Supplement to
"Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique"
(Proofs of theorems, corollaries, and selected remarks)

Winston Lin*

1. Additional notation and definitions. In the main paper:

- Section 2 defines the basic notation;
- Section 4.2.1 states Conditions 1–3;
- Section 4.2.2 defines the vectors $\mathbf{Q}_a$ and $\mathbf{Q}_b$ and the prediction errors a_i^* and b_i^*, and introduces the σ_x^2 and $\sigma_{x,y}$ notation for population variances and covariances;
- Section 5 defines the vector $\mathbf{Q}$ and the prediction errors a_i^{**} and b_i^{**}.

Let $\tilde{p}_A = n_A/n$ [as in remark (iii) after Corollary 1.2].

Extend Section 2's notation for population and group means to cover any scalar, vector, or matrix expression. For example:

$$\overline{ab}_A = \frac{1}{n_A}\sum_{i\in A} a_i b_i, \qquad \overline{a\mathbf{z}}_A = \frac{1}{n_A}\sum_{i\in A} a_i \mathbf{z}_i, \qquad \overline{\mathbf{z}'\mathbf{z}}_A = \frac{1}{n_A}\sum_{i\in A} \mathbf{z}_i'\mathbf{z}_i.$$

Extend Freedman's (2008b) angle bracket notation to cover all the finite limits assumed in Condition 2. For example:

$$\langle a\mathbf{z}\rangle = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} a_i\mathbf{z}_i, \qquad \langle \mathbf{z}'\mathbf{z}\rangle = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} \mathbf{z}_i'\mathbf{z}_i.$$

(The second limit exists since it is a submatrix of $\lim_{n\to\infty} n^{-1}\mathbf{Z}'\mathbf{Z}$.)

Condition 4 (centering) will sometimes be assumed for convenience. The proofs will explain why this can be done without loss of generality.

CONDITION 4. The population means of the potential outcomes and the covariates are zero: $\overline{a} = \overline{b} = 0$ and $\overline{\mathbf{z}} = \mathbf{0}$.

Some transformations of the regressors will be useful in the proofs. Define the pooled-slopes regression estimator of mean potential outcomes, $\widehat{\boldsymbol{\beta}}_{\text{adj}}$, as the 2×1 vector containing the estimated coefficients on T_i and $1-T_i$ from the no-intercept OLS regression of Y_i on T_i, $1-T_i$, and $\mathbf{z}_i - \overline{\mathbf{z}}$. Let $\widehat{\mathbf{Q}}$ denote the vector of estimated coefficients on $\mathbf{z}_i - \overline{\mathbf{z}}$ from the same regression.

*Ph.D. student, Dept. of Statistics, UC Berkeley (Linston@gmail.com).

The vector $\widehat{\boldsymbol{\beta}}_{\text{adj}}$ is an estimate of $\boldsymbol{\beta} = (\overline{a}, \overline{b})'$. By well-known invariance properties of least squares, $\widehat{ATE}_{\text{adj}}$ is the difference between the two elements of $\widehat{\boldsymbol{\beta}}_{\text{adj}}$.

Similarly, define the separate-slopes regression estimator of mean potential outcomes, $\widehat{\boldsymbol{\beta}}_{\text{interact}}$, as the 2×1 vector containing the estimated coefficients on T_i and $1 - T_i$ from the no-intercept OLS regression of Y_i on T_i, $1 - T_i$, $\mathbf{z}_i - \overline{\mathbf{z}}$, and $T_i(\mathbf{z}_i - \overline{\mathbf{z}})$. Then $\widehat{ATE}_{\text{interact}}$ is the difference between the two elements of $\widehat{\boldsymbol{\beta}}_{\text{interact}}$.

Let $\widehat{\mathbf{Q}}_a$ and $\widehat{\mathbf{Q}}_b$ denote the vectors of estimated coefficients on $\mathbf{z}_i$ in the OLS regressions of Y_i on $\mathbf{z}_i$ in groups A and B, respectively.

Conditions 1–3 do not rule out the possibility that under some realizations of random assignment, the regressors are perfectly collinear. The probability of this event converges to zero by Conditions 2 and 3, so it is irrelevant to the asymptotic results. For concreteness, whenever $\widehat{ATE}_{\text{adj}}$ cannot be computed because of collinearity, let $\widehat{ATE}_{\text{adj}} = \overline{Y}_A - \overline{Y}_B$, $\widehat{\mathbf{Q}} = \mathbf{0}$, and $\widehat{\boldsymbol{\beta}}_{\text{adj}} = (\overline{Y}_A, \overline{Y}_B)'$; whenever $\widehat{ATE}_{\text{interact}}$ cannot be computed, let $\widehat{ATE}_{\text{interact}} = \overline{Y}_A - \overline{Y}_B$, $\widehat{\mathbf{Q}}_a = \mathbf{0}$, $\widehat{\mathbf{Q}}_b = \mathbf{0}$, and $\widehat{\boldsymbol{\beta}}_{\text{interact}} = (\overline{Y}_A, \overline{Y}_B)'$. Other arbitrary values could be used.

2. Lemmas. Lemma 1 is a finite-population version of the Weak Law of Large Numbers.

LEMMA 1. *Assume Conditions 1–3. The means over group A or group B of a_i, b_i, $\mathbf{z}_i$, a_i^2, b_i^2, $\mathbf{z}_i'\mathbf{z}_i$, a_ib_i, $a_i\mathbf{z}_i$, and $b_i\mathbf{z}_i$ converge in probability to the limits of the population means. For example:*

$$\begin{aligned} \overline{a}_A &\xrightarrow{p} \langle a \rangle, \\ \overline{a^2}_A \equiv \frac{1}{n_A}\sum_{i \in A} a_i^2 &\xrightarrow{p} \langle a^2 \rangle, \\ \overline{ab}_A &\xrightarrow{p} \langle ab \rangle, \\ \overline{a\mathbf{z}}_A &\xrightarrow{p} \langle a\mathbf{z} \rangle, \\ \overline{\mathbf{z}'\mathbf{z}}_A &\xrightarrow{p} \langle \mathbf{z}'\mathbf{z} \rangle. \end{aligned}$$

PROOF. From basic results on simple random sampling [e.g., Freedman's (2008b) Proposition 1], $E(\overline{a}_A) = \overline{a}$ and

$$\operatorname{var}(\overline{a}_A) = \frac{1}{n-1}\frac{1 - \tilde{p}_A}{\tilde{p}_A}\sigma_a^2.$$

As $n \to \infty$, $\tilde{p}_A \to p_A > 0$ and $\sigma_a^2 \to \langle a^2 \rangle - \langle a \rangle^2$, so $\operatorname{var}(\overline{a}_A) \to 0$. By Chebyshev's inequality, $\overline{a}_A - \overline{a} \xrightarrow{p} 0$. Therefore,

$$\overline{a}_A \xrightarrow{p} \lim_{n \to \infty} \overline{a} = \langle a \rangle.$$

The proofs that $\overline{a^2}_A \xrightarrow{p} \langle a^2 \rangle$ and $\overline{ab}_A \xrightarrow{p} \langle ab \rangle$ are similar but rely on Condition 1 to show that $\text{var}(\overline{a^2}_A) \to 0$ and $\text{var}(\overline{ab}_A) \to 0$. First note that

$$\text{var}(\overline{a^2}_A) = \frac{1}{n-1} \frac{1 - \tilde{p}_A}{\tilde{p}_A} \sigma^2_{(a^2)}$$

and

$$\text{var}(\overline{ab}_A) = \frac{1}{n-1} \frac{1 - \tilde{p}_A}{\tilde{p}_A} \sigma^2_{(ab)}.$$

By Condition 1, $\sigma^2_{(a^2)}$ is bounded:

$$\sigma^2_{(a^2)} \leq \overline{a^4} < L.$$

Therefore, $\text{var}(\overline{a^2}_A) \to 0$. Next note that $\sigma^2_{(ab)}$ is bounded, using the Cauchy–Schwarz inequality:

$$\sigma^2_{(ab)} \leq \frac{1}{n} \sum_{i=1}^{n} a_i^2 b_i^2 \leq \left(\frac{1}{n} \sum_{i=1}^{n} a_i^4 \right)^{1/2} \left(\frac{1}{n} \sum_{i=1}^{n} b_i^4 \right)^{1/2} < L.$$

Therefore, $\text{var}(\overline{ab}_A) \to 0$.

