# INTERFERENCE AND COMPLEX EXPERIMENTS

Guido Imbens – Stanford University

Methods Lectures, NBER,

July 25, 2024, 4:35-6:00pm

# OUTLINE: INTERFERENCE AND COMPLEX EXPERIMENTS

1. Introduction: General Problem and Challenges
2. Clustering Designs
3. Exposure Mappings
4. Networks
5. Equilibrium Designs
6. Multiple Randomization Designs
7. Bipartite Experiments
8. Proof-of-Concept Experiments

# INTRODUCTION: GENERAL PROBLEM

- $N$ units, all exposed to binary treatment $W_i \in \{0, 1\}$
- Outcome $Y_i$ observed for unit $i$.
- Potential outcomes $Y_i(\mathbf{W})$, indexed by $N$-vector of treatments $\mathbf{W}$.
- Interest in average effect of all exposed versus none exposed:

$$\tau \equiv \frac{1}{N} \sum_{i=1}^{N} \Big( Y_i(\mathbf{1}) - Y_i(\mathbf{0}) \Big)$$

- (but estimating $\tau$ could be tough)
- Could be interested in other questions
  - Other estimands may be easier to estimate
  - Testing null hypothesis of no effect may be easier.

# 1. INTERFERENCE: GENERAL CHALLENGE

- If all units are affected in unrestricted way by all treatments, little can be learned.
- Need some structure/limits on interference to make progress.
- Structure is context specific.
  - In online settings the platform may determine structure.
  - It may come from networks/choices.
  - Many different structures possible.
- Exciting area, lots of new ideas and approaches, more to come, used in tech companies
- Asymptotics challenging: as sample size increases, what remains the same? Not random sampling from large population.
- No(t yet an) overarching / general framework.

# 1. INTERFERENCE: TODAY

- Some insights
  - Choice of estimand is often important.
- Some cases with clear results
  - e.g., clustering with known clusters
- Some general/useful concepts
  - e.g., exposure mapping
- Some discussion of critical assumptions
  - local interference,
  - Interference through sufficient statistics
  - interference through treatments or outcomes
- some innovative experimental designs
- experiments dont solve everything, but always help.

# 2. CLUSTERING DESIGNS

- Suppose we partition the population of units $[N] = \{1, 2, \ldots, N\}$ into $C$ clusters (eg individuals in states).
- Suppose that $C_i \in \{1, \ldots, C\}$, with

$$Y_i(\mathbf{W}) = Y_i(\mathbf{W}') \qquad \text{if } W_j = W'_j \ \forall j \text{ such that } C_j = C_i$$

  $\Longrightarrow$ outcomes are only affected by treatments within cluster.
- If in addition assignment is the same within clusters,

$$W_i = W_j \qquad \text{if } C_j = C_i$$

- Then analysis is standard with data aggregated to cluster (e.g., Hudgens & Halloran, 2008). well understood

# 2. CLUSTERING DESIGNS: ESTIMANDS

- One comment: choice of estimands matters:
  - Average effect for population

$$\tau^{\text{pop}} = \frac{1}{N}\sum_{i=1}^{N}\Big(Y_i(\mathbf{1}) - Y_i(\mathbf{0})\Big)$$

  - or average effect of cluster average effects:

$$\tau^{\text{cluster}} = \frac{1}{C}\sum_{c=1}^{C}\frac{1}{N_c}\sum_{i:C_i=c}\Big(Y_i(\mathbf{1}) - Y_i(\mathbf{0})\Big)$$

- If cluster sizes vary substantially (e.g., one megacluster) $\tau^{\text{cluster}}$ is easier to estimate than $\tau^{\text{pop}}$, but may be less interesting.

# 2. CLUSTERING DESIGNS

- But: often clustering structure is not known (Viviano et al, 2023)
  - You may have a big network of units (eg Facebook), and you want to partition (cut) the network into smaller ones that will be used as clusters in experiment
  - Assortment of products: Should the cluster be whole milk in different sizes, also fat free milk, or also non-dairy milk, or all dairy products?
  - bias/variance tradeoff
- Choosing optimal clusters is challenging because it depends on largely unknown quantities (magnitude of spillovers).

