Randomization tests for peer effects in group formation experiments

Panos Toulis
panos.toulis@chicagobooth.edu

Econometrics and Statistics
University of Chicago, Booth School of Business

Wharton School, University of Pennsylvania, 2023

Introduction

Standard causal inference assumes no interference;
i.e., a unit's treatment cannot affect other units.

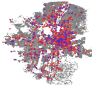
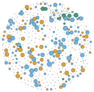
This describes a simple, static world.

In many interesting problems, units interact in a complex way.
e.g., spillovers, peer effects, contagion, equilibrium effects.

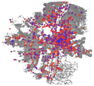
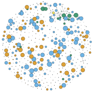
Pervasive in most social studies.

New methods and tools are needed. Many applications:
e.g., policy making, marketplace algorithms, climate science, healthcare.

Some applications

Peer effects in households	(Basse et al, 2019)	
Crime spillovers in a city	(Puelz et al, 2021)	
Peer effects in dorms (**this talk**)	(Basse et al, 2023+)	
Tax audit spillovers across firms	(ongoing)	

Some applications

Peer effects in households	(Basse et al, 2019)	
Crime spillovers in a city	(Puelz et al, 2021)	
Peer effects in dorms (**this talk**)	(Basse et al, 2023+)	
Tax audit spillovers across firms	(ongoing)	

- The tax application, in particular, is a complex and sensitive setting with $>$ 400,000 units sharing 10mil. connections.
- Procedures that are fast and finite-sample valid are highly desirable.

State-of-art

Current approaches tend to be heavily model-based.

In complex domains, this causes problems with inference and even with identification (e.g., "Perils of peer effects" by J. Angrist).

Randomization tests are nonparametric procedures that are **model-agnostic** and **finite-sample exact**.

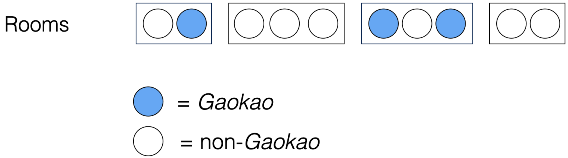
However, they tend to be limited in scope.

A lot of recent research work in extending the scope of randomization tests to complex domains. I will present such a line of work today.

Randomization-based and model-based methods can be synergetic.

Motivation: Peer effects in uni dorms (Li et al, 2019)

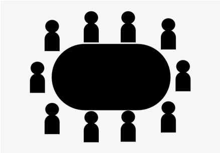
- Consider an experiment where students in a Chinese university are randomly assigned into dorm rooms.
- Each student has a binary attribute depending on whether they passed an entrance exam ($A_i = 1$) known as *Gaokao*.

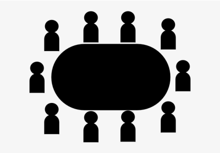


1. Is there an effect on academic outcomes from being roommates with a Gaokao student?
2. Can we test this *via simple permutations*?

More motivation: Interfirm relationships

- Cai and Szeidl (2017, QJE) randomized CEOs into working groups and tracked various firm performance metrics.





- Complex design (multi-stage, ∼1,300 firms) but group formation was exchangeable conditional on firm size, sector, and location.
- Here, the salient attribute is 3-dim, $A_i = (\text{size}_i, \text{sector}_i, \text{region}_i)$. Our methodology can still be applied. (coming later)

Business applications

Randomized group formation is especially interesting for business and management; e.g.,

- Diffusion of business practices across random groupings of African manufacturing firms (Fafchamps and Quinn, 2018).
- Random groupings of freshmen at USAF Academy to 'optimize' academic performance (Carrell et al, 2013).
- Peer effects in the workplace (Cornelissen et al, 2017).
- Random groupings in professional golf tournaments (Guryan et al, 2009).

Setup (Gaokao experiment)

- Units (students) indexed by $i = 1, 2 \ldots, N$.
- K rooms of max size $M + 1$.
- $\mathbf{L} = (L_1, \ldots, L_N) \in \{1, \ldots, K\}^N$, room assignment.
- $\mathbf{A} = (A_1, \ldots, A_N) \in \{0, 1\}^N$, binary attributes.
- $\mathbf{Y} = (Y_1, \ldots, Y_N) \in \mathbb{R}^N$, outcomes (e.g., grade improvement).
- Will use $(\mathbf{Y}^*, \mathbf{L}^*)$ for *counterfactual* outcomes/treatments.