The same logic can be used to show the remaining results. Those involving $\mathbf{z}_i$ can be proved element by element. □

LEMMA 2. *The pooled-slopes estimator of mean potential outcomes is*

$$\widehat{\beta}_{\text{adj}} = \left[\overline{Y}_A - (\overline{\mathbf{z}}_A - \overline{\mathbf{z}}) \widehat{\mathbf{Q}}, \overline{Y}_B - (\overline{\mathbf{z}}_B - \overline{\mathbf{z}}) \widehat{\mathbf{Q}} \right]'.$$

PROOF. The residuals from the regression defining $\widehat{\beta}_{\text{adj}}$ are uncorrelated with T_i and $1 - T_i$. Therefore, the regression line passes through the points of means within groups A and B, and the result follows. □

LEMMA 3. *The separate-slopes estimator of mean potential outcomes is*

$$\widehat{\beta}_{\text{interact}} = \left[\overline{Y}_A - (\overline{\mathbf{z}}_A - \overline{\mathbf{z}}) \widehat{\mathbf{Q}}_a, \overline{Y}_B - (\overline{\mathbf{z}}_B - \overline{\mathbf{z}}) \widehat{\mathbf{Q}}_b \right]'.$$

PROOF. In the regression defining $\widehat{\beta}_{\text{interact}}$, the coefficient on $\mathbf{z}_i - \overline{\mathbf{z}}$ is $\widehat{\mathbf{Q}}_b$ and the coefficient on $T_i(\mathbf{z}_i - \overline{\mathbf{z}})$ is $\widehat{\mathbf{Q}}_a - \widehat{\mathbf{Q}}_b$. (This can be shown from the equivalence of the minimization problems.) The rest of the proof is similar to that of Lemma 2. □

LEMMA 4. *Assume Conditions 1–3. Then* $\widehat{\mathbf{Q}} \xrightarrow{p} \mathbf{Q}$.

PROOF. We can assume Condition 4 without loss of generality: Let $\widehat{\gamma}$ be the estimated coefficient vector from a no-intercept OLS regression of Y_i on T_i, $1-T_i$, and $\mathbf{z}_i - \overline{\mathbf{z}}$. Let $\tilde{a}_i = a_i - \overline{a}$ and $\tilde{b}_i = b_i - \overline{b}$, so that Condition 4 holds for $\tilde{a}_i$ and $\tilde{b}_i$. Let $\tilde{Y}_i = \tilde{a}_i T_i + \tilde{b}_i(1-T_i)$. By a well-known property of OLS [e.g., Freedman's (2008b) Lemma A.1], the estimated coefficient vector from a no-intercept OLS regression of $\tilde{Y}_i$ on T_i, $1-T_i$, and $\mathbf{z}_i - \overline{\mathbf{z}}$ is $\widehat{\gamma} - (\overline{a}, \overline{b}, 0)'$, so $\widehat{\mathbf{Q}}$ is unchanged. Similarly, $\mathbf{Q}$ is unchanged. Finally, centering $\mathbf{z}_i$ has no effect on the slope vectors $\widehat{\mathbf{Q}}$ and $\mathbf{Q}$.

By the Frisch–Waugh–Lovell theorem, $\widehat{\mathbf{Q}}$ can be computed from auxiliary regressions: Let

$$\begin{aligned} e_i &= Y_i - \overline{Y}_A T_i - \overline{Y}_B(1-T_i), \\ \mathbf{f}_i &= \mathbf{z}_i - \overline{\mathbf{z}}_A T_i - \overline{\mathbf{z}}_B(1-T_i). \end{aligned}$$

Then

$$\widehat{\mathbf{Q}} = \left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{f}_i'\mathbf{f}_i\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{f}_i' e_i\right).$$

Some algebra yields

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{f}_i'\mathbf{f}_i = \overline{\mathbf{z}'\mathbf{z}} - \tilde{p}_A \overline{\mathbf{z}}_A' \overline{\mathbf{z}}_A - (1-\tilde{p}_A)\overline{\mathbf{z}}_B'\overline{\mathbf{z}}_B.$$

By Condition 4 and Lemma 1, $\overline{\mathbf{z}}_A \xrightarrow{p} \mathbf{0}$ and $\overline{\mathbf{z}}_B \xrightarrow{p} \mathbf{0}$. Therefore,

$$\frac{1}{n}\sum_{i=1}^{n} \mathbf{f}_i'\mathbf{f}_i \xrightarrow{p} \langle \mathbf{z}'\mathbf{z} \rangle.$$

Now note that

$$\begin{aligned} e_i &= (a_i - \overline{a}_A)T_i + (b_i - \overline{b}_B)(1-T_i), \\ \mathbf{f}_i &= (\mathbf{z}_i - \overline{\mathbf{z}}_A)T_i + (\mathbf{z}_i - \overline{\mathbf{z}}_B)(1-T_i). \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} \mathbf{f}_i' e_i &= \frac{1}{n}\sum_{i \in A} (\mathbf{z}_i - \overline{\mathbf{z}}_A)'(a_i - \overline{a}_A) + \frac{1}{n}\sum_{i \in B} (\mathbf{z}_i - \overline{\mathbf{z}}_B)'(b_i - \overline{b}_B) \\ &= \tilde{p}_A(\overline{a\mathbf{z}}_A - \overline{a}_A \overline{\mathbf{z}}_A)' + (1-\tilde{p}_A)(\overline{b\mathbf{z}}_B - \overline{b}_B\overline{\mathbf{z}}_B)' \\ &\xrightarrow{p} p_A\langle a\mathbf{z}\rangle' + (1-p_A)\langle b\mathbf{z}\rangle'. \end{aligned}$$

(Convergence to the last expression follows from Lemma 1 and Conditions 3–4.)