# 2. CLUSTERING DESIGNS: BIAS-VARIANCE TRADEOFF IN NETWORK SETTING

- Suppose we have an adjacency matrix **A**, and

$$Y_i(\mathbf{W}) = Y_i(\mathbf{W}') \quad \text{if} \quad W_i = W_i', \quad \text{and} \quad W_j = W_j' \quad \forall \quad j : A_{ij} = 1$$

- A particular cluster configuration has a bias from impurity of the treatments within the cluster, and a variance.
  - Increasing cluster size makes bias smaller, but increases variance.
  - Putting bounds on interference effects allows for operationalizing tradeoff and selecting optimal clusters.

# 2. CLUSTERING DESIGNS: BIAS-VARIANCE TRADEOFF IN NETWORK SETTING

- $\mathcal{N}(i)$ is set of neigbhors of unit $i$ (set of $j$ with $A_{ij} = 1$)

$$Y_i(\mathbf{W}) = Y_i(W_i, \mathbf{W}_{\mathcal{N}(i)})$$

- Suppose $|Y_i(w, \mathbf{1}) - Y_i(w, \mathbf{w})| \leq \phi \sum_{j \in \mathcal{N}(i)} |1 - w_j| / |\mathcal{N}(i)|$
- Variance is approximately (with $\psi$ bound on squared outcome)

$$\psi \sum_{c=1}^{C} \frac{N_c^2}{N^2}$$

- We can try to balance these through choice of clusters.

# 3. EXPOSURE MAPPINGS / EFFECTIVE TREATMENTS

- Aronow & Samii (2017), Manski (2013)
  - $N$ units, treatments $W_i \in \mathbb{W}$, characteristics $X_i \in \mathbb{X}$
  - Potential outcomes $Y_i : \mathbb{W}^N \mapsto \mathbb{R}$
- An exposure mapping is a function $\Delta : \mathbb{W}^N \times \mathbb{X} \mapsto \mathbb{D}$, such that if

$$\Delta(W_i, \mathbf{W}_{(i)}, X_i) = \Delta(W_i', \mathbf{W}_{(i)}', X_i)$$

  then potential outcomes given $\mathbf{W}$ and $\mathbf{W}'$ are the same for unit $i$:

$$Y_i(\mathbf{W}) = Y_i(\mathbf{W}')$$

- Exposure mapping captures structure on interference.

# 3. EXAMPLES OF EXPOSURE MAPPINGS / EFFECTIVE TREATMENTS

- Cluster design: $\Delta(W_i, \mathbf{W}_{(i)}, X_i) = (W_i, \sum_{j:C_j=C_i} W_i)$ so that the outcomes depend on the number of treated units in the cluster (or the fraction of treated units in a cluster)
- Network setting: $A_{ij} \in \{0, 1\}$ is the adjacency matrix, and
  - $\Delta(W_i, \mathbf{W}_{(i)}, X_i) = (W_i, \sum_{j:A_{ij}=1} W_i)$ so that the outcome depends on the number (or fraction) of treated friends.
  - Exposure mapping may depend on friends-of-friends.

# 3. LIMITATIONS OF EXPOSURE MAPPINGS

- Exposure mappings do not accomodate cascades where treatment assignment for unit *i* can affect outcomes for all other units.

- Suppose we have a network with adjacency matrix **A**.