Moreover,

- $P(\mathbf{L})$ is known and under our control (experimental study).
  e.g., completely randomized given fixed room compositions.
- $Y_i(\boldsymbol{\ell})$, potential outcome of i under room assignment $\boldsymbol{\ell}$.
- We make the typical consistency assumption: $Y_i = Y_i(\mathbf{L})$ for every i.

Potential outcome, $Y_i(\boldsymbol{\ell})$

- In classical causal inference ("Rubin Causal Model"), every unit i has only two potential outcomes, namely "$Y_i(0), Y_i(1)$" for control and treatment, respectively.
- With "$Y_i(\boldsymbol{\ell})$" we allow the treatment of other units to affect i's outcome. This is known as **interference**.
- Under interference, a unit is exposed to "something more" than just its own room assignment, perhaps a sum effect from the attributes of its roommates, and/or neighbors, etc.
- However, "$Y_i(\boldsymbol{\ell})$" may take ostensibly K^N possible values.

Potential outcome, $Y_i(\boldsymbol{\ell})$

- In classical causal inference ("Rubin Causal Model"), every unit i has only two potential outcomes, namely "$Y_i(0), Y_i(1)$" for control and treatment, respectively.
- With "$Y_i(\boldsymbol{\ell})$" we allow the treatment of other units to affect i's outcome. This is known as **interference**.
- Under interference, a unit is exposed to "something more" than just its own room assignment, perhaps a sum effect from the attributes of its roommates, and/or neighbors, etc.
- However, "$Y_i(\boldsymbol{\ell})$" may take ostensibly K^N possible values.

▷ We need to put some structure and reduce this space.

Effective treatment

- A common approach is to model interference on i through a pre-defined function $w_i()$:

$$W_i = w_i(\mathbf{L}) \in \mathbb{W}.$$

Potential outcomes are assumed to be a function of W_i :

Assumption 1.

$$Y_i(\boldsymbol{\ell}) = Y_i(\boldsymbol{\ell}') \text{ for all } \boldsymbol{\ell}, \boldsymbol{\ell}', i \text{ if } w_i(\boldsymbol{\ell}) = w_i(\boldsymbol{\ell}').$$

- W_i is the known as the *treatment exposure*; e.g., (Verbitsky and Raudenbush, 2004), (Hong and Raudenbush, 2006), (T. and Kao, 2013), (Aronow and Samii, 2017), (Athey et al, 2018), (Basse et al, 2019).
- Also known as the *effective treatment* (Manski, 2013).
- $\mathbb{W}$ may be arbitrary but it is typically much smaller than K^N.

Effective treatment — Examples

- The definition of $w_i(\cdot)$ usually depends on the domain and subject-matter experts; e.g.,
  - $W_i = \sum_{j \neq i} A_j 1\{L_i = L_j\} = \#$of Gaokao roommates (**this talk**).
  - $W_i =$% of Gaokao roommates.
  - May depend on covariates, etc.

Effective treatment — Examples

- The definition of $w_i(\cdot)$ usually depends on the domain and subject-matter experts; e.g.,
  - $W_i = \sum_{j \neq i} A_j 1\{L_i = L_j\} = \#$of Gaokao roommates (**this talk**).
  - $W_i =$% of Gaokao roommates.
  - May depend on covariates, etc.