It follows that

$$\begin{aligned}\widehat{\mathbf{Q}} &\xrightarrow{p} \langle \mathbf{z}'\mathbf{z}\rangle^{-1}\left[p_A\langle a\mathbf{z}\rangle' + (1-p_A)\langle b\mathbf{z}\rangle'\right] \\ &= p_A \lim_{n\to\infty}\left[\left(\sum_{i=1}^n \mathbf{z}_i'\mathbf{z}_i\right)^{-1}\sum_{i=1}^n \mathbf{z}_i' a_i\right] + (1-p_A)\lim_{n\to\infty}\left[\left(\sum_{i=1}^n \mathbf{z}_i'\mathbf{z}_i\right)^{-1}\sum_{i=1}^n \mathbf{z}_i' b_i\right] \\ &= p_A\mathbf{Q}_a + (1-p_A)\mathbf{Q}_b = \mathbf{Q}.\end{aligned}$$

□

LEMMA 5. *Assume Conditions 1–3. Then* $\widehat{\mathbf{Q}}_a \xrightarrow{p} \mathbf{Q}_a$ *and* $\widehat{\mathbf{Q}}_b \xrightarrow{p} \mathbf{Q}_b$.

PROOF. The proof is similar to that of Lemma 4 but simpler. Again, we can assume Condition 4 without loss of generality. By the Frisch–Waugh–Lovell theorem,

$$\widehat{\mathbf{Q}}_a = \left[\frac{1}{n_A}\sum_{i\in A}(\mathbf{z}_i - \overline{\mathbf{z}}_A)'(\mathbf{z}_i - \overline{\mathbf{z}}_A)\right]^{-1}\left[\frac{1}{n_A}\sum_{i\in A}(\mathbf{z}_i - \overline{\mathbf{z}}_A)'(a_i - \overline{a}_A)\right].$$

Some algebra, Lemma 1, and Condition 4 yield

$$\frac{1}{n_A}\sum_{i\in A}(\mathbf{z}_i - \overline{\mathbf{z}}_A)'(\mathbf{z}_i - \overline{\mathbf{z}}_A) = \overline{\mathbf{z}'\mathbf{z}}_A - \overline{\mathbf{z}}_A'\overline{\mathbf{z}}_A \xrightarrow{p} \langle \mathbf{z}'\mathbf{z}\rangle$$

and

$$\frac{1}{n_A}\sum_{i\in A}(\mathbf{z}_i - \overline{\mathbf{z}}_A)'(a_i - \overline{a}_A) = (\overline{a\mathbf{z}}_A - \overline{a}_A\overline{\mathbf{z}}_A)' \xrightarrow{p} \langle a\mathbf{z}\rangle'$$

so

$$\begin{aligned}\widehat{\mathbf{Q}}_a &\xrightarrow{p} \langle \mathbf{z}'\mathbf{z}\rangle^{-1}\langle a\mathbf{z}\rangle' \\ &= \lim_{n\to\infty}\left[\left(\sum_{i=1}^n \mathbf{z}_i'\mathbf{z}_i\right)^{-1}\sum_{i=1}^n \mathbf{z}_i' a_i\right] = \mathbf{Q}_a.\end{aligned}$$

The proof that $\widehat{\mathbf{Q}}_b \xrightarrow{p} \mathbf{Q}_b$ is similar. □

Lemma 6 is similar to part of Freedman's (2008b) Theorem 2.

LEMMA 6. *Assume Conditions 1–3. Then*

$$\sqrt{n}(\widehat{\boldsymbol{\beta}}_{\text{adj}} - \boldsymbol{\beta}) \xrightarrow{d} N(0, \mathbf{V})$$

where

$$\mathbf{V} \quad = \quad \begin{bmatrix} \frac{1-p_A}{p_A} \lim_{n\to\infty} \sigma^2_{a^{**}} & -\lim_{n\to\infty} \sigma_{a^{**},b^{**}} \\ -\lim_{n\to\infty} \sigma_{a^{**},b^{**}} & \frac{p_A}{1-p_A} \lim_{n\to\infty} \sigma^2_{b^{**}} \end{bmatrix}.$$

PROOF. We can assume Condition 4 without loss of generality: Centering a_i, b_i, and $\mathbf{z}_i$ has no effect on $\widehat{\mathbf{Q}}$ and $\mathbf{Q}$, as shown in the proof of Lemma 4, so it subtracts $(\overline{a}, \overline{b})'$ from both $\widehat{\boldsymbol{\beta}}_{\text{adj}}$ (see Lemma 2) and $\boldsymbol{\beta}$, and it has no effect on the elements of $\mathbf{V}$.

Condition 4 and Lemma 2 imply that

$$\begin{aligned} \sqrt{n}(\widehat{\boldsymbol{\beta}}_{\text{adj}} - \boldsymbol{\beta}) &= \sqrt{n}(\overline{Y}_A - \overline{\mathbf{z}}_A \widehat{\mathbf{Q}}, \overline{Y}_B - \overline{\mathbf{z}}_B \widehat{\mathbf{Q}})' \\ &= \sqrt{n}(\overline{a}_A - \overline{\mathbf{z}}_A \mathbf{Q}, \overline{b}_B - \overline{\mathbf{z}}_B \mathbf{Q})' - [\sqrt{n}\overline{\mathbf{z}}_A(\widehat{\mathbf{Q}} - \mathbf{Q}), \sqrt{n}\overline{\mathbf{z}}_B(\widehat{\mathbf{Q}} - \mathbf{Q})]'. \end{aligned}$$

By a finite-population Central Limit Theorem [Freedman's (2008b) Theorem 1], $\sqrt{n}\overline{\mathbf{z}}_A$ and $\sqrt{n}\overline{\mathbf{z}}_B$ are $O_p(1)$, and by Lemma 4, $\widehat{\mathbf{Q}} - \mathbf{Q}$ is $o_p(1)$. Therefore,

$$[\sqrt{n}\overline{\mathbf{z}}_A(\widehat{\mathbf{Q}} - \mathbf{Q}), \sqrt{n}\overline{\mathbf{z}}_B(\widehat{\mathbf{Q}} - \mathbf{Q})]' \xrightarrow{p} \mathbf{0}.$$

The conclusion follows from Freedman's (2008b) Theorem 1 with a and b replaced by $a - \mathbf{z}\mathbf{Q}$ and $b - \mathbf{z}\mathbf{Q}$. $\square$

Lemma 7 is an application of the Weak Law of Large Numbers (Lemma 1).

LEMMA 7. *Assume Conditions 1–3. Let $\boldsymbol{\theta}$ be any $K \times 1$ vector that is constant as $n \to \infty$. Then*

$$\begin{aligned} \frac{1}{n_A} \sum_{i \in A} (a_i + \mathbf{z}_i \boldsymbol{\theta})^2 &\xrightarrow{p} \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (a_i + \mathbf{z}_i \boldsymbol{\theta})^2, \\ \frac{1}{n - n_A} \sum_{i \in B} (b_i + \mathbf{z}_i \boldsymbol{\theta})^2 &\xrightarrow{p} \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (b_i + \mathbf{z}_i \boldsymbol{\theta})^2. \end{aligned}$$

PROOF. Using Lemma 1,

$$\begin{aligned} \frac{1}{n_A} \sum_{i \in A} (a_i + \mathbf{z}_i \boldsymbol{\theta})^2 &= \overline{a^2}_A + 2\overline{a\mathbf{z}}_A \boldsymbol{\theta} + \boldsymbol{\theta}' \overline{\mathbf{z}'\mathbf{z}}_A \boldsymbol{\theta} \\ &\xrightarrow{p} \langle a^2 \rangle + 2\langle a\mathbf{z} \rangle \boldsymbol{\theta} + \boldsymbol{\theta}' \langle \mathbf{z}'\mathbf{z} \rangle \boldsymbol{\theta} \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (a_i + \mathbf{z}_i \boldsymbol{\theta})^2. \end{aligned}$$

The proof of the other assertion is analogous. $\square$

Lemma 8 shows that the sandwich variance estimator for $\widehat{ATE}_{\text{adj}}$ is invariant to the transformation of the regressors that was used to define $\widehat{\beta}_{\text{adj}}$.