  - Model 1: Interference mediated through treatments of friends.

$$Y_i(\mathbf{W}) = \tau W_i + \theta \sum_{j=1}^{N} A_{ij} W_j$$

  - Model 2: Interference mediated through outcomes of friends. Does not fit exposure mapping (other than trivial one).

$$Y_i(\mathbf{W}) = \tau W_i + \theta \sum_{j=1}^{N} A_{ij} Y_j(\mathbf{W})$$

# 3. LIMITATIONS OF EXPOSURE MAPPINGS

- Bond et al (2012) Facebook voting experiment: treated individuals can choose to reveal to their friends whether they voted or not.
- Their treatment status can only effect their friends if they do reveal whether they voted (an outcome).
- It can potentially affect friends-of-friends who cannot see whether the treated individual revealed their voting behavior.
- No exposure mapping with small number of arguments (trivial exposure mapping depends on full vector $\mathbf{W}$).

# 4. NETWORKS EXPERIMENTS: EXACT TESTS

- $N$ individuals, adjacency matrix $\mathbf{A}$ with $A_{ij} \in \{0, 1\}$.
- Suppose the null hypothesis is "no interference"

$$H_0 : \qquad Y_i(\mathbf{W}) = Y_i(\mathbf{W}') \qquad \text{if} \quad W_i = W_i'.$$

- How can we test this hypothesis given random assignment of the treatment?
- Alternative null hypothesis: no second order interference from friends-of-friends:

$$H_0' : \qquad Y_i(\mathbf{W}) = Y_i(\mathbf{W}') \qquad \text{if } W_i = W_i' \quad \text{and } W_j = W_j' \, \forall \, j : A_{ij} = 1.$$

# 4. NETWORKS EXPERIMENTS: EXACT TESTS

- Choose a set of focal units $\mathbb{F} \subset \{1, \ldots, N\}$
- Given actual assignment $\mathbf{W}_0$, define set of assignments $\mathbb{W}(\mathbf{W}_0)$ such that $\mathbf{W} \in \mathbb{W}$ implies

$$Y_i(\mathbf{W}) = Y_i(\mathbf{W}_0) \qquad \forall \quad i \in \mathbb{F}$$

- Then
  - the distribution of $\mathbf{W}$ conditional on $\mathbf{W} \in \mathbb{W}(\mathbf{W}_0)$ is known (because of randomization)
  - and outcomes for all focal units are known for assignment vectors in this set, so we can do Fisher exact finite sample tests.

# 4. NETWORK EXPERIMENTS: CHALLENGES

- Asymptotics: As $N$ gets large
  - Does degree distribution of $\sum_{j=1}^{N} A_{ij}$ converge or diverge?
- Does $Y_i(\mathbf{W})$ depend on values $W_j$ such that the distance $\ell_{\mathbf{A}}(i, j) > K$ for some fixed $K$? (awkward assumption because it rules out cascades)
- Leung (2022): approximate neighborhood interference requires that expected outcomes are close if all treatments within a neighborhood are the same. Formally:

$$\lim_{s \to \infty} \max_{i} \mathbb{E}|Y_i(\mathbf{W}) - Y_i(\mathbf{W}')| \to 0$$

$$\text{for } \mathbf{W}' \text{ s.t. } W_j = W'_j \text{ if } \ell_{\mathbf{A}}(i, j) < s$$

# 5. EQUILIBRIUM DESIGNS

- Munro, Wager & Xu (2021) consider a market setting.
  - Individuals are assigned to binary treatment (eg discounts).
  - The equilibrium price sets total excess demand equal to 0:

$$P \text{ solves } \sum_i Z_i(W_i,\, p) = 0$$

  - The outcome depends on treatment and prices: $Y_i(W_i, P)$
  - Interfence is mediated by prices, so exposure mapping

$$Y_i(\mathbf{W}) = Y_i(W_i, P(\mathbf{W}))$$

# 5. EQUILIBRIUM DESIGNS

- Experimental Design: perturb prices randomly at individual level in addition to assigning treatments randomly, so observed outcome is

$$Y_i(W_i, P + U_i)$$

- $P$ now solves

$$P \text{ solves } \sum_i Z_i(W_i,\ p + U_i) = 0$$

- Now we can learn direct effects of treatment, and indirect effects through prices.