- <u>Notation:</u> Under Assumption 1, we may use "$Y_i^{\omega}(\cdot)$" to denote potential outcomes in the "exposure space":

$$Y_i^{\omega}(\mathsf{w}) := Y_i(\mathbf{L}) \text{ where } \mathsf{w} = w_i(\mathbf{L}).$$

Main hypothesis under interference

A large class of hypotheses under interference may be expressed as:

$$H_0 : Y_i^{\omega}(\mathsf{w}) = Y_i^{\omega}(\mathsf{w}') \text{ for all } i \text{ and } \mathsf{w}, \mathsf{w}' \in \mathbb{W}_0 \subseteq \mathbb{W}.$$

Main hypothesis under interference

A large class of hypotheses under interference may be expressed as:

$$H_0 : Y_i^\omega(\mathsf{w}) = Y_i^\omega(\mathsf{w}') \text{ for all } i \text{ and } \mathsf{w}, \mathsf{w}' \in \mathbb{W}_0 \subseteq \mathbb{W}.$$

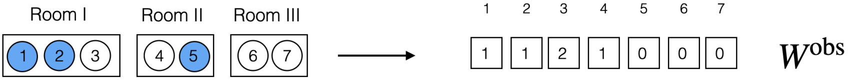
- e.g., $\mathbb{W}_0 = \{0, 1\}$ whereas $\mathbb{W} = \{0, \ldots, \mathsf{M}\}$. This suggests the null

  $$H_0 : Y_i^\omega(0) = Y_i^\omega(1), \text{ for all } i.$$

  That is, there is "no difference in outcomes from having 0 or 1 Gaokao roommate".
- For simplicity, we will focus on the special null above.

Illustration

$$H_0 : Y_i^{\omega}(0) = Y_i^{\omega}(1), \text{ for all } i.$$



- The null hypothesis implies that the outcomes of all units except 3 should be "similar".
- Testing the null, however, is challenging because it is defined in the "exposure space", not the "treatment space".
- Naive randomization/permutation can fail.

Fisher's Randomization Test

Let's start with a simple problem.

If $\mathbb{W}_0 = \mathbb{W}$ then all exposures give identical outcomes under the null.
This is equivalent to the "sharp null" of no effect:

$$H_0 : Y_i^\omega(0) = \ldots = Y_i^\omega(\mathsf{M}) \text{ for all } i.$$

Fisher's Randomization Test

Let's start with a simple problem.

If $\mathbb{W}_0 = \mathbb{W}$ then all exposures give identical outcomes under the null.
This is equivalent to the "sharp null" of no effect:

$$H_0 : Y_i^\omega(0) = \ldots = Y_i^\omega(\mathsf{M}) \text{ for all } i.$$

This can be tested through Fisher's randomization test (Fisher, 1935),

1. Calculate test statistic, $T = t(\mathbf{W}, \mathbf{Y})$; e.g., regression coefficient, ML.
2. pval $= \mathsf{E}[t(\mathbf{W}^*, \mathbf{Y}) > T],\ \mathbf{W}^* = w(\mathbf{L}^*),\ \mathbf{L}^* \sim P$.

Fisher's Randomization Test

Let's start with a simple problem.

If $\mathbb{W}_0 = \mathbb{W}$ then all exposures give identical outcomes under the null. This is equivalent to the "sharp null" of no effect:

$$H_0 : Y_i^\omega(0) = \ldots = Y_i^\omega(\mathsf{M}) \text{ for all } i.$$

This can be tested through Fisher's randomization test (Fisher, 1935),

1. Calculate test statistic, $T = t(\mathbf{W}, \mathbf{Y})$; e.g., regression coefficient, ML.
2. pval $= \mathsf{E}[t(\mathbf{W}^*, \mathbf{Y}) > T],\ \mathbf{W}^* = w(\mathbf{L}^*),\ \mathbf{L}^* \sim P$.

The p-value from FRT is finite-sample exact.
Proof. The null implies $\mathbf{Y}^* = \mathbf{Y}$ a.s. Thus, $t(\mathbf{W}^*, \mathbf{Y}) \overset{H_0}{=} t(\mathbf{W}^*, \mathbf{Y}^*) \overset{d}{=} T$.

An assessment of FRT

Main advantages:

- The test is exact in finite samples. No asymptotics.
- Y-model may be misspecified.(affects power but not validity)
- Robustness: Same answer under transformations of Y.

Common criticism:

- Can only test "strong" hypotheses. (This talk. Also, a lot of related research activity recently).
- Cannot generalize to population.

An assessment of FRT

Main advantages:

- The test is exact in finite samples. No asymptotics.
- Y-model may be misspecified.(affects power but not validity)
- Robustness: Same answer under transformations of Y.