LEMMA 8. *Let*

$$\mathbf{W} \quad = \quad (\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\left(\sum_{i=1}^{n}\hat{e}_i^2\tilde{\mathbf{x}}_i'\tilde{\mathbf{x}}_i\right)(\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}$$

where $\tilde{\mathbf{X}}$ *is the* $n \times (K+2)$ *matrix with row* i *equal to* $\tilde{\mathbf{x}}_i = (T_i, 1-T_i, \mathbf{z}_i - \bar{\mathbf{z}})$ *and* $\hat{e}_i$ *is the residual from the no-intercept OLS regression of* Y_i *on* $\tilde{\mathbf{x}}_i$. *Then* $\widehat{v}_{\text{adj}} = W_{11} + W_{22} - 2W_{12}$, *where* W_{ij} *is the* (i,j) *element of* $\mathbf{W}$.

PROOF. By definition, $\widehat{v}_{\text{adj}}$ is the $(2,2)$ element of

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\text{diag}(\hat{\varepsilon}_1^2,\ldots,\hat{\varepsilon}_n^2)\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \quad = \quad (\mathbf{X}'\mathbf{X})^{-1}\left(\sum_{i=1}^{n}\hat{\varepsilon}_i^2\mathbf{x}_i'\mathbf{x}_i\right)(\mathbf{X}'\mathbf{X})^{-1}$$

where $\mathbf{X}$ is the $n \times (K+2)$ matrix whose ith row is $\mathbf{x}_i = (1, T_i, \mathbf{z}_i)$ and $\hat{\varepsilon}_i$ is the residual from the OLS regression of Y_i on $\mathbf{x}_i$.

The OLS residuals are invariant to the linear transformation of regressors, so $\hat{e}_i = \hat{\varepsilon}_i$ for $i = 1, 2, \ldots, n$. Also, $\mathbf{X} = \tilde{\mathbf{X}}\mathbf{R}\mathbf{S}$ where

$$\mathbf{R} = \begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_K \end{bmatrix}, \qquad \mathbf{S} = \begin{bmatrix} \mathbf{I}_2 & \mathbf{L} \\ \mathbf{0} & \mathbf{I}_K \end{bmatrix},$$

and

$$\mathbf{M} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}^{-1}, \qquad \mathbf{L} = \begin{bmatrix} \bar{\mathbf{z}} \\ \mathbf{0} \end{bmatrix}.$$

Note that $\mathbf{R}$ is symmetric but $\mathbf{S}$ is not, and

$$\mathbf{S}^{-1} \quad = \quad \begin{bmatrix} \mathbf{I}_2 & -\mathbf{L} \\ \mathbf{0} & \mathbf{I}_K \end{bmatrix}.$$

Therefore,

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\text{diag}(\hat{\varepsilon}_1^2,\ldots,\hat{\varepsilon}_n^2)\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \quad = \quad \mathbf{S}^{-1}\mathbf{R}^{-1}\mathbf{W}\mathbf{R}^{-1}(\mathbf{S}^{-1})'.$$

The (2, 2) element is $W_{11} + W_{22} - 2W_{12}$. □

Lemma 9 is important for the proof of Theorem 2.

LEMMA 9. *Assume Conditions 1–4. Let $\hat{e}_i$ denote the residual from the no-intercept OLS regression of Y_i on T_i, $1-T_i$, and $\mathbf{z}_i$. Then*

$$\frac{1}{n_A}\sum_{i\in A}\hat{e}_i^2 \xrightarrow{p} \lim_{n\to\infty}\sigma^2_{a^{**}}, \qquad \frac{1}{n-n_A}\sum_{i\in B}\hat{e}_i^2 \xrightarrow{p} \lim_{n\to\infty}\sigma^2_{b^{**}},$$

and $n^{-1}\sum_{i\in A}\hat{e}_i^2\mathbf{z}_i$, $n^{-1}\sum_{i\in B}\hat{e}_i^2\mathbf{z}_i$, and $n^{-1}\sum_{i=1}^n\hat{e}_i^2\mathbf{z}_i'\mathbf{z}_i$ are all $O_p(1)$.

PROOF. Let $\widehat{\beta}_{\mathrm{adj}(1)}$ and $\widehat{\beta}_{\mathrm{adj}(2)}$ denote the estimated coefficients on T_i and $1-T_i$, respectively. Then

$$\begin{aligned}
\hat{e}_i &= Y_i - \widehat{\beta}_{\mathrm{adj}(1)}T_i - \widehat{\beta}_{\mathrm{adj}(2)}(1-T_i) - \mathbf{z}_i\widehat{\mathbf{Q}} \\
&= T_i[(a_i - \mathbf{z}_i\widehat{\mathbf{Q}}) - \widehat{\beta}_{\mathrm{adj}(1)}] + (1-T_i)[(b_i - \mathbf{z}_i\widehat{\mathbf{Q}}) - \widehat{\beta}_{\mathrm{adj}(2)}] \\
&= T_i[a_i^{**} - \mathbf{z}_i(\widehat{\mathbf{Q}}-\mathbf{Q}) - \widehat{\beta}_{\mathrm{adj}(1)}] + (1-T_i)[b_i^{**} - \mathbf{z}_i(\widehat{\mathbf{Q}}-\mathbf{Q}) - \widehat{\beta}_{\mathrm{adj}(2)}].
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{1}{n_A}\sum_{i\in A}\hat{e}_i^2 &= \frac{1}{n_A}\sum_{i\in A}[a_i^{**} - \mathbf{z}_i(\widehat{\mathbf{Q}}-\mathbf{Q}) - \widehat{\beta}_{\mathrm{adj}(1)}]^2 \\
&= S_1 + S_2 + S_3 - 2S_4 - 2S_5 - 2S_6
\end{aligned}$$

where

$$\begin{aligned}
S_1 &= \frac{1}{n_A}\sum_{i\in A}(a_i^{**})^2, \\
S_2 &= (\widehat{\mathbf{Q}}-\mathbf{Q})'\overline{\mathbf{z}'\mathbf{z}}_A(\widehat{\mathbf{Q}}-\mathbf{Q}), \\
S_3 &= \widehat{\beta}^2_{\mathrm{adj}(1)}, \\
S_4 &= \left(\frac{1}{n_A}\sum_{i\in A}a_i^{**}\mathbf{z}_i\right)(\widehat{\mathbf{Q}}-\mathbf{Q}), \\
S_5 &= \widehat{\beta}_{\mathrm{adj}(1)}\overline{a^{**}}_A, \\
S_6 &= \widehat{\beta}_{\mathrm{adj}(1)}\overline{\mathbf{z}}_A(\widehat{\mathbf{Q}}-\mathbf{Q}).
\end{aligned}$$

$S_1 \xrightarrow{p} \lim_{n\to\infty}\sigma^2_{a^{**}}$ by Lemma 7 and Condition 4.
The other terms are all $o_p(1)$:

- $S_2 \xrightarrow{p} 0$ because $\widehat{\mathbf{Q}} \xrightarrow{p} \mathbf{Q}$ (by Lemma 4) and $\overline{\mathbf{z}'\mathbf{z}}_A \xrightarrow{p} \langle\mathbf{z}'\mathbf{z}\rangle$ (by Lemma 1).
- $S_3 \xrightarrow{p} 0$ because $\widehat{\beta}_{\mathrm{adj}(1)} \xrightarrow{p} \overline{a} = 0$ (by Condition 4 and Lemma 6).

