-\theta_{z0}-(X_i-E(X_i\mid Z_i=z))^\top\gamma_0\Big)\\ &\quad + \sum_z \mathrm{pr}(Z=z)(\pi\gamma_1-(1-\pi)\gamma_0)^\top(\bar{X}_z - E(X_i\mid Z_i=z)) + o_p(n^{-1/2}).\\ &:= B_1 - B_2 + B_3 + o_p(n^{-1/2}). \end{aligned} \tag{S7}$$

Next, we show that $\sqrt{n}(B_1-B_2+B_3)$ is asymptotically normal. Let $\mu_{Xz} = E(X_i \mid Z_i = z)$. Consider the random vector

$$\sqrt{n}\begin{pmatrix}\left(E_n\big[I_i\mathcal{I}(Z_i=z)(O_{zi1}-\theta_{z1}-(X_i-\mu_{Xz})^\top\gamma_1)\big], z\in\mathcal{Z}\right)^\top\\ \left(E_n\big[(1-I_i)\mathcal{I}(Z_i=z)(O_{zi0}-\theta_{z0}-(X_i-\mu_{Xz})^\top\gamma_0)\big], z\in\mathcal{Z}\right)^\top\\ \left(E_n\big[\mathcal{I}(Z_i=z)(X_i-\mu_{Xz})\big], z\in\mathcal{Z}\right)^\top\end{pmatrix}. \tag{S8}$$

Conditional on $\mathcal{I},\mathcal{S}$, every component in (S8) is an average of independent terms. From Lindeberg's Central Limit Theorem, as $n\to\infty$, (S8) is asymptotically normal with mean 0 conditional on $\mathcal{I},\mathcal{S}$. This implies that $\sqrt{n}(B_1-B_2+B_3)$ is asymptotically normal with mean 0 conditional on $\mathcal{I},\mathcal{S}$. Then, we calculate its variance. Note that

$$\begin{aligned}&\mathrm{var}(\sqrt{n}(B_1-B_2)\mid\mathcal{I},\mathcal{S})\\ &= \sum_z\frac{1}{n}\sum_{i=1}^n I_i\mathcal{I}(Z_i=z)\mathrm{var}(O_{zi1}-(X_i-\mu_{Xz})^\top\gamma_1\mid Z_i=z)\\ &\quad + \sum_z\frac{1}{n}\sum_{i=1}^n (1-I_i)\mathcal{I}(Z_i=z)\mathrm{var}(O_{zi0}-(X_i-\mu_{Xz})^\top\gamma_0\mid Z_i=z)\\ &= \pi\sum_z P(Z_i=z)\mathrm{var}(O_{zi1}-(X_i-\mu_{Xz})^\top\gamma_1\mid Z_i=z)\\ &\quad + (1-\pi)\sum_z P(Z_i=z)\mathrm{var}(O_{zi0}-(X_i-\mu_{Xz})^\top\gamma_0\mid Z_i=z) + o_p(1)\end{aligned}$$

$$
\begin{aligned}
&= \pi \sum_z P(Z_i = z)\big\{\mathrm{var}(O_{zi1} \mid Z_i = z) + \gamma_1^\top \mathrm{var}(X_i \mid Z_i = z)\gamma_1 - 2\gamma_1^\top \mathrm{cov}(X_i, O_{zi1} \mid Z_i = z)\big\} \\
&\quad + (1-\pi) \sum_z P(Z_i = z)\big\{\mathrm{var}(O_{zi0} \mid Z_i = z) + \gamma_0^\top \mathrm{var}(X_i \mid Z_i = z)\gamma_0 \\
&\quad - 2\gamma_0^\top \mathrm{cov}(X_i, O_{zi0} \mid Z_i = z)\big\} + o_p(1) \\
&= \sigma^2_{\mathrm{SL}} - \pi\gamma_1^\top E\{\mathrm{var}(X_i \mid Z_i)\}\gamma_1 - (1-\pi)\gamma_0^\top E\{\mathrm{var}(X_i \mid Z_i)\}\gamma_0 + o_p(1),
\end{aligned}
$$

where the last line is because

$$
\gamma_j = \left\{ \sum_z P(Z_i = z)\mathrm{var}(X_i \mid Z_i = z)\right\}^{-1} \left\{ \sum_z P(Z_i = z)\mathrm{cov}(X_i, O_{zij} \mid Z_i = z)\right\}
$$

for $j = 0, 1$, from Lemma S1. Moreover,

$$
\begin{aligned}
&\mathrm{var}(\sqrt{n}B_3 \mid \mathcal{I}, \mathcal{S}) \\
&= (\pi\gamma_1 - (1-\pi)\gamma_0)^\top E\{\mathrm{var}(X_i \mid Z_i)\}(\pi\gamma_1 - (1-\pi)\gamma_0) + o_p(1).
\end{aligned}
$$