Common criticism:

- Can only test "strong" hypotheses. (This talk. Also, a lot of related research activity recently).
- Cannot generalize to population.

$\triangleright$ Can we use FRTs to test for spillovers?

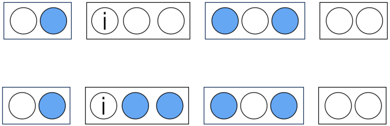
FRT problems under interference

Now, consider testing $H_0 : Y_i^\omega(0) = Y_i^\omega(1)$.

In the FRT, suppose we naively resample $\mathbf{L}^*$ as shown below:

Observed $\mathbf{L}$

Resampled $\mathbf{L}^*$



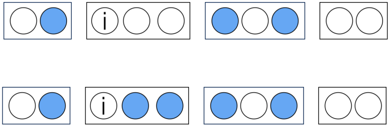
FRT problems under interference

Now, consider testing $H_0 : Y_i^{\omega}(0) = Y_i^{\omega}(1)$.

In the FRT, suppose we naively resample $\mathbf{L}^*$ as shown below:

Observed $\mathbf{L}$

Resampled $\mathbf{L}^*$



- Under $\mathbf{L}$ we observed outcome $Y_i^{\omega}(0)$ for unit i.
- Under $\mathbf{L}^*$, the unit has outcome $Y_i^{\omega}(2)$. But this outcome **cannot be imputed** under the null hypothesis (the null is "weak").
  Thus, the standard FRT is invalid.

A recent development

Recently, a general approach to apply FRTs under interference has been put forward (Aronow, 2012); (Athey et al, 2018); (Basse et al, 2019):

- Let $\mathbf{U} = (U_1, \ldots, U_N) \in \{0, 1\}^N$ denote a subset of units.
- Then, the idea is to run FRT on the subset of the *focal units*:

$$(Y_i, W_i, .. : U_i = 1)$$

under the following requirements:

I. The potential outcomes of *all* focal units should be *imputable* under the null, H_0.
II. The resulting conditional randomization test should be easy to implement. (Only implicit in prior work)

A conditional FRT

Specifically:

1. $P(\mathbf{U}) \sim$ Unif; i.e., pick focal units uniformly at random.
2. Enumerate:

   $$\mathcal{W}_U = \left\{ \mathbf{W}' : Y_i^{\omega}(W_i') \text{ imputable under } H_0 \text{ for all } i \text{ with } U_i = 1. \right\}$$

3. Define test statistic only on data from focals ($U_i = 1$).
4. Run a *conditional* FRT by resampling from:

$$P(\mathbf{W}^* \mid \mathbf{U}) \propto 1\{\mathbf{W}^* \in \mathcal{W}_U\} P(\mathbf{W}^*). \tag{1}$$

A conditional FRT

Specifically:

1. $P(\mathbf{U}) \sim$ Unif; i.e., pick focal units uniformly at random.
2. Enumerate:

$$\mathcal{W}_U = \{ \ \mathbf{W}' : Y_i^{\omega}(W_i') \text{ imputable under } H_0 \text{ for all } i \text{ with } U_i = 1.\}$$

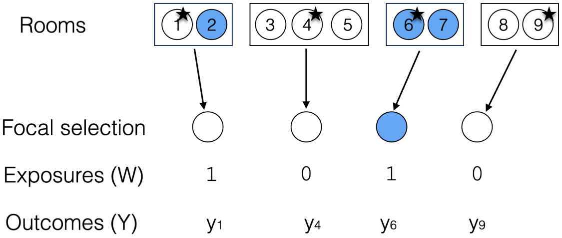
3. Define test statistic only on data from focals ($U_i = 1$).
4. Run a *conditional* FRT by resampling from:

$$P(\mathbf{W}^* \mid \mathbf{U}) \propto 1\{\mathbf{W}^* \in \mathcal{W}_U\} P(\mathbf{W}^*). \tag{1}$$

- This construction satisfies Condition I of imputability (Step 2).
- However, distribution (1) is usually very hard to sample from; cf. Puelz et al (2021) connects this to graph clique decomposition (NP-hard).
- In particular, (1) does not generally imply a permutation test.