- $S_4 \xrightarrow{p} 0$ because

$$\frac{1}{n_A}\sum_{i\in A} a_i^{**}\mathbf{z}_i = \frac{1}{n_A}\sum_{i\in A}(a_i - \mathbf{Q}'\mathbf{z}_i')\mathbf{z}_i \xrightarrow{p} \langle a\mathbf{z}\rangle - \mathbf{Q}'\langle \mathbf{z}'\mathbf{z}\rangle$$

  (by Lemma 1) and $\widehat{\mathbf{Q}} \xrightarrow{p} \mathbf{Q}$.
- $S_5 \xrightarrow{p} 0$ because $\overline{a^{**}}_A \xrightarrow{p} \langle a\rangle - \langle \mathbf{z}\rangle\mathbf{Q} = 0$ (by Lemma 1 and Condition 4) and $\widehat{\beta}_{\text{adj}(1)} \xrightarrow{p} 0$.
- $S_6 \xrightarrow{p} 0$ because $\overline{\mathbf{z}}_A \xrightarrow{p} \mathbf{0}$ (by Lemma 1 and Condition 4), $\widehat{\beta}_{\text{adj}(1)} \xrightarrow{p} 0$, and $\widehat{\mathbf{Q}} \xrightarrow{p} \mathbf{Q}$.

Therefore,

$$\frac{1}{n_A}\sum_{i\in A}\hat{e}_i^2 \xrightarrow{p} \lim_{n\to\infty}\sigma^2_{a^{**}}.$$

Similarly,

$$\frac{1}{n-n_A}\sum_{i\in B}\hat{e}_i^2 \xrightarrow{p} \lim_{n\to\infty}\sigma^2_{b^{**}}.$$

Now note that

$$\begin{aligned} n^{-1}\sum_{i\in A}\hat{e}_i^2\mathbf{z}_i &= \frac{1}{n}\sum_{i\in A}[a_i - \mathbf{z}_i\widehat{\mathbf{Q}} - \widehat{\beta}_{\text{adj}(1)}]^2\mathbf{z}_i \\ &= \mathbf{R}_1 + \mathbf{R}_2 + \mathbf{R}_3 - 2\mathbf{R}_4 - 2\mathbf{R}_5 - 2\mathbf{R}_6 \end{aligned}$$

where

$$\begin{aligned} \mathbf{R}_1 &= \frac{1}{n}\sum_{i\in A} a_i^2\mathbf{z}_i, \\ \mathbf{R}_2 &= \frac{1}{n}\sum_{i\in A}(\mathbf{z}_i\widehat{\mathbf{Q}})^2\mathbf{z}_i, \\ \mathbf{R}_3 &= \tilde{p}_A\widehat{\beta}^2_{\text{adj}(1)}\overline{\mathbf{z}}_A, \\ \mathbf{R}_4 &= \widehat{\mathbf{Q}}'\frac{1}{n}\sum_{i\in A} a_i\mathbf{z}_i'\mathbf{z}_i, \\ \mathbf{R}_5 &= \tilde{p}_A\widehat{\beta}_{\text{adj}(1)}\overline{a\mathbf{z}}_A, \\ \mathbf{R}_6 &= \tilde{p}_A\widehat{\beta}_{\text{adj}(1)}\widehat{\mathbf{Q}}'\overline{\mathbf{z}'\mathbf{z}}_A. \end{aligned}$$

$\mathbf{R}_3$, $\mathbf{R}_5$, and $\mathbf{R}_6$ are $o_p(1)$ because $\widehat{\beta}_{\text{adj}(1)} \xrightarrow{p} 0$, $\overline{\mathbf{z}}_A \xrightarrow{p} \mathbf{0}$, and $\tilde{p}_A$, $\overline{a\mathbf{z}}_A$, $\overline{\mathbf{z}'\mathbf{z}}_A$, and $\widehat{\mathbf{Q}}$ converge to finite limits (by Condition 3, Lemma 1, and Lemma 4).

$\mathbf{R}_1$, $\mathbf{R}_2$, and $\mathbf{R}_4$ are $O_p(1)$, by Condition 1, Lemma 4, and repeated application of the Cauchy–Schwarz inequality. For example, for $k = 1, \ldots, K$, the kth element of $\mathbf{R}_2$ is

$$\frac{1}{n}\sum_{i\in A}\left(\sum_{j=1}^{K} z_{ij}\widehat{Q}_j\right)^2 z_{ik} \quad = \quad \sum_{j=1}^{K}\sum_{\ell=1}^{K}\left(\widehat{Q}_j\widehat{Q}_\ell\frac{1}{n}\sum_{i\in A} z_{ij}z_{i\ell}z_{ik}\right).$$

$\widehat{Q}_j$ and $\widehat{Q}_\ell$ are $O_p(1)$, and $n^{-1}\sum_{i\in A} z_{ij}z_{i\ell}z_{ik}$ is $O(1)$:

$$\begin{aligned}\left|\frac{1}{n}\sum_{i\in A} z_{ij}z_{i\ell}z_{ik}\right| &\leq \frac{1}{n}\sum_{i=1}^{n}|z_{ij}||z_{i\ell}z_{ik}| \leq \left(\frac{1}{n}\sum_{i=1}^{n} z_{ij}^2\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n} z_{i\ell}^2 z_{ik}^2\right)^{1/2} \\ &\leq \left(\frac{1}{n}\sum_{i=1}^{n} z_{ij}^4\right)^{1/4}\left(\frac{1}{n}\sum_{i=1}^{n} 1\right)^{1/4}\left(\frac{1}{n}\sum_{i=1}^{n} z_{i\ell}^4\right)^{1/4}\left(\frac{1}{n}\sum_{i=1}^{n} z_{ik}^4\right)^{1/4} \\ &< L^{3/4}.\end{aligned}$$

Therefore, $\mathbf{R}_2$ is $O_p(1)$.

