The "Graph of Why": Randomization Inference for Spillover Effects

Panagiotis (Panos) Toulis
panos.toulis@chicagobooth.edu

Econometrics and Statistics
University of Chicago, Booth School of Business

JSM 2023, Toronto, Canada

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.
—spillovers, peer effects, contagion, equilibrium effects, etc.

Pervasive in most social studies. Can either be a nuisance to be addressed by design, or the quantity of interest.

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

Overview

Current approaches tend to be heavily model-based.

In complex domains, this causes problems with inference and even with identification.

Randomization tests (e.g., permutations) are nonparametric procedures that are **model-agnostic** and **finite-sample exact** under certain conditions.

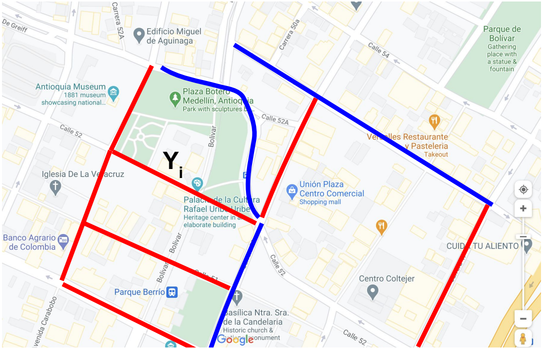
However, they have been limited in scope.

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.

Motivation: Crime spillovers in Medellin, Colombia

(Collazos, 2019), (Puelz et al, 2021)

Crime spillovers from nearby **treated streets** on **control streets**?



treatment = increased policing; control = baseline policing.

- What is a proper definition of a spillover effect?
- How to estimate it?

Causal Inference

Suppose data $\{(Y_i, Z_i, X_i)\}$, $i = 1, \ldots, N$.

$\mathbf{Y}$ = outcomes, $\mathbf{Z}$ = treatments, $\mathbf{X}$ = covariates (features).

$Y_i(\mathbf{z})$ is the *potential outcome* of unit i under treatment $\mathbf{z} \in \{0, 1\}^N$,

Consistency assumption: $\mathbf{Y} = \mathbf{Y}(\mathbf{Z})$.

Outcomes are only a function of treatment.Variation only comes from treatment assignment ("design-based inference"). See, e.g., (Abadie et al, 2020).

No Interference

In classical causal inference, every unit i has only two potential outcomes, namely "$Y_i(0), Y_i(1)$" for control and treatment, respectively.

This is equivalent to assuming that

$$Y_i(\mathbf{z}) = Y_i(\mathbf{z}') \text{ for all } \mathbf{z}, \mathbf{z}', i \text{ if } z_i = z_i' \; - \text{"SUTVA" (Rubin, 1974).}$$

No Interference

In classical causal inference, every unit i has only two potential outcomes, namely "$Y_i(0), Y_i(1)$" for control and treatment, respectively.

This is equivalent to assuming that

$$Y_i(\mathbf{z}) = Y_i(\mathbf{z}') \;\text{ for all } \mathbf{z}, \mathbf{z}', i \text{ if } z_i = z_i' \;-\text{"SUTVA" (Rubin, 1974).}$$

However, in many problems there is treatment interference.

Under interference, a unit is exposed to "something more" than Z_i, a sum effect from the entire population treatment, $\mathbf{Z}$.

Think of a vaccine trial. A control unit (unvaccinated) is still "protected" by treated units (vaccinated) in proximity.

Effective treatments

Under interference, it is popular to use treatment exposures, $f_i(\mathbf{Z}) \in \mathbb{F}$.

Although not necessary, it is useful to think that the exposure is the "effective treatment" (Manski, 2013); i.e.,

Assumption

$$Y_i(\mathbf{z}) = Y_i(\mathbf{z}') \quad \textit{for all } \mathbf{z}, \mathbf{z}', i \textit{ if } f_i(\mathbf{z}) = f_i(\mathbf{z}').$$

Examples of treatment exposure:

- $f_i(\mathbf{z}) = z_i$. Standard, no interference setting.
- $f_i(\mathbf{z}) = \left(z_i, \sum_{j:d(i,j)<d_0} z_j\right)$. "Treatments nearby matter".