Permutation test for spillovers?

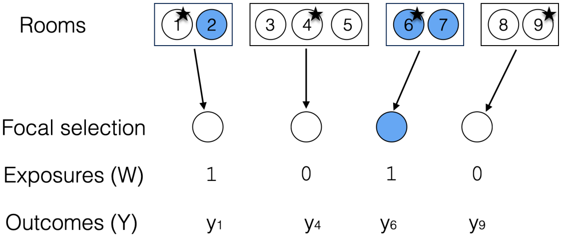
- $H_0 : Y_i^\omega(0) = Y_i^\omega(1)$.



- The conditional FRT requires enumerating *all* assignments for which the focal units are exposed to $\{0, 1\}$. This grows exponentially in N.

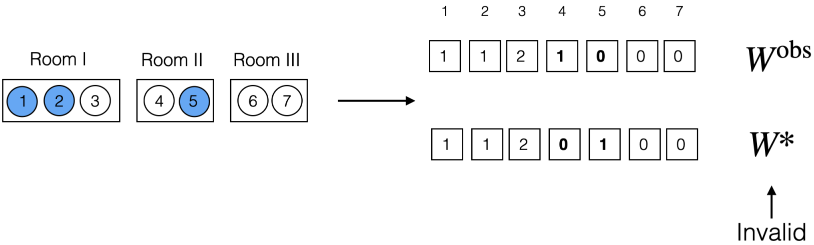
Permutation test for spillovers?

- $H_0 : Y_i^{\omega}(0) = Y_i^{\omega}(1)$.



- The conditional FRT requires enumerating *all* assignments for which the focal units are exposed to $\{0, 1\}$. This grows exponentially in N.
- Couldn't we just run a permutation test between Y and W on the focal units shown above?

Naive permutation fails



- In this example, we permute the exposures of units 4 and 5.
- However, the resulting $\mathbf{W}^*$ is invalid. It cannot be generated from the design since it would require that 1,2, and 5 (Gaokao students) all have exactly one Gaokao roommate.

The problem, in summary

To summarize:

- Room assignment $\mathbf{L}$ according to known design, $P(\mathbf{L})$.
- Eff. treatment due to interference: $\mathbf{W} = w(\mathbf{L}) = (W_1, \ldots, W_N)$.
- $H_0 : Y_i^\omega(\mathsf{w}) = Y_i^\omega(\mathsf{w}')$ for all $\mathsf{w}, \mathsf{w}' \in \mathbb{W}_0$.

▷ Can we test H_0 via permutations on (a subvector of) $\mathbf{W}$?

Main theorem

Theorem

Let $\mathbf{U} = u(\mathbf{L})$ *be the focal selection function. Let* $\mathsf{S}_{A,U}$ *be the permutation subgroup that leaves* **A** *(attributes) and* **U** *(focals) unchanged. Suppose:*

- (a) $P(\mathbf{L}) = P(\pi\mathbf{L})$ *for all* $\pi \in \mathsf{S}_{A,U}$*.*
- (b) $w(\mathbf{L})$ *is* **equivariant** *with respect to* $\mathsf{S}_{A,U}$*; i.e.,* $w(\pi\mathbf{L}) = \pi w(\mathbf{L})$*.*
- (c) $u(\mathbf{L})$ *is equivariant with respect to* $\mathsf{S}_{A,U}$*.*

Then, **W** *is uniformly distributed conditional on an orbit generated by* $\mathsf{S}_{A,U}$*.*

- The theorem shows that the procedure that permutes the exposures of focal units *stratified by attribute* (i.e., permutations in $\mathsf{S}_{A,U}$) is finite-sample valid.
- Note the "interaction" between (a) design, (b) exposure definition, and (c) focal unit selection.

Proof sketch

Let $\mathcal{O} = \{\pi \mathbf{W}^{\text{obs}} : \pi \in \mathsf{S}_{A,U}\}$ = orbit generated by observed exposure.
Goal is to show that $\mathbf{W} \mid \mathbf{U}$ is uniform in $\mathcal{O}$ (implies permutation test!)

