Statistica Sinica: Supplement

CAUSAL INFERENCE FROM POSSIBLY UNBALANCED SPLIT-PLOT DESIGNS: A RANDOMIZATION-BASED PERSPECTIVE

Rahul Mukerjee and Tirthankar Dasgupta

Indian Institute of Managament Calcutta and Rutgers University

Supplementary Material

Proof of all results

In what follows, E_1 and cov_1 denote unconditional expectation and covariance with respect to the randomization at the WP stage, while E_2 and cov_2 denote expectation and covariance with respect to the randomization at the SP stage, conditional on the WP stage assignment.

Proof of Proposition 1. Follows from straightforward conditioning arguments.

$\square$

Proof of Theorem 2. Recall that

$$\begin{aligned}\overline{U}_w^{\mathrm{obs}}(z_1z_2) &= \frac{1}{r_{w2}(z_2)}\sum_{i\in T_{w2}(z_2)} U_i(z_1z_2), w\in T_1(z_1)\\ \overline{U}^{\mathrm{obs}}(z_1z_2) &= \frac{1}{r_1(z_1)}\sum_{w\in T_1(z_1)} \overline{U}_w^{\mathrm{obs}}(z_1z_2).\end{aligned}$$

Consequently,

$$E_2\left\{\overline{U}^{\mathrm{obs}}(z_1z_2)\right\} = \frac{1}{r_1(z_1)}\sum_{w\in T_1(z_1)} \overline{U}_w(z_1z_2),$$

and,

$$E_2\left\{\overline{U}^{\mathrm{obs}}(z_1^*z_2^*)\right\} = \frac{1}{r_1(z_1^*)}\sum_{w\in T_1(z_1^*)} \overline{U}_w(z_1^*z_2^*).$$

Defining $\delta(z_1, z_1^*)$ as an indicator that equals one if $z_1 = z_1^*$ and zero otherwise, we have

$$\begin{aligned}&\mathrm{cov}_1\left[E_2\left\{\overline{U}^{\mathrm{obs}}(z_1z_2)\right\}, E_2\left\{\overline{U}^{\mathrm{obs}}(z_1^*z_2^*)\right\}\right]\\ =\ & \frac{1}{(W-1)Wr_1(z_1)}\sum_{w=1}^{W}\left\{\overline{U}_w(z_1z_2)-\overline{U}(z_1z_2)\right\}\left\{\overline{U}_w(z_1^*z_2^*)-\overline{U}(z_1^*z_2^*)\right\}\left\{W\delta(z_1,z_1^*)-r_1(z_1)\right\}\\ =\ & \frac{1}{W\overline{M}r_1(z_1)}S_{\mathrm{bt}}(z_1z_2, z_1^*z_2^*)\left\{W\delta(z_1,z_1^*)-r_1(z_1)\right\}.\end{aligned}$$

Next,

$$\begin{aligned}&\mathrm{cov}_2\left\{\overline{U}^{\mathrm{obs}}(z_1z_2), \overline{U}^{\mathrm{obs}}(z_1^*z_2^*)\right\}\\ =\ & \delta(z_1,z_1^*)\sum_{w\in T_1(z_1)}\frac{S_{\mathrm{in},w}(z_1z_2, z_1^*z_2^*)\left\{M_w\delta(z_2,z_2^*)-r_{w2}(z_2)\right\}}{M_wr_{w2}(z_2)\left\{r_1(z_1)\right\}^2}.\end{aligned}$$

so that

$$
\begin{aligned}
& E_1\left[\operatorname{cov}_2\left\{\overline{U}^{\mathrm{obs}}(z_1z_2), \overline{U}^{\mathrm{obs}}(z_1^*z_2^*)\right\}\right] \\
= \; & \delta(z_1, z_1^*) \sum_{w=1}^{W} \frac{S_{\mathrm{in},w}(z_1z_2, z_1^*z_2^*)\left\{M_w\delta(z_2, z_2^*) - r_{w2}(z_2)\right\}}{WM_w r_1(z_1) r_{w2}(z_2)}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
& \operatorname{cov}\left\{\overline{U}^{\mathrm{obs}}(z_1z_2), \overline{U}^{\mathrm{obs}}(z_1^*z_2^*)\right\} \\
= \; & \delta(z_1, z_1^*)\left\{\frac{S_{\mathrm{bt}}(z_1z_2, z_1^*z_2^*)}{\overline{M} r_1(z_1)} - \sum_{w=1}^{W} \frac{S_{\mathrm{in},w}(z_1z_2, z_1^*z_2^*)}{WM_w r_1(z_1)}\right\} \\
+ \; & \delta(z_1, z_1^*)\delta(z_2, z_2^*) \sum_{w=1}^{W} \frac{S_{\mathrm{in},w}(z_1z_2, z_1^*z_2^*)}{W r_1(z_1) r_{w2}(z_2)} - \frac{S_{\mathrm{bt}}(z_1z_2, z_1^*z_2^*)}{N}. \quad \text{(S1.1)}
\end{aligned}
$$

