Supplemental Appendix

Causal inference for social network formation

Maximilian Kasy Elizabeth Linos Sanaz Mobasseri

April 3, 2026

A Proofs

Proof of Lemma 1:

- By the structural relationship and the exclusion restriction,

$$Y_{ij} \cdot \mathbf{1}(D_{ij} = d) = Y_{ij}^d \cdot \mathbf{1}(D_{ij} = d).$$

- By the support condition, $P(D_{ij} = d) > 0$
- Because Y_{ij}^d is a function of f,

$$E\left[Y_{ij}^d \cdot \frac{\mathbf{1}(D_{ij} = d)}{P(D_{ij} = d)} \middle| f\right] = Y_{ij}^d \cdot E\left[\frac{\mathbf{1}(D_{ij} = d)}{P(D_{ij} = d)} \middle| f\right]$$

- By the randomization condition,

$$E\left[\frac{\mathbf{1}(D_{ij} = d)}{P(D_{ij} = d)} \middle| f\right] = 1.$$

The claim follows. □

Proof of Lemma 2:

1. By item 3 of Assumption 1 and Assumption 2, A^1 is a uniform random draw from $\mathcal{A} = \{A_\pi : \pi \in \Pi\}$.

2. By definition, $p_{\pi(i)\pi(j)}(d) = P(d_{\pi(i)\pi(j)}(A^1) = d)$.

3. By equivariance, $d_{\pi(i)\pi(j)}(A^1) = d_{ij}(A^1_\pi)$.

4. By assumption, A^1 is a uniform draw from $\mathcal{A} = \{A'_\pi : \pi' \in \Pi\}$. The mapping from A to A_π is one-to-one on $\mathcal{A}$, because Π is a group: $\pi \circ \pi' \in \Pi$ and $\pi^{-1} \circ \pi'$ in Π, and thus $A_{\pi\circ\pi'}, A_{\pi^{-1}\circ\pi'} \in \mathcal{A}$. It follows that A^1_π is also a uniform random draw from $\mathcal{A}$.

5. Therefore the distribution of $d_{ij}(A^1_\pi)$ is the same as the distribution of $d_{ij}(A^1)$. The first claim follows.

6. The second claim follows because $P_{\pi(i)\pi(j)} = p_{\pi(i)\pi(j)}(D_{\pi(i)\pi(j)}) = p_{ij}(D_{\pi(i)\pi(j)}) \sim p_{ij}(D_{ij}) = P_{ij}$, where the first equality holds by definition, the second equality uses the first claim (applied at $d = D_{\pi(i)\pi(j)}$), and the distributional step uses $D_{\pi(i)\pi(j)} \sim D_{ij}$.

$\square$

Proof of Proposition 1:

1. By the first order condition for the least squares estimator,

$$\widehat{\beta} = \left(\sum_{(i,j)\in\mathcal{E}} \frac{1}{P_{ij}} D_{ij} \cdot D'_{ij} \right)^{-1} \cdot \left(\sum_{(i,j)\in\mathcal{E}} \frac{1}{P_{ij}} Y_{ij} \cdot D_{ij} \right).$$

2. Consider the second term of this expression. By the structural relationship and the exclusion restriction, we can rewrite

$$\frac{Y_{ij} \cdot D_{ij}}{P_{ij}} = \sum_d Y^d_{ij} \cdot d \cdot \frac{\mathbf{1}(D_{ij} = d)}{P(D_{ij} = d)}.$$

Therefore, because Y_{ij}^d is a function of f, and by the randomization condition and support condition (as in Lemma 1),

$$E\left[\frac{Y_{ij} \cdot D_{ij}}{P_{ij}} \middle| f\right] = \sum_d Y_{ij}^d \cdot d.$$

3. Consider next the first term of the expression for $\widehat{\beta}$. By a similar argument as for the second term, for all i, j

$$E\left[\frac{1}{P_{ij}} D_{ij} \cdot D'_{ij} \middle| f\right] = \sum_d d \cdot d'.$$