  We use this in the Medellin application.
- $f_i(\mathbf{z}) = (z_i, \mathbf{z}_{\text{neighborhood}_i})$.

Effective treatments

Under interference, it is popular to use treatment exposures, $f_i(\mathbf{Z}) \in \mathbb{F}$.

Although not necessary, it is useful to think that the exposure is the "effective treatment" (Manski, 2013); i.e.,

Assumption

$$Y_i(\mathbf{z}) = Y_i(\mathbf{z}') \quad \textit{for all } \mathbf{z}, \mathbf{z}', i \textit{ if } f_i(\mathbf{z}) = f_i(\mathbf{z}').$$

Examples of treatment exposure:

- $f_i(\mathbf{z}) = z_i$. Standard, no interference setting.
- $f_i(\mathbf{z}) = \left(z_i, \sum_{j:d(i,j)<d_0} z_j\right)$. "Treatments nearby matter".

  We use this in the Medellin application.
- $f_i(\mathbf{z}) = (z_i, \mathbf{z}_{\text{neighborhood}_i})$.

Goal: Learn the effect of $f_i(\mathbf{z})$ on outcome Y_i?

Hypotheses for spillovers

We will consider a large family of hypotheses about $f_i(\mathbf{z})$:

$$H_0 : Y_i(\mathbf{z}) = Y_i(\mathbf{z}') \text{ for all } i, \mathbf{z}, \mathbf{z}' \text{ st } f_i(\mathbf{z}), f_i(\mathbf{z}') \in \mathbb{F}_0 \subseteq \mathbb{F}.$$

(Manski, 2009), (Aronow, 2012), (T. and Kao, 2013), (Bowers et al., 2013), (Athey et al., 2019), Basse et al, 2019), (Puelz et al, 2021).

This null tests whether certain kinds of exposures are equivalent in their outcomes;
e.g., $\mathbb{F}_0 = \{$"control-spillovers", "pure-control"$\}$ (coming soon).

If $\mathbb{F}_0 = \mathbb{F}$, then the null can be tested **exactly** through the celebrated Fisherian randomization test (Fisher, 1935) (Lehmann and Romano, 2005).

General Idea: Fisher's Randomization Test

Let's start with the simple problem.

When $\mathbb{F}_0 = \mathbb{F}$, then all exposures give identical outcomes under the null. This is **equivalent to** the global null of no effect:

$$H_0 : Y_i(\mathbf{z}) = Y_i(\mathbf{z}') \text{ for all } \mathbf{z}, \mathbf{z}', i.$$

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

1. Calculate test statistic, $T = t(\mathbf{Z}, \mathbf{Y})$; e.g., ANOVA statistic
2. pval $= \mathsf{E}[t(\mathbf{Z}', \mathbf{Y}) > T], \quad \mathbf{Z}' \sim P$.

General Idea: Fisher's Randomization Test

Let's start with the simple problem.

When $\mathbb{F}_0 = \mathbb{F}$, then all exposures give identical outcomes under the null. This is **equivalent to** the global null of no effect:

$$H_0 : Y_i(\mathbf{z}) = Y_i(\mathbf{z}') \text{ for all } \mathbf{z}, \mathbf{z}', i.$$

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

1. Calculate test statistic, $T = t(\mathbf{Z}, \mathbf{Y})$; e.g., ANOVA statistic
2. pval $= \mathrm{E}[t(\mathbf{Z}', \mathbf{Y}) > T], \quad \mathbf{Z}' \sim P$.

♠ Works because $t(\mathbf{Z}', \mathbf{Y}) \overset{H_0}{=} t(\mathbf{Z}', \mathbf{Y}') \overset{d}{=} T$.

An assessment of FRT

Main advantages:

- The test is exact in finite samples. No asymptotics.
- Not necessary to have correct Y-model specification.
- The test is robust. Same answer under transformations of Y.

Some disadvantages:

- Can only test "strong" hypotheses. (Currently, a lot of research activity in this area).
- Cannot generalize to population.

An assessment of FRT

Main advantages:

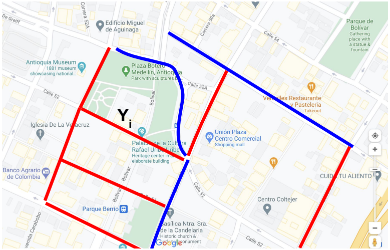
- The test is exact in finite samples. No asymptotics.
- Not necessary to have correct Y-model specification.
- The test is robust. Same answer under transformations of Y.