Next, we can calculate the covariance between $\sqrt{n}B_1$ and $\sqrt{n}B_3$ when conditional on $\mathcal{I}, \mathcal{S}$ as

$$
\begin{aligned}
&\mathrm{cov}(\sqrt{n}B_3, \sqrt{n}B_1 \mid \mathcal{I}, \mathcal{S}) \\
&= n\mathrm{cov}\left(\sum_z \Pr(Z = z)(\pi\gamma_1 - (1-\pi)\gamma_0)^\top (\bar{X}_z - \mu_{Xz}), \right. \\
&\quad \left. \sum_z \frac{1}{n} \sum_{i=1}^n I_i \mathcal{I}(Z_i = z)(O_{zi1} - (X_i - \mu_{Xz})^\top \gamma_1) \mid \mathcal{I}, \mathcal{S} \right) \\
&= \sum_z \sum_{i=1}^n I_i \mathcal{I}(Z_i = z) \Pr(Z = z)(\pi\gamma_1 - (1-\pi)\gamma_0)^\top \mathrm{cov}(\bar{X}_z - \mu_{Xz}, O_{zi1} - (X_i - \mu_{Xz})^\top \gamma_1 \mid \mathcal{I}, \mathcal{S}) \\
&= \sum_z \sum_{i=1}^n I_i \mathcal{I}(Z_i = z) \Pr(Z = z)(\pi\gamma_1 - (1-\pi)\gamma_0)^\top \frac{1}{n_z} \mathrm{cov}(X_i - \mu_{Xz}, O_{zi1} - (X_i - \mu_{Xz})^\top \gamma_1 \mid \mathcal{I}, \mathcal{S}) \\
&= \sum_z \pi \Pr(Z_i = z)(\pi\gamma_1 - (1-\pi)\gamma_0)^\top \left\{\mathrm{cov}(X_i, O_{zi1} \mid Z_i) - \mathrm{var}(X_i \mid Z_i)\gamma_1\right\} + o_p(1) \\
&= o_p(1)
\end{aligned}
$$

Similarly, $\mathrm{cov}(\sqrt{n}B_3, \sqrt{n}B_2 \mid \mathcal{I}, \mathcal{S}) = o_p(1)$. Hence, $\mathrm{cov}(\sqrt{n}B_3, \sqrt{n}(B_1 - B_2) \mid \mathcal{I}, \mathcal{S}) = o_p(1)$. Combining all the above derivations, we have that

$$
\begin{aligned}
&\mathrm{var}\{\sqrt{n}(B_1 - B_2 + B_3) \mid \mathcal{I}, \mathcal{S}\} \\
&= \sigma^2_{\mathrm{SL}} - \pi(1-\pi)(\gamma_1 + \gamma_0)^\top E\{\mathrm{var}(X_i \mid Z_i)\}(\gamma_1 + \gamma_0) + o_p(1).
\end{aligned}
$$

The asymptotic distribution then follows from the Slutsky's theorem and bounded convergence theorem.

(b) It is straightforward to show the assumed conditions imply $\theta_{z1} = \theta_{z0} = 0$ for any z from applying Lemma S2 separately for every stratum z.

Next, under H_0, $\widehat{\sigma}^2_{\mathrm{SL}} \xrightarrow{p} \sigma^2_{\mathrm{SL}}$ is proved in Ye & Shao (2020), and $\widehat{\sigma}^2_{CSL} \xrightarrow{p} \sigma^2_{CSL}$ is from $\widehat{\gamma}_j = \gamma_j + o_p(1), j = 0, 1$ from Lemma S1, $\widehat{\Sigma}_{X|z} = \mathrm{var}(X_i \mid Z_i = z) + o_p(1)$, and $n_z/n = P(Z_i = z) + o_p(1)$. The result that $\mathcal{T}_{\mathrm{CL}} \xrightarrow{d} N(0, 1)$ and validity of $\mathcal{T}_{\mathrm{CL}}$ follows from Slutsky theorem.