Thus, $n^{-1}\sum_{i\in A}\hat{e}_i^2\mathbf{z}_i$ is $O_p(1)$. The proofs for $n^{-1}\sum_{i\in B}\hat{e}_i^2\mathbf{z}_i$ and $n^{-1}\sum_{i=1}^{n}\hat{e}_i^2\mathbf{z}_i'\mathbf{z}_i$ are similar. □

3. Proof of Theorem 1. We can assume Condition 4 without loss of generality, by an argument similar to that given in the proof of Lemma 6. Then $ATE = 0$, and by Lemma 3 and Condition 4,

$$\begin{aligned}\sqrt{n}(\widehat{ATE}_{\text{interact}} - ATE) &= \sqrt{n}[(\overline{a}_A - \overline{\mathbf{z}}_A\widehat{\mathbf{Q}}_a) - (\overline{b}_B - \overline{\mathbf{z}}_B\widehat{\mathbf{Q}}_b)] \\ &= \sqrt{n}[(\overline{a}_A - \overline{\mathbf{z}}_A\mathbf{Q}_a) - (\overline{b}_B - \overline{\mathbf{z}}_B\mathbf{Q}_b)] - \\ &\qquad \sqrt{n}\overline{\mathbf{z}}_A(\widehat{\mathbf{Q}}_a - \mathbf{Q}_a) + \sqrt{n}\overline{\mathbf{z}}_B(\widehat{\mathbf{Q}}_b - \mathbf{Q}_b).\end{aligned}$$

By a finite-population Central Limit Theorem [Freedman's (2008b) Theorem 1], $\sqrt{n}\overline{\mathbf{z}}_A$ and $\sqrt{n}\overline{\mathbf{z}}_B$ are $O_p(1)$, and by Lemma 5, $\widehat{\mathbf{Q}}_a - \mathbf{Q}_a$ and $\widehat{\mathbf{Q}}_b - \mathbf{Q}_b$ are $o_p(1)$. Therefore, $\sqrt{n}\overline{\mathbf{z}}_A(\widehat{\mathbf{Q}}_a - \mathbf{Q}_a)$ and $\sqrt{n}\overline{\mathbf{z}}_B(\widehat{\mathbf{Q}}_b - \mathbf{Q}_b)$ are $o_p(1)$.

The conclusion follows from Freedman's (2008b) Theorem 1 with a and b replaced by $a - \mathbf{z}\mathbf{Q}_a$ and $b - \mathbf{z}\mathbf{Q}_b$.

4. Proof of Corollary 1.1. We can assume Condition 4 without loss of generality: Centering a_i, b_i, and $\mathbf{z}_i$ has no effect on $\widehat{ATE}_{\text{interact}} - ATE$, $\widehat{ATE}_{\text{unadj}} - ATE$, $\mathbf{Q}_a$, $\mathbf{Q}_b$, or σ_E^2.

Note that:

$$\begin{aligned}
\lim_{n\to\infty}\sigma^2_{a^*} &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(a_i-\mathbf{z}_i\mathbf{Q}_a)^2\\
&= \langle a^2\rangle - \langle a\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle a\mathbf{z}\rangle',\\
\lim_{n\to\infty}\sigma^2_{b^*} &= \langle b^2\rangle - \langle b\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle',\\
\lim_{n\to\infty}\sigma_{a^*,b^*} &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}(a_i-\mathbf{z}_i\mathbf{Q}_a)(b_i-\mathbf{z}_i\mathbf{Q}_b)\\
&= \langle ab\rangle - \langle a\mathbf{z}\rangle\mathbf{Q}_b - \langle b\mathbf{z}\rangle\mathbf{Q}_a + \mathbf{Q}_a'\langle\mathbf{z}'\mathbf{z}\rangle\mathbf{Q}_b\\
&= \langle ab\rangle - \langle a\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle'.
\end{aligned}$$

By Freedman's (2008b) Theorem 1,

$$\begin{aligned}
\mathrm{avar}(\sqrt{n}[\widehat{ATE}_{\mathrm{unadj}}-ATE]) &= \mathrm{avar}(\sqrt{n}[\overline{a}_A-\overline{b}_B])\\
&= \frac{1-p_A}{p_A}\langle a^2\rangle + \frac{p_A}{1-p_A}\langle b^2\rangle + 2\langle ab\rangle.
\end{aligned}$$

Let

$$\Delta = \mathrm{avar}(\sqrt{n}[\widehat{ATE}_{\mathrm{unadj}}-ATE]) - \mathrm{avar}(\sqrt{n}[\widehat{ATE}_{\mathrm{interact}}-ATE]).$$

Then

$$\begin{aligned}
\Delta &= \frac{1-p_A}{p_A}\langle a\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle a\mathbf{z}\rangle' + \frac{p_A}{1-p_A}\langle b\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle' + 2\langle a\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle'\\
&= \frac{1}{p_A(1-p_A)}\mathbf{Q}_E'\langle\mathbf{z}'\mathbf{z}\rangle\mathbf{Q}_E = \frac{1}{p_A(1-p_A)}\lim_{n\to\infty}\sigma_E^2 \geq 0.
\end{aligned}$$

The matrix $\langle\mathbf{z}'\mathbf{z}\rangle$ is positive definite, so $\Delta/n = 0$ if and only if $\mathbf{Q}_E = \mathbf{0}$.

5. Proof of remark (iv) after Corollary 1.1. Suppose there are three treatment groups, A, B, and C, with associated dummy variables U_i, V_i, and W_i and potential outcomes a_i, b_i, and c_i. Let $ATE = \overline{a}-\overline{b}$, and let $\widehat{ATE}_{\mathrm{interact}}$ be the difference between the estimated coefficients on U_i and V_i in the no-intercept OLS regression of Y_i on U_i, V_i, W_i, $\mathbf{z}_i-\overline{\mathbf{z}}$, $U_i(\mathbf{z}_i-\overline{\mathbf{z}})$, and $W_i(\mathbf{z}_i-\overline{\mathbf{z}})$.

Assume the three groups are of fixed sizes n_A, n_B, and $n-n_A-n_B$. Assume regularity conditions analogous to Conditions 1–3: for example, $n_A/n\to p_A$ and $n_B/n\to p_B$, where $p_A>0$, $p_B>0$, and $p_A+p_B<1$. Without loss of generality, assume Condition 4.

Then $\sqrt{n}(\widehat{ATE}_{\mathrm{interact}}-ATE)$ converges in distribution to a Gaussian random variable with mean 0 and variance

$$\frac{1-p_A}{p_A}\lim_{n\to\infty}\sigma^2_{a^*} + \frac{1-p_B}{p_B}\lim_{n\to\infty}\sigma^2_{b^*} + 2\lim_{n\to\infty}\sigma_{a^*,b^*}.$$

The proof is essentially the same as that of Theorem 1.

Let $\widehat{ATE}_{\text{unadj}} = \overline{Y}_A - \overline{Y}_B$. By Freedman's (2008b) Theorem 1, the asymptotic variance of $\sqrt{n}(\widehat{ATE}_{\text{unadj}} - ATE)$ is

$$\frac{1-p_A}{p_A}\langle a^2\rangle + \frac{1-p_B}{p_B}\langle b^2\rangle + 2\langle ab\rangle.$$

Let

$$\Delta \quad = \quad \operatorname{avar}(\sqrt{n}[\widehat{ATE}_{\text{unadj}} - ATE]) - \operatorname{avar}(\sqrt{n}[\widehat{ATE}_{\text{interact}} - ATE]).$$

Then

$$\begin{aligned}
\Delta &= \frac{1-p_A}{p_A}\langle a\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle a\mathbf{z}\rangle' + \frac{1-p_B}{p_B}\langle b\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle' + 2\langle a\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle' \\
&= \frac{1-p_A}{p_A}\langle a\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle a\mathbf{z}\rangle' + \frac{p_A}{1-p_A}\langle b\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle' + 2\langle a\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle' + \\
&\qquad \left(\frac{1-p_B}{p_B} - \frac{p_A}{1-p_A}\right)\langle b\mathbf{z}\rangle\langle\mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle' \\
&= \frac{1}{p_A(1-p_A)}\lim_{n\to\infty}\sigma_E^2 + \left(\frac{1-p_B}{p_B} - \frac{p_A}{1-p_A}\right)\mathbf{Q}_b'\langle\mathbf{z}'\mathbf{z}\rangle\mathbf{Q}_b,
\end{aligned}$$

where $E_i = (\mathbf{z}_i - \overline{\mathbf{z}})\mathbf{Q}_E$ and $\mathbf{Q}_E = (1-p_A)\mathbf{Q}_a + p_A\mathbf{Q}_b$.