4. By Lemma 2, $\frac{1}{P_{ij}} D_{ij} \cdot D'_{ij}$ is replaced by $\frac{1}{P_{\pi(i)\pi(j)}} D_{\pi(i)\pi(j)} \cdot D'_{\pi(i)\pi(j)}$ when the network A^1 is replaced by A^1_π. By the invarance condition for $\mathcal{E}$ of Assumption 2,

$$\sum_{(i,j)\in\mathcal{E}} \frac{1}{P_{\pi(i)\pi(j)}} D_{\pi(i)\pi(j)} \cdot D'_{\pi(i)\pi(j)} = \sum_{(i,j)\in\pi(\mathcal{E})} \frac{1}{P_{ij}} D_{ij} \cdot D'_{ij} = \sum_{(i,j)\in\mathcal{E}} \frac{1}{P_{ij}} D_{ij} \cdot D'_{ij},$$

where we denote $\pi(\mathcal{E}) = \{(\pi(i), \pi(j)) : (i,j) \in \mathcal{E}\}$. This implies that $\sum_{(i,j)\in\mathcal{E}} \frac{1}{P_{ij}} D_{ij} \cdot D'_{ij}$ is invariant under the permutations $\pi \in \Pi$ of A, and therefore non-random.

5. It follows from the previous two items that

$$\frac{1}{|\mathcal{E}|} \sum_{(i,j)\in\mathcal{E}} \frac{1}{P_{ij}} D_{ij} \cdot D'_{ij} = \sum_d d \cdot d'$$

with probability 1. The claim of the proposition follows, by collecting terms.

□

Proof of Proposition 2:

- By the first order condition for the least squares estimator,

$$\widehat{\gamma}_j = \left(\sum_{i \in \mathcal{I}} D_{ij} \cdot D'_{ij} \right)^{-1} \cdot \left(\sum_{i \in \mathcal{I}} Y_{ij} \cdot D_{ij} \right).$$

- By definition of $\mathcal{D}_j$,

$$\left(\sum_{i \in \mathcal{I}} D_{ij} \cdot D'_{ij} \right) = \left(\sum_{d \in \mathcal{D}_j} d \cdot d' \right).$$

  By assumption, the right-hand side is invertible for all j.

- By equivariance of d_{ij} (Assumption 2), and because $\pi(j) = j$ for all $j \in \mathcal{J}$ and $\pi(\mathcal{I}) = \mathcal{I}$ (Assumption 3), the set $\mathcal{D}_j$ is invariant under permutations. Therefore the right hand side is non-random.

- Consider next the second term in the expression for $\widehat{\gamma}_j$. By the structural relationship and the exclusion restriction, we can rewrite $Y_{ij} \cdot D_{ij} = Y_{ij}^{D_{ij}} \cdot D_{ij}$.

- By Assumption 2, by the assumed structure of Π (Assumption 3), and by the definition of $\mathcal{D}_j$, $D_{\pi(i)j}$ is a uniform draw from $\mathcal{D}_j$, and therefore the same holds for D_{ij}. Because $|\mathcal{D}_j| = |\mathcal{I}|$, we thus get

$$E[Y_{ij} \cdot D_{ij} | f] = \frac{1}{|\mathcal{I}|} \sum_{d \in \mathcal{D}_j} Y_{ij}^d \cdot d.$$

  The first claim of the proposition follows.

- Equality of γ and β is immediate from their definitions, if $\mathcal{D}_j = \mathcal{D}$.

$\square$

Proof of Proposition 3: Suppose that, under the null, the sampling distribution of $\widehat{\beta}$ is the same as the distribution of $\widehat{\beta}_\pi$, for a uniform draw of $\pi \in \Pi$. Under this assumption, the claim of the proposition follows. Under the null, $\widehat{\beta}$ is uniform over $\{\widehat{\beta}_\pi : \pi \in \Pi\}$; a permutation π' satisfies $p_{\pi'} \leq \alpha$ iff at most $\alpha|\Pi|$ permutations give

$\widehat{\beta}_\pi \geq \widehat{\beta}_{\pi'}$, and the number of such π''s is at most $\lfloor \alpha|\Pi| \rfloor \leq \alpha|\Pi|$, so $P(p \leq \alpha) \leq \alpha$.