(c) Under the local alternative, applying Lemmas S3-S4 within each stratum $Z_i = z$ gives $\widehat{\sigma}^2_{\mathrm{SL}} = \sigma^2_{\mathrm{SL}} + o_p(1)$. In addition, from Lemma S1 and $\widehat{\Sigma}_{X|z} = \mathrm{var}(X_i \mid Z_i = z) + o_p(1)$, we have $\widehat{\sigma}^2_{\mathrm{C}SL} = \sigma^2_{\mathrm{C}SL} + o_p(1)$. Then,

$$\begin{aligned}
&\mathcal{T}_{\mathrm{SL}} - \frac{\sum_z \Pr(Z=z)\{\pi c_{z1} - (1-\pi)c_{z0}\}}{\sigma_{\mathrm{SL}}} \\
&= \frac{\sqrt{n}\widehat{U}_{\mathrm{SL}}}{\widehat{\sigma}_{\mathrm{SL}}} - \frac{\sum_z \Pr(Z=z)\{\pi c_{z1} - (1-\pi)c_{z0}\}}{\sigma_{\mathrm{SL}}} \\
&= \frac{\sqrt{n}\left(\widehat{U}_{\mathrm{SL}} - \sum_z \frac{n_{z1}\theta_{z1} - n_{z0}\theta_{z0}}{n}\right)}{\widehat{\sigma}_{\mathrm{SL}}} + \sum_z \frac{n_{z1}c_{z1} - n_{z0}c_{z0}}{n\widehat{\sigma}_{\mathrm{SL}}} - \frac{\sum_z \Pr(Z=z)\{\pi c_{z1} - (1-\pi)c_{z0}\}}{\sigma_{\mathrm{SL}}} \\
&= \frac{\sqrt{n}\left(\widehat{U}_{\mathrm{SL}} - \sum_z \frac{n_{z1}\theta_{z1} - n_{z0}\theta_{z0}}{n}\right)}{\widehat{\sigma}_{\mathrm{SL}}} + o_p(1) \\
&\xrightarrow{d} N(0,1),
\end{aligned}$$

and

$$\begin{aligned}
&\mathcal{T}_{CSL} - \frac{\sum_z \Pr(Z=z)\{\pi c_{z1} - (1-\pi)c_{z0}\}}{\sigma_{\mathrm{C}SL}} \\
&= \frac{\sqrt{n}\widehat{U}_{\mathrm{C}SL}}{\widehat{\sigma}_{\mathrm{C}SL}} - \frac{\sum_z \Pr(Z=z)\{\pi c_{z1} - (1-\pi)c_{z0}\}}{\sigma_{\mathrm{C}SL}} \\
&= \frac{\sqrt{n}\left(\widehat{U}_{\mathrm{C}SL} - \sum_z \frac{n_{z1}\theta_{z1} - n_{z0}\theta_{z0}}{n}\right)}{\widehat{\sigma}_{\mathrm{C}SL}} + \sum_z \frac{n_{z1}c_{z1} - n_{z0}c_{z0}}{n\widehat{\sigma}_{\mathrm{C}SL}} - \frac{\sum_z \Pr(Z=z)\{\pi c_{z1} - (1-\pi)c_{z0}\}}{\sigma_{\mathrm{C}SL}} \\
&= \frac{\sqrt{n}\left(\widehat{U}_{\mathrm{C}SL} - \sum_z \frac{n_{z1}\theta_{z1} - n_{z0}\theta_{z0}}{n}\right)}{\widehat{\sigma}_{\mathrm{C}SL}} + o_p(1) \\
&\xrightarrow{d} N(0,1).
\end{aligned}$$

REFERENCES

ANDERSEN, P. K. & GILL, R. D. (1982). Cox's regression model for counting processes: A large sample study. *Annals of Statistics* **10**, 1100–1120.

JIANG, H., SYMANOWSKI, J., PAUL, S., QU, Y., ZAGAR, A. & HONG, S. (2008). The type i error and power of non-parametric logrank and wilcoxon tests with adjustment for covariates–a simulation study. *Statistics in Medicine* **27**, 5850–5860.

LIN, D. Y. & WEI, L. J. (1989). The robust inference for the cox proportional hazards model. *Journal of the American Statistical Association* **84**, 1074–1078.

TANGEN, C. M. & KOCH, G. G. (1999). Nonparametric analysis of covariance for hypothesis testing with logrank and wilcoxon scores and survival-rate estimation in a randomized clinical trial. *Journal of Biopharmaceutical Statistics* **9**, 307–338.

YE, T. & SHAO, J. (2020). Robust tests for treatment effect in survival analysis under covariate-adaptive randomization. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **82**, 1301–1323.

YE, T., SHAO, J., YI, Y. & ZHAO, Q. (2022). Toward better practice of covariate adjustment in analyzing randomized clinical trials. *Journal of the American Statistical Association* .