Similarly,

$$\Delta \quad = \quad \frac{1}{p_B(1-p_B)}\lim_{n\to\infty}\sigma_F^2 + \left(\frac{1-p_A}{p_A} - \frac{p_B}{1-p_B}\right)\mathbf{Q}_a'\langle\mathbf{z}'\mathbf{z}\rangle\mathbf{Q}_a,$$

where $F_i = (\mathbf{z}_i - \overline{\mathbf{z}})\mathbf{Q}_F$ and $\mathbf{Q}_F = p_B\mathbf{Q}_a + (1-p_B)\mathbf{Q}_b$.

The condition $p_A + p_B < 1$ implies

$$\frac{1-p_B}{p_B} - \frac{p_A}{1-p_A} > 0, \qquad \frac{1-p_A}{p_A} - \frac{p_B}{1-p_B} > 0.$$

Also, $\langle\mathbf{z}'\mathbf{z}\rangle$ is positive definite. Therefore, $\Delta \geq 0$, and the inequality is strict unless $\mathbf{Q}_a = \mathbf{0}$ and $\mathbf{Q}_b = \mathbf{0}$.

The proof extends to designs with more than three treatment groups.

6. Proof of Corollary 1.2. Again, we can assume Condition 4 without loss of generality. By Lemma 6,

$$\begin{aligned}
\mathrm{avar}(\sqrt{n}[\widehat{ATE}_{\mathrm{adj}} - ATE]) &= \frac{1-p_A}{p_A}\lim_{n\to\infty}\sigma^2_{a^{**}} + \frac{p_A}{1-p_A}\lim_{n\to\infty}\sigma^2_{b^{**}} + 2\lim_{n\to\infty}\sigma_{a^{**},b^{**}} \\
&= \frac{1-p_A}{p_A}[\langle a^2\rangle + \mathbf{Q}'\langle \mathbf{z}'\mathbf{z}\rangle\mathbf{Q} - 2\mathbf{Q}'\langle a\mathbf{z}\rangle'] + \\
&\quad \frac{p_A}{1-p_A}[\langle b^2\rangle + \mathbf{Q}'\langle \mathbf{z}'\mathbf{z}\rangle\mathbf{Q} - 2\mathbf{Q}'\langle b\mathbf{z}\rangle'] + \\
&\quad 2[\langle ab\rangle + \mathbf{Q}'\langle \mathbf{z}'\mathbf{z}\rangle\mathbf{Q} - \mathbf{Q}'\langle a\mathbf{z}\rangle' - \mathbf{Q}'\langle b\mathbf{z}\rangle'] \\
&= \frac{1-p_A}{p_A}\langle a^2\rangle + \frac{p_A}{1-p_A}\langle b^2\rangle + 2\langle ab\rangle + \\
&\quad \frac{1}{p_A(1-p_A)}\mathbf{Q}'\langle \mathbf{z}'\mathbf{z}\rangle\mathbf{Q} - \frac{2}{p_A}\mathbf{Q}'\langle a\mathbf{z}\rangle' - \frac{2}{1-p_A}\mathbf{Q}'\langle b\mathbf{z}\rangle'.
\end{aligned}$$

Let

$$\Delta = \mathrm{avar}(\sqrt{n}[\widehat{ATE}_{\mathrm{adj}} - ATE]) - \mathrm{avar}(\sqrt{n}[\widehat{ATE}_{\mathrm{interact}} - ATE]).$$

Then

$$\begin{aligned}
\Delta &= \frac{1}{p_A(1-p_A)}\mathbf{Q}'\langle \mathbf{z}'\mathbf{z}\rangle\mathbf{Q} - \frac{2}{p_A}\mathbf{Q}'\langle a\mathbf{z}\rangle' - \frac{2}{1-p_A}\mathbf{Q}'\langle b\mathbf{z}\rangle' + \\
&\quad \frac{1-p_A}{p_A}\langle a\mathbf{z}\rangle\langle \mathbf{z}'\mathbf{z}\rangle^{-1}\langle a\mathbf{z}\rangle' + \frac{p_A}{1-p_A}\langle b\mathbf{z}\rangle\langle \mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle' + 2\langle a\mathbf{z}\rangle\langle \mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle' \\
&= \left(\frac{p_A}{1-p_A} - 2 + \frac{1-p_A}{p_A}\right)\left(\langle a\mathbf{z}\rangle\langle \mathbf{z}'\mathbf{z}\rangle^{-1}\langle a\mathbf{z}\rangle' + \langle b\mathbf{z}\rangle\langle \mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle' - 2\langle a\mathbf{z}\rangle\langle \mathbf{z}'\mathbf{z}\rangle^{-1}\langle b\mathbf{z}\rangle'\right) \\
&= \frac{(2p_A-1)^2}{p_A(1-p_A)}(\mathbf{Q}_a - \mathbf{Q}_b)'\langle \mathbf{z}'\mathbf{z}\rangle(\mathbf{Q}_a - \mathbf{Q}_b) \\
&= \frac{(2p_A-1)^2}{p_A(1-p_A)}\lim_{n\to\infty}\sigma^2_D \geq 0.
\end{aligned}$$

7. Outline of proof of remark (iii) after Corollary 1.2. Without loss of generality, assume Condition 4. From the proof of Theorem 1,

$$\sqrt{n}\widehat{ATE}_{\mathrm{interact}} = \sqrt{n}[(\overline{a}_A - \overline{\mathbf{z}}_A\mathbf{Q}_a) - (\overline{b}_B - \overline{\mathbf{z}}_B\mathbf{Q}_b)] + o_p(1).$$

By Condition 4, $\tilde{p}_A\overline{\mathbf{z}}_A + (1-\tilde{p}_A)\overline{\mathbf{z}}_B = \mathbf{0}$. Therefore, $\overline{\mathbf{z}}_A = (1-\tilde{p}_A)(\overline{\mathbf{z}}_A - \overline{\mathbf{z}}_B)$ and $\overline{\mathbf{z}}_B = -\tilde{p}_A(\overline{\mathbf{z}}_A - \overline{\mathbf{z}}_B)$. It follows that

$$\sqrt{n}\widehat{ATE}_{\mathrm{interact}} = \sqrt{n}\{\overline{a}_A - \overline{b}_B - (\overline{\mathbf{z}}_A - \overline{\mathbf{z}}_B)[(1-p_A)\mathbf{Q}_a + p_A\mathbf{Q}_b]\} + o_p(1).$$

Now let $\widehat{ATE}_{\text{tyranny}}$ and $\widehat{\mathbf{Q}}_{\text{tyranny}}$ be the estimated coefficients on T_i and $\mathbf{z}_i$ from a weighted least squares regression of Y_i on T_i and $\mathbf{z}_i$, with weights

$$w_i = \frac{1-\tilde{p}_A}{\tilde{p}_A} T_i + \frac{\tilde{p}_A}{1-\tilde{p}_A}(1-T_i).$$

It can be shown that $\widehat{\mathbf{Q}}_{\text{tyranny}} \xrightarrow{p} (1-p_A)\mathbf{Q}_a + p_A \mathbf{Q}_b$. The proof is similar to that of Lemma 4, after noting that weighted least squares is equivalent to OLS with all data values (including the constant) multiplied by $\sqrt{w_i}$.