It remains to show that our initial assumption was correct, so that the sampling distribution of $\widehat{\beta}$ is indeed the same as the distribution of $\widehat{\beta}_\pi$, for a uniform draw of $\pi \in \Pi$:

1. The group structure of Π implies that
$$\mathcal{A} = \{A_\pi :\ \pi \in \Pi\} = \{A^1_\pi :\ \pi \in \Pi\}.$$
We can therefore obtain the sampling distribution of $\widehat{\beta}$ by considering a uniform random draw of a permutation $\pi \in \Pi$, the counterfactual network $\tilde{A}^1 = A^1_\pi$, and the corresponding counterfactual realizations of $(\tilde{D}_{ij}, \tilde{P}_{ij}, \tilde{Y}_{ij})$, and calculating the corresponding counterfactual estimate $\tilde{\beta}$.

2. By equivariance (Assumption 2), the counterfactual $\tilde{D}_{ij}$ for the permutation π satisfies
$$\tilde{D}_{ij} = d_{ij}(A^1_\pi) = d_{\pi(i)\pi(j)}(A^1) = D_{\pi(i)\pi(j)}.$$

3. By Lemma 2, the counterfactual $\tilde{P}_{ij}$ for the permutation π satisfies
$$\tilde{P}_{ij} = p_{ij}(\tilde{D}_{ij}) = p_{ij}(D_{\pi(i)\pi(j)}) = p_{\pi(i)\pi(j)}(D_{\pi(i)\pi(j)}) = P_{\pi(i)\pi(j)}.$$

4. Under the null hypothesis, the counterfactual $\tilde{Y}_{ij}$ are the same as the realized Y_{ij}.

5. Collecting the three preceding claims, we thus get that under the permutation π of the entries of A^1, (Y_{ij}, D_{ij}, P_{ij}) is replaced by $(Y_{ij}, D_{\pi(i)\pi(j)}, P_{\pi(i)\pi(j)})$. The counterfactual estimate $\tilde{\beta}$ is therefore equal to $\widehat{\beta}_\pi$, as defined in the statement of the proposition.

6. Since π is a uniform random draw from Π, the claim follows.

The claim for $\widehat{\gamma}$ and $\widehat{\gamma}_\pi$ follows by the same argument, once we note that under Assumption 3, $d_{ij}(A_\pi) = D_{\pi(i)j}$, for $(i,j) \in \mathcal{E}$. $\square$

Proof of Proposition 5: Denote $b_o = \frac{m_o}{m}\widehat{\beta}_o$, $\bar{b}_o = E[b_o|f]$, and $\epsilon_o = b_o - \bar{b}_o$. Since $\widehat{\beta} = \sum_o b_o = \mathbf{1}'\mathbf{b}$, we have $\frac{m_o}{m}\widehat{\beta}_o - \frac{\widehat{\beta}}{N} = b_o - \frac{1}{N}\mathbf{1}'\mathbf{b} = (P\mathbf{b})_o$, where $P = I - \frac{1}{N}\mathbf{1}\mathbf{1}'$ is the $N \times N$ centering matrix. Hence

$$\widehat{V}_\beta = \tfrac{N}{N-1}\,\mathbf{b}'P\mathbf{b}.$$

Since P is symmetric and idempotent ($P^2 = P$), we can write

$$\begin{aligned}\tfrac{N-1}{N} \cdot E[\widehat{V}_\beta|f] &= E[\mathbf{b}'P\mathbf{b}|f] \\ &= \bar{\mathbf{b}}'P\bar{\mathbf{b}} + E[\epsilon'P\epsilon|f] \\ &= \|P\bar{\mathbf{b}}\|^2 + \mathrm{tr}(P\Sigma),\end{aligned}$$