Proof sketch

Let $\mathcal{O} = \{\pi \mathbf{W}^{\mathsf{obs}} : \pi \in \mathsf{S}_{A,U}\}$ = orbit generated by observed exposure.
Goal is to show that $\mathbf{W} \mid \mathbf{U}$ is uniform in $\mathcal{O}$ (implies permutation test!)

$$
\begin{aligned}
P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \pi\mathbf{L}) &= \ldots \\
&= 1\{w(\pi\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\pi\mathbf{L})\} && \textsf{L fully determines } \mathbf{W}, \mathbf{U} \\
&= 1\{\pi w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = \pi u(\mathbf{L})\} && \text{equivariance conditions (b),(c)} \\
&= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\pi^{-1}\mathbf{U} = u(\mathbf{L})\} && \text{orbit property} \\
&= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\mathbf{L})\} && \pi^{-1}\mathbf{U} = \mathbf{U} \text{ since } \pi \in \mathsf{S}_{A,U} \\
&= P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \mathbf{L}).
\end{aligned}
$$

Proof sketch

Let $\mathcal{O} = \{\pi \mathbf{W}^{\text{obs}} : \pi \in \mathsf{S}_{A,U}\}$ = orbit generated by observed exposure.
Goal is to show that $\mathbf{W} \mid \mathbf{U}$ is uniform in $\mathcal{O}$ (implies permutation test!)

$P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \pi \mathbf{L}) = \ldots$

$= 1\{w(\pi \mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\pi \mathbf{L})\}$ — $\mathbf{L}$ fully determines $\mathbf{W}, \mathbf{U}$

$= 1\{\pi w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = \pi u(\mathbf{L})\}$ — equivariance conditions (b),(c)

$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\pi^{-1}\mathbf{U} = u(\mathbf{L})\}$ — orbit property

$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\mathbf{L})\}$ — $\pi^{-1}\mathbf{U} = \mathbf{U}$ since $\pi \in \mathsf{S}_{A,U}$

$= P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \mathbf{L}).$

Proof sketch

Let $\mathcal{O} = \{\pi \mathbf{W}^{\text{obs}} : \pi \in \mathsf{S}_{A,U}\}$ = orbit generated by observed exposure.
Goal is to show that $\mathbf{W} \mid \mathbf{U}$ is uniform in $\mathcal{O}$ (implies permutation test!)

$$P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \pi \mathbf{L}) = \ldots$$

$$= 1\{w(\pi \mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\pi \mathbf{L})\} \qquad \mathbf{L} \text{ fully determines } \mathbf{W}, \mathbf{U}$$

$$= 1\{\pi w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = \pi u(\mathbf{L})\} \qquad \text{equivariance conditions (b),(c)}$$

$$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\pi^{-1}\mathbf{U} = u(\mathbf{L})\} \qquad \text{orbit property}$$

$$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\mathbf{L})\} \qquad \pi^{-1}\mathbf{U} = \mathbf{U} \text{ since } \pi \in \mathsf{S}_{A,U}$$

$$= P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \mathbf{L}).$$

Proof sketch

Let $\mathcal{O} = \{\pi \mathbf{W}^{\text{obs}} : \pi \in \mathsf{S}_{A,U}\}$ = orbit generated by observed exposure.
Goal is to show that $\mathbf{W} \mid \mathbf{U}$ is uniform in $\mathcal{O}$ (implies permutation test!)

$P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \pi \mathbf{L}) = \ldots$

$= 1\{w(\pi \mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\pi \mathbf{L})\}$ — $\mathbf{L}$ fully determines $\mathbf{W}, \mathbf{U}$

$= 1\{\pi w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = \pi u(\mathbf{L})\}$ — equivariance conditions (b),(c)

$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\pi^{-1}\mathbf{U} = u(\mathbf{L})\}$ — orbit property

$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\mathbf{L})\}$ — $\pi^{-1}\mathbf{U} = \mathbf{U}$ since $\pi \in \mathsf{S}_{A,U}$

$= P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \mathbf{L}).$

Proof sketch

Let $\mathcal{O} = \{\pi \mathbf{W}^{\text{obs}} : \pi \in \mathsf{S}_{A,U}\}$ = orbit generated by observed exposure.
Goal is to show that $\mathbf{W} \mid \mathbf{U}$ is uniform in $\mathcal{O}$ (implies permutation test!)