It follows that

$$\sqrt{n}\widehat{ATE}_{\text{tyranny}} = \sqrt{n}\{\overline{a}_A - \overline{b}_B - (\overline{\mathbf{z}}_A - \overline{\mathbf{z}}_B)[(1-p_A)\mathbf{Q}_a + p_A\mathbf{Q}_b]\} + o_p(1).$$

The proof is similar to arguments in the proofs of Lemmas 2 and 6.

Therefore, $\sqrt{n}(\widehat{ATE}_{\text{tyranny}} - \widehat{ATE}_{\text{interact}}) \xrightarrow{p} 0$.

8. Proof of Theorem 2. We can assume Condition 4 without loss of generality, by arguments similar to those given in the proofs of Lemmas 4, 6, and 8.

By Lemma 8, $n\widehat{v}_{\text{adj}} = M_{11} + M_{22} - 2M_{12}$, where

$$\mathbf{M} = (n^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}\left(n^{-1}\sum_{i=1}^{n}\hat{e}_i^2\tilde{\mathbf{x}}_i'\tilde{\mathbf{x}}_i\right)(n^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1}.$$

Using Condition 4,

$$n^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{C} & \mathbf{D} \\ \mathbf{D}' & \underline{\mathbf{z}'\mathbf{z}} \end{bmatrix},$$

where

$$\mathbf{C} = \begin{bmatrix} \tilde{p}_A & 0 \\ 0 & 1-\tilde{p}_A \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} \tilde{p}_A\overline{\mathbf{z}}_A \\ (1-\tilde{p}_A)\overline{\mathbf{z}}_B \end{bmatrix}.$$

By Conditions 2–4 and Lemma 1, $\tilde{p}_A \to p_A$, $\overline{\mathbf{z}}_A \xrightarrow{p} \mathbf{0}$, $\overline{\mathbf{z}}_B \xrightarrow{p} \mathbf{0}$, and $\langle\mathbf{z}'\mathbf{z}\rangle$ is invertible. Therefore,

$$(n^{-1}\tilde{\mathbf{X}}'\tilde{\mathbf{X}})^{-1} \xrightarrow{p} \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \langle\mathbf{z}'\mathbf{z}\rangle^{-1} \end{bmatrix}$$

where

$$\mathbf{F} = \begin{bmatrix} 1/p_A & 0 \\ 0 & 1/(1-p_A) \end{bmatrix}.$$

Also,

$$\tilde{\mathbf{x}}_i'\tilde{\mathbf{x}}_i = \begin{bmatrix} \mathbf{G} & \mathbf{H} \\ \mathbf{H}' & \mathbf{z}_i'\mathbf{z}_i \end{bmatrix},$$

where

$$\mathbf{G} = \begin{bmatrix} T_i & 0 \\ 0 & 1-T_i \end{bmatrix}, \qquad \mathbf{H} = \begin{bmatrix} T_i\mathbf{z}_i \\ (1-T_i)\mathbf{z}_i \end{bmatrix}.$$

So

$$n^{-1}\sum_{i=1}^{n} \hat{e}_i^2 \tilde{\mathbf{x}}_i'\tilde{\mathbf{x}}_i = \begin{bmatrix} \mathbf{K} & \mathbf{L} \\ \mathbf{L}' & n^{-1}\sum_{i=1}^{n} \hat{e}_i^2 \mathbf{z}_i'\mathbf{z}_i \end{bmatrix},$$

where

$$\begin{aligned} \mathbf{K} &= \begin{bmatrix} n^{-1}\sum_{i\in A}\hat{e}_i^2 & 0 \\ 0 & n^{-1}\sum_{i\in B}\hat{e}_i^2 \end{bmatrix} = \begin{bmatrix} \tilde{p}_A n_A^{-1}\sum_{i\in A}\hat{e}_i^2 & 0 \\ 0 & (1-\tilde{p}_A)(n-n_A)^{-1}\sum_{i\in B}\hat{e}_i^2 \end{bmatrix}, \\ \mathbf{L} &= \begin{bmatrix} n^{-1}\sum_{i\in A}\hat{e}_i^2\mathbf{z}_i \\ n^{-1}\sum_{i\in B}\hat{e}_i^2\mathbf{z}_i \end{bmatrix}. \end{aligned}$$

By Lemma 9 and Condition 3, $\mathbf{L}$ and $n^{-1}\sum_{i=1}^{n}\hat{e}_i^2\mathbf{z}_i'\mathbf{z}_i$ are $O_p(1)$, and

$$\mathbf{K} \xrightarrow{p} \begin{bmatrix} p_A \lim_{n\to\infty}\sigma_{a^{**}}^2 & 0 \\ 0 & (1-p_A)\lim_{n\to\infty}\sigma_{b^{**}}^2 \end{bmatrix}.$$

The above results imply that the upper-left 2×2 block of $\mathbf{M}$ converges in probability to

$$\begin{aligned} &\begin{bmatrix} 1/p_A & 0 \\ 0 & 1/(1-p_A) \end{bmatrix} \begin{bmatrix} p_A \lim_{n\to\infty}\sigma_{a^{**}}^2 & 0 \\ 0 & (1-p_A)\lim_{n\to\infty}\sigma_{b^{**}}^2 \end{bmatrix} \begin{bmatrix} 1/p_A & 0 \\ 0 & 1/(1-p_A) \end{bmatrix} \\ &\qquad = \begin{bmatrix} p_A^{-1} \lim_{n\to\infty}\sigma_{a^{**}}^2 & 0 \\ 0 & (1-p_A)^{-1}\lim_{n\to\infty}\sigma_{b^{**}}^2 \end{bmatrix}. \end{aligned}$$

Thus,

$$n\widehat{v}_{\text{adj}} \xrightarrow{p} \frac{1}{p_A}\lim_{n\to\infty}\sigma_{a^{**}}^2 + \frac{1}{1-p_A}\lim_{n\to\infty}\sigma_{b^{**}}^2.$$

Lemma 6 implies

$$\text{avar}(\sqrt{n}[\widehat{ATE}_{\text{adj}} - ATE]) = \frac{1-p_A}{p_A}\lim_{n\to\infty}\sigma_{a^{**}}^2 + \frac{p_A}{1-p_A}\lim_{n\to\infty}\sigma_{b^{**}}^2 + 2\lim_{n\to\infty}\sigma_{a^{**},b^{**}}.$$

Let $\Delta = \text{plim } n\widehat{v}_{\text{adj}} - \text{avar}(\sqrt{n}[\widehat{ATE}_{\text{adj}} - ATE])$. Then

$$\begin{aligned}\Delta &= \lim_{n\to\infty} \sigma^2_{a^{**}} + \lim_{n\to\infty} \sigma^2_{b^{**}} - 2\lim_{n\to\infty} \sigma_{a^{**},b^{**}} \\ &= \lim_{n\to\infty} \sigma^2_{(a^{**}-b^{**})} = \lim_{n\to\infty} \sigma^2_{(a-b)} \geq 0.\end{aligned}$$

The proof for $n\widehat{v}_{\text{interact}}$ is similar.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720-3860
USA
LINSTON@GMAIL.COM