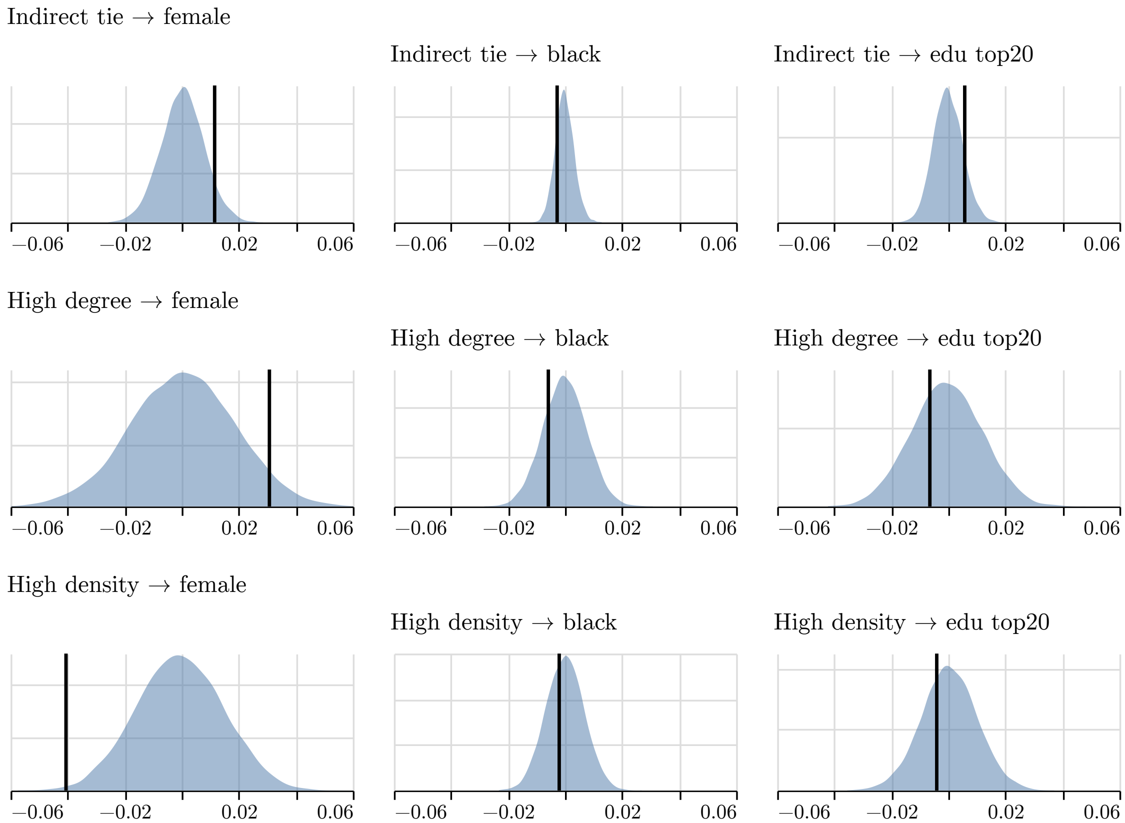
where $\Sigma = \mathrm{diag}(\mathrm{Var}(\epsilon_o|f))$, using the independence of offices. Since $P_{oo} = 1 - \frac{1}{N} = \frac{N-1}{N}$ for all o, we have $\mathrm{tr}(P\Sigma) = \frac{N-1}{N}\sum_o \mathrm{Var}(\epsilon_o|f)$, and therefore

$$E[\widehat{V}_\beta|f] = \tfrac{N}{N-1}\|P\bar{\mathbf{b}}\|^2 + \sum_o \mathrm{Var}\left(\tfrac{m_o}{m}\widehat{\beta}_o \mid f\right).$$