$P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \pi \mathbf{L}) = ...$

$$= 1\{w(\pi\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\pi\mathbf{L})\} \qquad \mathbf{L} \text{ fully determines } \mathbf{W}, \mathbf{U}$$

$$= 1\{\pi w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = \pi u(\mathbf{L})\} \qquad \text{equivariance conditions (b),(c)}$$

$$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\pi^{-1}\mathbf{U} = u(\mathbf{L})\} \qquad \text{orbit property}$$

$$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\mathbf{L})\} \qquad \pi^{-1}\mathbf{U} = \mathbf{U} \text{ since } \pi \in \mathsf{S}_{A,U}$$

$$= P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \mathbf{L}).$$

Proof sketch

Let $\mathcal{O} = \{\pi \mathbf{W}^{\text{obs}} : \pi \in \mathsf{S}_{A,U}\}$ = orbit generated by observed exposure.
Goal is to show that $\mathbf{W} \mid \mathbf{U}$ is uniform in $\mathcal{O}$ (implies permutation test!)

$P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \pi\mathbf{L}) = ...$

$= 1\{w(\pi\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\pi\mathbf{L})\}$	$\mathbf{L}$ fully determines $\mathbf{W}, \mathbf{U}$
$= 1\{\pi w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = \pi u(\mathbf{L})\}$	equivariance conditions (b),(c)
$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\pi^{-1}\mathbf{U} = u(\mathbf{L})\}$	orbit property
$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\mathbf{L})\}$	$\pi^{-1}\mathbf{U} = \mathbf{U}$ since $\pi \in \mathsf{S}_{A,U}$
$= P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \mathbf{L})$.	

Proof sketch

Let $\mathcal{O} = \{\pi \mathbf{W}^{\text{obs}} : \pi \in \mathsf{S}_{A,U}\}$ = orbit generated by observed exposure.
Goal is to show that $\mathbf{W} \mid \mathbf{U}$ is uniform in $\mathcal{O}$ (implies permutation test!)

$P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \pi \mathbf{L}) = \ldots$

$= 1\{w(\pi \mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\pi \mathbf{L})\}$ — $\mathbf{L}$ fully determines $\mathbf{W}, \mathbf{U}$

$= 1\{\pi w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = \pi u(\mathbf{L})\}$ — equivariance conditions (b),(c)

$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\pi^{-1}\mathbf{U} = u(\mathbf{L})\}$ — orbit property

$= 1\{w(\mathbf{L}) \in \mathcal{O}\}\ 1\{\mathbf{U} = u(\mathbf{L})\}$ — $\pi^{-1}\mathbf{U} = \mathbf{U}$ since $\pi \in \mathsf{S}_{A,U}$

$= P(\mathbf{W} \in \mathcal{O}, \mathbf{U} \mid \mathbf{L})$.

- From Bayes and condition (a), we obtain

$$P(\pi \mathbf{L} \mid \mathbf{W} \in \mathcal{O}, \mathbf{U}) = P(\mathbf{L} \mid \mathbf{W} \in \mathcal{O}, \mathbf{U}).$$

  That is, the design "maintains" its invariance even *conditional on focal selection* within the subspace $\mathsf{S}_{A,U}$.

- Equivariance of $w(\mathbf{L})$ then implies the theorem.

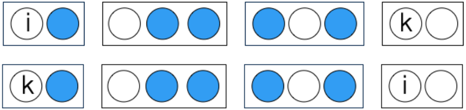
Application of theory — Condition (b)

Recall that as exposure we use

$$w_i(\mathbf{L}) = f(\{A_j : L_j = L_i, j \neq i\}) \quad \text{(for some known } f\text{)}.$$

Then, $w(\mathbf{L})$ satisfies equivariance (b).

To see this, if we swap the rooms of i, k (with $A_i = A_k$), then the exposures of i, k are transposed but all other exposures are unchanged.