Some disadvantages:

- Can only test "strong" hypotheses. (Currently, a lot of research activity in this area).
- Cannot generalize to population.

* But can we use it for the Medellin application?

Crime spillovers from nearby **treated streets** on **control streets**?



Recall

$$f_i(\mathbf{z}) = \Big(z_i, \sum_{j:d(i,j)<d_0} z_j\Big).$$

Let $\mathbb{F}_0 = \{$"control-spillovers", "pure-control"$\}$ where

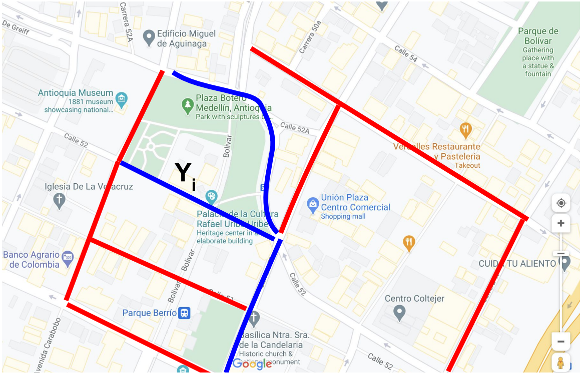
- "control-spillovers" if $f_i(\mathbf{z}) = (0, +)$.
- "pure-control" if $f_i(\mathbf{z}) = (0, 0)$.

Thus, we wish to test that

$$H_0 : Y_i(\text{"control-spill"}) = Y_i(\text{"pure-control"}).$$

FRT problems under interference

Suppose we resample $\mathbf{z}'$ in the FRT as shown below:



The exposure of i is not in $\mathbb{F}_0$. Thus, $Y_i(\mathbf{z}')$ **cannot be imputed** under H_0.

This means that, under interference, we cannot naively apply FRT.

Recent developments

Athey et al (2019), Basse et al (2019) recently proposed a general approach to apply FRTs under interference:

1. Pick a set of *focal units* U according to $P(U \mid \mathbf{Z})$.
2. The test will only depend on data from units in U (we through out all other data).
3. Run FRT by resampling from the correct conditional randomization distribution:

$$P(\mathbf{Z} \mid U) \propto P(U \mid \mathbf{Z}) \cdot P(\mathbf{Z}).$$

Practical issue:

- Make sure that the outcomes of all focal units are *imputable* under H_0 for every $\mathbf{z}$ in the support, $\mathcal{Z}_U = \{\mathbf{z} : P(\mathbf{z} \mid U) > 0\}$.

Conditioning the FRT for spillovers

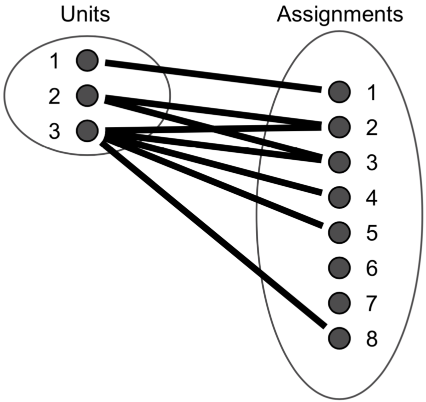
Puelz et al. (2021) developed a general method to construct such valid conditioning for FRTs under spillovers.

Main idea: Connect pair $(i, \mathbf{z})$ iff $f_i(\mathbf{z}) \in \mathbb{F}_0 \Rightarrow$ null exposure graph.

Conditioning the FRT for spillovers

Puelz et al. (2021) developed a general method to construct such valid conditioning for FRTs under spillovers.

Main idea: Connect pair $(i, \mathbf{z})$ iff $f_i(\mathbf{z}) \in \mathbb{F}_0 \Rightarrow$ null exposure graph.



Null exposure graph

The null exposure graph has some very nice properties.

- It encodes the problem structure (only a function of H_0).
- The density of the graph reveals the "support" for testing H_0. (is H_0 easy or hard to test?)
- The bipartite structure describes the power of the test:

$$\text{power} \geq \frac{1}{1 + e^{-\sqrt{n}}} - O(1/\sqrt{m})$$

  ($n=$ #units; $m=$ #assignments in biclique)
- An edge in the graph is equivalent to imputability.