By independence across offices, $\sigma^2_\beta = \mathrm{Var}(\widehat{\beta}|f) = \mathrm{Var}(\sum_o b_o|f) = \sum_o \mathrm{Var}(b_o|f)$. Since $\|P\bar{\mathbf{b}}\|^2 \geq 0$, it follows that $E[\widehat{V}_\beta|f] \geq \sigma^2_\beta$. The claim for $\widehat{V}_\gamma$ follows by the same argument. □

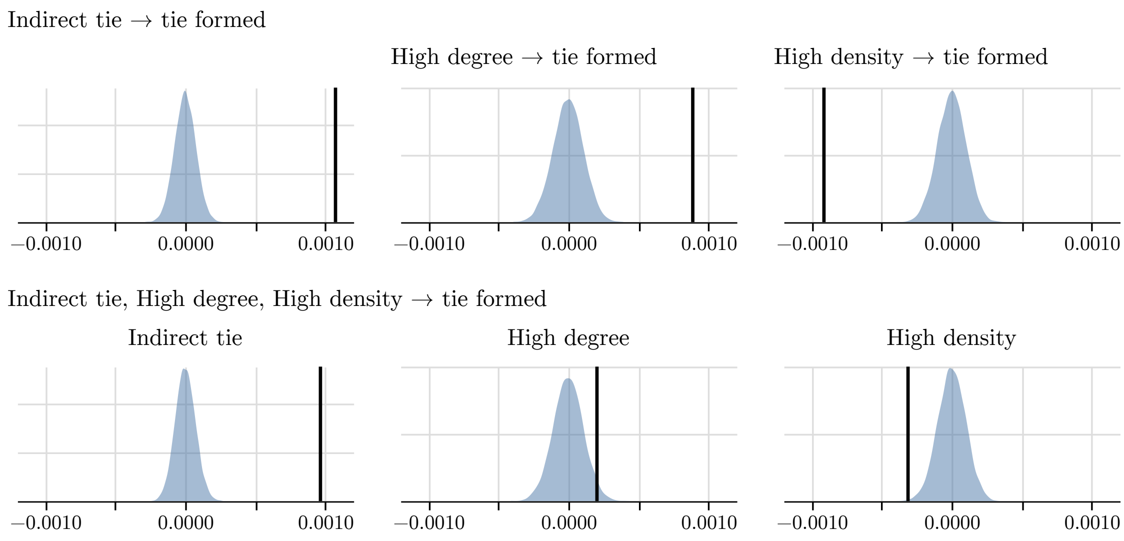
B Additional plots

Figure 1: Permutation distribution of placebo estimates



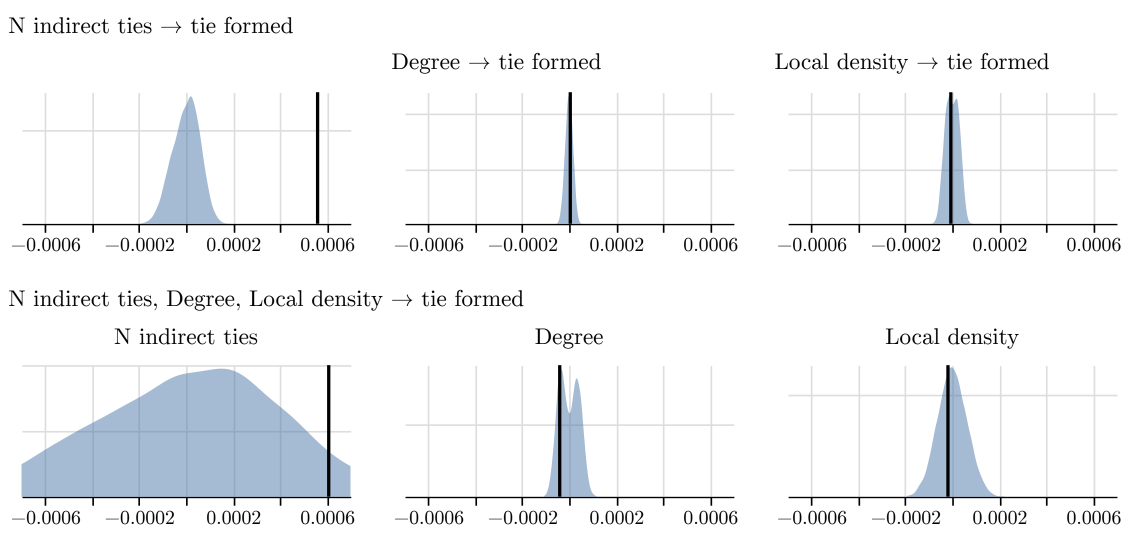
Notes: These figures show the distribution of counterfactual estimates of placebo regressions under 10,000 permutations of new hires within offices. The black lines indicate the actual point estimate. The quantile of the point estimate in the randomization distribution corresponds to the p-values shown in Table 4.