Application of theory — Condition (c)

Suppose we simply define the focal units as:

$$\mathbf{U} = u(\mathbf{L}) = 1\{w(\mathbf{L}) \in \mathbb{W}_0\};$$

i.e., "Focus on units that are exposed to the null exposure levels."

Then, $u(\mathbf{L})$ satisfies equivariance (c). To see this:

Application of theory — Condition (c)

Suppose we simply define the focal units as:

$$\mathbf{U} = u(\mathbf{L}) = 1\{w(\mathbf{L}) \in \mathbb{W}_0\};$$

i.e., "Focus on units that are exposed to the null exposure levels."

Then, $u(\mathbf{L})$ satisfies equivariance (c). To see this:

$$u(\pi\mathbf{L}) = 1\{w(\pi\mathbf{L}) \in \mathbb{W}_0\} = 1\{\pi w(\mathbf{L}) \in \mathbb{W}_0\} = \pi 1\{w(\mathbf{L}) \in \mathbb{W}_0\} = \pi u(\mathbf{L}).$$

Condition (a): Design symmetry

The last condition is $P(\pi \mathbf{L}) = P(\mathbf{L})$. This is very mild under reasonable randomized designs; e.g.,

- Completely randomized design with a fixed number of units assigned to each room.
- Stratified randomized design with a fixed number of units assigned to each pair (room, attribute).
- ... etc

Re-analysis of (Li et al, 2019)

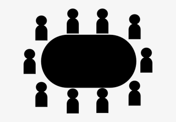
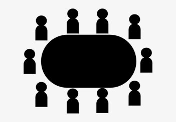
- Hypothesis of no difference between "0 or 3 Gaokao roommates" denoted as "$H_0^{0,3}$".
- Also test within subgroups: (0)=non-Gaokao; (1)=Gaokao students.

	p-value	estimate	confidence interval
$H_0^{0,3}$	0.04	−0.31	(−0.67, −0.02)
$H_0^{0,3}(0)$	0.02	−0.37	(−0.73, −0.05)
$H_0^{0,3}(1)$	0.23	−0.28	(−0.81, 0.12)

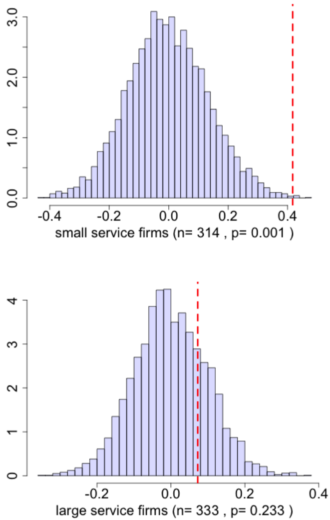
- Main difference with the design-based analysis of Li et al (2019) is for the subgroup of Gaokao students (they find strong significance).
- Could be explained by the asymptotic approximations in (Li et al, 2019) , which may be unwarranted given the small sample size (see paper for simulation study).

Re-analysis of (Cai and Szeidl, 2017)

- The design randomized CEOs into working groups that met monthly for a year, and then tracked various firm performance metrics.

- The authors studied heterogeneity in "direct effects". They showed that larger firms benefited more from the meetings.
- Our method can analyze *heterogeneity in peer effects* by testing the global null within subgroups.
- Regression specifications cannot easily capture peer effect heterogeneity due to model saturation.



- Peer effects only on small service firms.

Concluding remarks

- Recently, conditional randomization tests have been devised to test complex causal effects under interference.
- However, conditional FRTs are computationally demanding.
- Permutation tests are computationally simple, but conditional FRTs are not always permutation tests.
- We proved *sufficient* theoretical conditions to make the connection. But, are these conditions necessary?
- What about other designs? (e.g., two-stage, cluster)

Thank you!

- (*) Basse, Ding, Feller, T., "Randomization tests for group formation experiments" (2023+, *cond. accept, Econometrica*)
- Puelz, Basse, Feller, T., "A graph-theoretic approach to randomization tests of causal effects under interference" (JRSSB, 2021)
- Basse, Feller, T., "Randomization tests of causal effects under interference" (Biometrika, 2019)