Null exposure graph

The null exposure graph has some very nice properties.

- It encodes the problem structure (only a function of H_0).
- The density of the graph reveals the "support" for testing H_0. (is H_0 easy or hard to test?)
- The bipartite structure describes the power of the test:

$$\text{power} \geq \frac{1}{1 + e^{-\sqrt{n}}} - O(1/\sqrt{m})$$

  ($n=$ #units; $m=$ #assignments in biclique)
- An edge in the graph is equivalent to imputability.

Theorem

Under any conditional randomization test that uses focal units U and assignments $\mathcal{Z}_U$, the potential outcomes are imputable if and only if the sets $(U, \mathcal{Z}_U)$ form a biclique in the null exposure graph.

FRT for spillovers

This leads to the following extension of the classical FRT.

1. Calculate NE graph based on H_0.
2. Calculate a *biclique decomposition* of NE.
3. Resample any $\mathbf{Z}'$ in the biclique proportional to $P(\mathbf{Z}')$.

FRT for spillovers

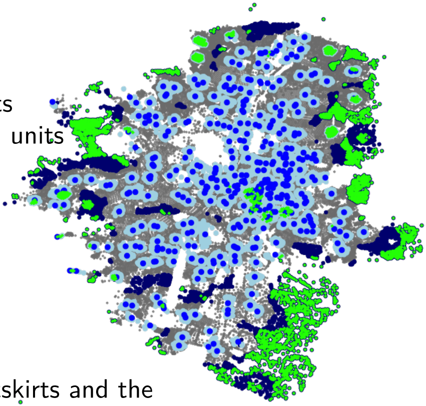
This leads to the following extension of the classical FRT.

1. Calculate NE graph based on H_0.
2. Calculate a *biclique decomposition* of NE.
3. Resample any $\mathbf{Z}'$ in the biclique proportional to $P(\mathbf{Z}')$.

This algorithm automatically generates a conditioning mechanism that fits the particular problem structure.

Medellin application

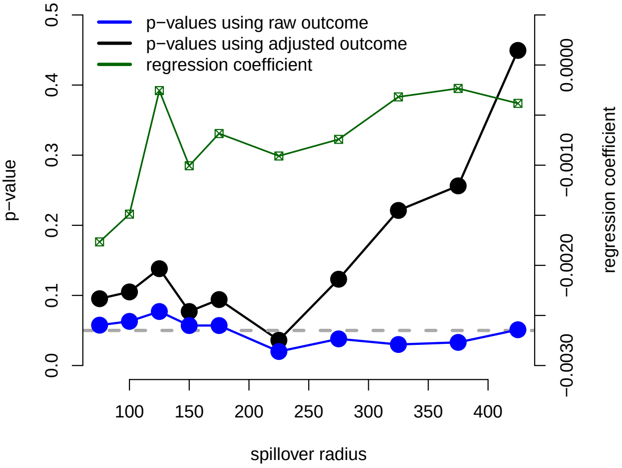
- treated units
- "pure-control" units
- "control-spillovers" units
- focal units



Only units in the outskirts and the city center are pertinent to H_0.

The picture reveals a complex conditioning structure for this particular H_0 that is hard to obtain otherwise.

Results



p−values using raw outcome
p−values using adjusted outcome
regression coefficient
p−value
0.0
0.1
0.2
0.3
0.4
0.5
regression coefficient
−0.0030
−0.0020
−0.0010
0.0000
100
150
200
250
300
350
400
spillover radius

Concluding remarks

Randomization tests can be extended to problems with interference.

These are robust, **finite-sample exact** procedures.

Many open problems remain. — Inference, average spillover effects, etc.

More challenges: Marketplace dynamics, game theory etc.

Thank you!

Basse, Ding, Feller, Toulis “Randomization tests for group formation experiments” , (R&R, 2023)

Puelz, Basse, Feller, Toulis “A graph-theoretic approach to randomization tests of causal effects under interference” , (JRSS-B, 2021)

Basse, Feller, Toulis, “Randomization tests of causal effects under interference” (Biometrika, 2019)