Since $\widehat{\overline{\tau}} = \sum_{z_1 \in Z_1} \sum_{z_2 \in Z_2} g(z_1z_2)\overline{U}^{\mathrm{obs}}(z_1z_2)$, we have that

$$
\operatorname{var}(\widehat{\overline{\tau}}) = \sum_{z_1 \in Z_1} \sum_{z_2 \in Z_2} \sum_{z_1^* \in Z_1} \sum_{z_2^* \in Z_2} g(z_1z_2) g(z_1^*z_2^*) \operatorname{cov}\left\{\overline{U}^{\mathrm{obs}}(z_1z_2), \overline{U}^{\mathrm{obs}}(z_1^*z_2^*)\right\}.
$$

Substituting the expression of $\operatorname{cov}\left\{\overline{U}^{\mathrm{obs}}(z_1z_2), \overline{U}^{\mathrm{obs}}(z_1^*z_2^*)\right\}$ from (S1.1) in the above, the first two terms in the expression of $\operatorname{var}(\widehat{\overline{\tau}})$ in Theorem 1 follow immediately. The last term can be explained as

$$
\begin{aligned}
& \sum_{z_1 \in Z_1} \sum_{z_2 \in Z_2} \sum_{z_1^* \in Z_1} \sum_{z_2^* \in Z_2} g(z_1z_2) g(z_1^*z_2^*) S_{\mathrm{bt}}(z_1z_2, z_1^*z_2^*)/N \\
= \; & \frac{\overline{M}}{(W-1)N} \sum_{w=1}^{W} \left[\sum_{z_1 \in Z_1} \sum_{z_2 \in Z_2} g(z_1z_2)\left\{\overline{U}_w(z_1z_2) - \overline{U}(z_1z_2)\right\}\right]^2 \\
= \; & \frac{1}{W(W-1)} \sum_{w=1}^{W} \left[\sum_{z_1 \in Z_1} \sum_{z_2 \in Z_2} g(z_1z_2)\left\{(M_w/\overline{M})\overline{Y}_w(z_1z_2) - \overline{Y}(z_1z_2)\right\}\right]^2 \\
= \; & \frac{1}{W(W-1)} \sum_{w=1}^{W} \left\{(M_w/\overline{M})\overline{\tau}_w - \overline{\tau}\right\}^2 = \Delta.
\end{aligned}
$$

□

Proof of Theorem 3.

$$
\begin{aligned}
E_2\left\{\widehat{S}(z_1z_2, z_1z_2^*)\right\} &= \frac{1}{r_1(z_1)}\sum_{w\in T_1(z_1)} \mathrm{cov}_2\left\{\overline{U}_w^{\mathrm{obs}}(z_1z_2), \overline{U}_w^{\mathrm{obs}}(z_1z_2^*)\right\} \\
&+ \frac{1}{r_1(z_1)-1}\sum_{w\in T_1(z_1)}\left\{\overline{U}_w(z_1z_2) - \widetilde{\overline{U}}(z_1z_2)\right\}\left\{\overline{U}_w(z_1z_2^*) - \widetilde{\overline{U}}(z_1z_2^*)\right\},
\end{aligned}
$$

where $\widetilde{\overline{U}}(z_1z_2) = \sum_{w\in T_1(z_1)} \overline{U}_w(z_1z_2)/r_1(z_1)$, and $\widetilde{\overline{U}}(z_1z_2^*)$ is similarly defined. For any $w \in T_1(z_1)$,

$$
\mathrm{cov}_2\left\{\overline{U}_w^{\mathrm{obs}}(z_1z_2), \overline{U}_w^{\mathrm{obs}}(z_1z_2^*)\right\} = \frac{S_{\mathrm{in},w}(z_1z_2, z_1z_2^*)\left\{M_w\delta(z_2, z_2^*) - r_{w2}(z_2)\right\}}{M_w r_{w2}(z_2)}.
$$