Figure 2: Permutation distribution of estimates



Notes: These figures show the distribution of counterfactual estimates under 10,000 permutations of new hires within offices. The black lines indicate the actual point estimate. The quantile of the point estimate in the randomization distribution corresponds to the p-values shown in Table 5.

Figure 3: Permutation distribution of estimates, continuous regressors



Notes: These figures show the distribution of counterfactual estimates under 10,000 permutations of new hires within offices. The black lines indicate the actual point estimate. The quantile of the point estimate in the randomization distribution corresponds to the p-values shown in Table 6.

C Additional discussion

Review: Causal inference in RCTs The framework as stated in Section 2 fairly abstract. To fix ideas and to motivate our estimands and estimators below, it is useful to relate this framework to the conventional setting of causal inference in randomized controlled trials with binary treatments (Imbens and Rubin, 2015).

This conventional setting describes the effect of binary treatments in the absence of interference (spillovers). There is a vector of outcomes Y, and a vector of treatments D, and the two are related by the causal relationship $Y = g(D)$. The structural function g (like the function f in the network setting) describes potential (or counterfactual) outcomes. Under the assumption of no interference (also known as stable unit treatment effects, which is a form of exclusion restriction), Y_j only depends on D_j, and it is common to define the potential outcome notation $Y_j^d = g(D_1, \ldots, D_{j-1}, d, D_{j+1}, \ldots, D_n)$. A common object of interest is the (sample) average treatment effect,

$$SATE = \tfrac{1}{n} \sum_j \left(Y_j^1 - Y_j^0 \right).$$

Note that this definition of the $SATE$ conditions on unobserved heterogeneity by conditioning on the function g. The potential outcomes Y_j^d are thus non-random, by definition; the only potential source of randomness is the treatment assignment D.

Under exogenous random assignment of D, the $SATE$ can be estimated without bias using inverse probability weighting (IPW). We can write the IPW estimator of the SATE as follows:

$$\widehat{Y}_j^d = Y_j \cdot \frac{\mathbf{1}(D_j = d)}{P(D_j = d)}, \qquad \widehat{SATE} = \tfrac{1}{n} \sum_j \left(\widehat{Y}_j^1 - \widehat{Y}_j^0 \right).$$

Suppose that the support condition $0 < P(D_j = 1) < 1$ holds for all j. Under this condition, it is easy to see that the IPW estimator is unbiased, $E\left[\widehat{SATE}\right] = SATE$.

We can furthermore perform randomization inference on $\widehat{SATE}$ and related objects: Randomization inference, for testing the null of zero treatment effects, assumes that $Y_j^1 = Y_j^0$ for all j. Under the no interference condition and this null hypothesis, $g(D)$ is known for all counterfactual treatment assignments D. Because the sampling distribution of D itself is known, the sampling distribution of any statistic based on D and Y is therefore known.

If heterogeneity or moderators of treatment effects are of interest, one might furthermore consider variants of the $SATE$ for subsets $\mathcal{I}$ of individuals i, where these subsets might be defined based on pre-determined individual covariates.

References

Imbens, Guido W and Donald B Rubin (2015), *Causal inference in statistics, social, and biomedical sciences*, Cambridge University Press.