Thus,

$$
\begin{aligned}
E\left\{\widehat{S}(z_1z_2, z_1z_2^*)\right\} &= E_1E_2\left\{\widehat{S}(z_1z_2, z_1z_2^*)\right\} \\
&= \sum_{w=1}^{W}\frac{S_{\mathrm{in},w}(z_1z_2, z_1z_2^*)\left\{M_w\delta(z_2, z_2^*) - r_{w2}(z_2)\right\}}{WM_w r_{w2}(z_2)} \\
&+ \frac{1}{W-1}\sum_{w=1}^{W}\left\{\overline{U}_w(z_1z_2) - \overline{U}(z_1z_2)\right\}\left\{\overline{U}_w(z_1z_2^*) - \overline{U}(z_1z_2^*)\right\} \\
&= \sum_{w=1}^{W}\frac{S_{\mathrm{in},w}(z_1z_2, z_1z_2^*)\left\{M_w\delta(z_2, z_2^*) - r_{w2}(z_2)\right\}}{WM_w r_{w2}(z_2)} + \frac{S_{\mathrm{bt}}(z_1z_2, z_1z_2^*)}{\overline{M}}.
\end{aligned}
$$

The result stated in Theorem 3 is evident from the above. □

Proof of Proposition 2. Because $w \neq w^*$, by (2.5) and the definition of G_w^{obs}, conditionally on the assignment of the WPs to the level combinations of the

WP factors, G_w^{obs} and $G_{w^*}^{\text{obs}}$ are independent and the conditional expectation of their product equals

$$\left\{ \sum_{z_2 \in Z_2} g(z_{1w} z_2) \overline{Y}_w(z_{1w} z_2) \right\} \left\{ \sum_{z_2 \in Z_2} g(z_{1w^*} z_2) \overline{Y}_{w^*}(z_{1w^*} z_2) \right\}.$$

The result now follows from (3.2), noting that the pair (z_{1w}, z_{1w^*}) equals any (z_1, z_1^*) with probability $\frac{r_1(z_1)\{r_1(z_1^*) - \delta(z_1, z_1^*)\}}{W(W-1)}$. □

Proof of the necessity part of Theorem 5. Suppose a psd matrix $B = (b_{ww^*})$ of order W and satisfying (c1)-(c3) exists. Then by (c1),

$$|b_{ww^*}| \leq M_w M_{w^*}, \quad w, w^* = 1, \ldots, W, \quad w \neq w^*. \tag{S1.2}$$

Hence using (c2), (S1.2), and (c1) in succession,

$$0 = b_{W1} + \ldots + b_{WW} \geq b_{WW} - M_W(M_1 + \ldots + M_{W-1}) = M_W(M_W - M_1 - \ldots - M_{W-1}), \tag{S1.3}$$

which implies $M_W \leq M_1 + \ldots + M_{W-1}$. If possible, let equality hold here. Then equality holds throughout in (S1.3), and invoking (S1.2), this yields

$$b_{Ww} = -M_W M_w, \quad w = 1, \ldots, W-1. \tag{S1.4}$$

For any w, w^* such that $w < w^* < W$, by (c1) and (S1.4), the principal minor of B, as given by its wth, w^*th and Wth rows and columns turns out to be $-M_W^2(b_{ww^*} - M_w M_{w^*})^2$. Because this principal minor is nonnegative due to psd-ness of B, it follows that $b_{ww^*} = M_w M_{w^*}$. This, in conjunction with (c1) and (S1.4), implies that $B = bb'$, where $b = (M_1, \ldots, M_{W-1}, -M_W)'$. But then $\text{rank}(B) = 1 < W - 1$, and (c3) is violated. This contradiction proves the necessity of the condition $M_W < M_1 + \ldots + M_{W-1}$. □

To prove the sufficiency part of Theorem 5, we first state a lemma that is crucial in this proof and also leads to the algorithm for construction of the symmetric psd matrix B of order W that satisfies conditions (c1)-(c3).

Lemma 1. *Let $W \geq 3$. Suppose $M_1, \ldots, M_W$ are not all equal and $M_1 \leq \ldots \leq M_W$, as per (5.1). Let e denote the $(W-1) \times 1$ vector of ones and $\mu = (M_1, \ldots, M_{W-1})'$.*

(a) Then there exists a $(W-1) \times 1$ vector x with elements ± 1 such that $|\mu' x| < M_W$.

(b) If, in addition, condition (5.3) holds, i.e., $M_W < M_1 + \ldots + M_{W-1}$, then, with the vector x as in (a) above, there exist nonnegative constants a_1, a_2 satisfying $a_1 + a_2 < 1$, such that equation (5.6) holds, i.e.,

$$a_1 \left\{ (\mu' x)^2 - \mu' \mu \right\} + a_2 \left\{ (\mu' e)^2 - \mu' \mu \right\} = M_W^2 - \mu' \mu.$$

Proof of Lemma 1. Part (a). It will suffice to show that there exist $x_1, \ldots, x_{W-1}$, each $+1$ or -1, such that $|\sum_{w=1}^{W-1} M_w x_w| < M_W$. One can then simply take $x = (x_1, \ldots, x_{W-1})'$. Recall that $M_1 \leq M_2 \leq \ldots \leq M_W$, as per (5.1). Because $M_1 \ldots, M_W$ are not all equal, this yields

$$M_1 < M_W. \tag{S1.5}$$

Let h be the largest nonnegative integer such that

$$M_{W-2h} = M_W. \tag{S1.6}$$

By (S1.5), $W - 2h \geq 2$. If $h \geq 1$, define

$$x_{W-h} = \ldots = x_{W-1} = 1, \quad x_{W-2h} = \ldots = x_{W-h-1} = -1, \tag{S1.7}$$

and note that

$$\sum_{w=W-2h}^{W-1} M_w x_w = 0, \tag{S1.8}$$

because by (5.1) and (S1.6), $M_w = M_W$ for $w = W - 2h, \ldots, W - 1$. Now, if $W - 2h = 2$, then with $x_1 = 1$ and $x_2, \ldots, x_{W-1}$ as in (S1.7), $|\sum_{w=1}^{W-1} M_w x_w| = M_1 < M_W$, by (S1.5) and (S1.8).

Next, let $W - 2h \geq 3$. Then, by (5.1),

$$\sum_{w=2}^{W-2h-1} M_w \geq (W - 2h - 2)M_2 \geq M_1.$$

Let w_1 be the largest integer in $\{1, \ldots, W - 2h - 2\}$ such that $\sum_{w=1}^{w_1} M_w \leq \sum_{w=w_1+1}^{W-2h-1} M_w$. If $w_1 = W - 2h - 2$, then $\sum_{w=1}^{W-2h-2} M_w \leq M_{W-2h-1}$. So, with $x_1 = \ldots = x_{W-2h-2} = -1$, $x_{W-2h-1} = 1$ and $x_{W-2h}, \ldots, x_{W-1}$ as in (S1.7) when $h \geq 1$,

$$\left|\sum_{w=1}^{W-1} M_w x_w\right| = M_{W-2h-1} - \sum_{w=1}^{W-2h-2} M_w < M_{W-2h-1} \leq M_W,$$

by (S1.8).

Now, suppose $1 \leq w_1 \leq W-2h-3$, in which case $W-2h \geq 4$. Then,

$$\sum_{w=1}^{w_1} M_w \leq \sum_{w=w_1+1}^{W-2h-1} M_w, \quad \text{and} \quad \sum_{w=1}^{w_1+1} M_w > \sum_{w=w_1+2}^{W-2h-1} M_w.$$

As a result, either

$$\text{(i)} \quad \left|\sum_{w=w_1+1}^{W-2h-1} M_w - \sum_{w=1}^{w_1} M_w\right| < M_W \quad \text{or (ii)} \quad \left|\sum_{w=1}^{w_1+1} M_w - \sum_{w=w_1+2}^{W-2h-1} M_w\right| < M_W.$$

Else,

$$\sum_{w=w_1+1}^{W-2h-1} M_w - \sum_{w=1}^{w_1} M_w \geq M_W, \quad \text{as well as} \quad \sum_{w=1}^{w_1+1} M_w - \sum_{w=w_1+2}^{W-2h-1} M_w \geq M_W.$$

Adding these two inequalities, we have $M_{w_1+1} \geq M_W$, which is impossible by the definition of h, because $w_1 + 1 \leq W - 2h - 2$.

If (i) holds, then the choice $x_1 = \ldots = x_{w_1} = -1$, $x_{w_1+1} = \ldots = x_{W-2h-1} = 1$, coupled with $x_{W-2h}, \ldots, x_{W-1}$ as in (S1.7) when $h \geq 1$, entails $\left|\sum_{w=1}^{W-1} M_w x_w\right| < M_W$, by (S1.8). Similarly, if (ii) holds, then the choice $x_1 = \ldots = x_{w_1+1} = -1$, $x_{w_1+2} = \ldots = x_{W-2h-1} = 1$, coupled with $x_{W-2h}, \ldots, x_{W-1}$ as in (S1.7) when $h \geq 1$, entails $\left|\sum_{w=1}^{W-1} M_w x_w\right| < M_W$.

Part (b): Let $M_W < M_1 + \ldots + M_{W-1} = \mu' e$, and let the vector x be as in part (a) above, so that $|\mu' x| < M_W$. Let $\phi_1 = (\mu' x)^2 - \mu'\mu$, $\phi = M_W^2 - \mu'\mu$ and $\phi_2 = (\mu' e)^2 - \mu'\mu$. Then $\phi_1 < \phi < \phi_2$, as $|\mu' x| < M_W < \mu' e$. As a result, there exist constants $\tilde{a}_1$ and $\tilde{a}_2$ such that $0 \leq \tilde{a}_1, \tilde{a}_2 < 1$ and $\tilde{a}_1\phi_1 < \phi < \tilde{a}_2\phi_2$. Let $\xi = (\tilde{a}_2\phi_2 - \phi) / (\tilde{a}_2\phi_2 - \tilde{a}_1\phi_1)$. Then $0 < \xi < 1$. Hence, if we take $a_1 = \tilde{a}_1\xi$, $a_2 = \tilde{a}_2(1-\xi)$, then $a_1, a_2 \geq 0$ and $a_1 + a_2 < 1$, because $a_1 + a_2$ is a weighted average of $\tilde{a}_1$ and $\tilde{a}_2$, both of which are less than one. Moreover, $a_1\phi_1 + a_2\phi_2 = \phi$ by the definition of ξ, i.e., a_1 and a_2 satisfy (5.6). $\square$

Proof of the sufficiency part of Theorem 5. In view of Lemma 1, this follows from steps 1-4 in Section 5, noting that (i) the matrix A there is positive definite, and hence the matrix B there is psd of rank $W-1$ with each row sum zero, (ii) A has diagonal elements $M_1^2, \ldots, M_{W-1}^2$, and (iii) by (29),

$$e' A e = a_1(\mu' x)^2 + a_2(\mu' e)^2 + (1 - a_1 - a_2)\mu'\mu = M_W^2,$$

because $De = \mu$.

$\square$

Symbol Chart

Table 1: Symbols used in the manuscript and their explanation

Symbol	Meaning	Symbol	Meaning
A	matrix, in sufficiency part of Theorem 4	a	constant, in sufficiency part of Theorem 4
B	matrix in new variance estimator	b	element of matrix B
D	diagonal matrix of whole-plot sizes	e	vector of ones
E	expectation	g	function defining treatment contrast
F	factor	h	integer, in proving Lemma A.1(a)
G	term associated with the new variance estimator	i	dummy subscript for unit
H	term used in defining the new variance estimator	k	dummy subscript
I	identity matrix	m	number of factors
J	matrix of ones	r	treatment replication
M	number of sub-plots in a whole plot	u	dummy subscript
N	total number of units	v	dummy subscript
S	similar to mean square/product component	w	dummy, whole-plot index
T	set of whole- or sub-plots	x	vector of ± 1, in sufficiency of Theorem 4
U	transformed outcome	z	treatment combination
V	variance estimator		
W	number of whole-plots	τ	treatment contrast
Y	potential outcome	δ	Kronecker delta
Z	set of level combinations	μ	subvector of whole-plot sizes
		ϕ	in the proof of Lemma A.1(b)
Ω	whole-plot	ζ	in the proof of Lemma A.1(b)
Δ	bias in variance estimation	λ	eigenvalue