# 6. MULTIPLE RANDOMIZATION DESIGNS

- Bajari et al (2023) focus on a different structure on the interference.
- They consider settings with two interacting populations, say customers and products (at Amazon), or movies and viewers (at Netflix), or properties and renters (at Airbnb/VRBO), or drivers and riders (at Uber/Lyft), or units and time periods
- Key features of set up:
  - Assignment can be assigned at level of pair viewer/movie.
  - Outcome can be measured at level of pair viewer/movie.
- cross-over / switchback designs are special case going back to 1930s (give cows one feed this week, different feed next week)

# 6. MULTIPLE RANDOMIZATION DESIGNS

- Can think of treatment assignment as matrix instead of vector:

$$\mathbf{W} = \begin{pmatrix} \textit{viewers} \rightarrow & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \textit{movies} & & & & & & & & \\ \downarrow & & & & & & & & \\ 1 & \mathrm{T} & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{T} \\ 2 & \mathrm{T} & \mathrm{C} & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{C} \\ 3 & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{C} \\ 4 & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{T} & \mathrm{T} & \mathrm{T} \\ 5 & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{C} & \mathrm{T} & \mathrm{C} & \mathrm{T} & \mathrm{T} \end{pmatrix}$$

# 6. MULTIPLE RANDOMIZATION DESIGNS

- This nests standard movie or viewer experiments:

Movie Exper.
Random. Movies

$$\mathbf{W} = \begin{pmatrix} \downarrow \textit{movies} \quad \textit{viewers} \rightarrow & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & C & C & C & C & C & C & C \\ 2 & C & C & C & C & C & C & C \\ 3 & T & T & T & T & T & T & T \\ 4 & T & T & T & T & T & T & T \end{pmatrix}$$

Viewer Exper.
Random. Viewers

$$\mathbf{W} = \begin{pmatrix} \downarrow \textit{movies} \quad \textit{viewers} \rightarrow & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & C & T & C & C & T & T & T \\ 2 & C & T & C & C & T & T & T \\ 3 & C & T & C & C & T & T & T \\ 4 & C & T & C & C & T & T & T \end{pmatrix}$$

- Concern: viewers may switch from control to treated movies.

# 6. MULTIPLE RANDOMIZATION DESIGNS

- But we can do more interesting things than movie or viewer experiments: Simple Multiple Randomization Design

$$\mathbf{W} = \begin{pmatrix} \textit{viewers} \rightarrow & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \textit{movies} & & & & & & & & \\ \downarrow & & & & & & & & \\ 1 & C & C & C & C & C & C & C & C \\ 2 & C & C & C & C & C & C & C & C \\ 3 & C & C & C & C & T & T & T & T \\ 4 & C & C & C & C & T & T & T & T \\ 5 & C & C & C & C & T & T & T & T \end{pmatrix}$$

- Three control groups that are *ex ante* comparable, but *ex post* have different experiences: C, C, C

# 6. MULTIPLE RANDOMIZATION, KEY ASSUMPTION

- Local Interference assumption

  For two assignment matrices $\mathbf{W}$ and $\mathbf{W}'$, the potential outcomes $Y_{ij}(\mathbf{W})$ and $Y_{ij}(\mathbf{W}')$ are identical if

$$W_{ik} = W'_{ik} \quad \forall k, \qquad \text{and } W_{kj} = W'_{kj} \quad \forall k$$

- Interference only extends within rows and within columns of $\mathbf{W}$.
- Unlikely to hold exactly, but may be reasonable approximation.

# 6. MULTIPLE RANDOMIZATION DESIGNS

- We can compare average outcomes in each of the four groups:

$$\overline{Y}_T,\ \overline{Y}_C,\ \overline{Y}_C,\ \overline{Y}_C$$

  - $\overline{Y}_T - \overline{Y}_C$ tells us about total effect.
  - $\overline{Y}_C - \overline{Y}_C$ tells us about indirect effect of within-viewer/across-movie spillovers.
  - $\overline{Y}_C - \overline{Y}_C$ tells us about indirect effect of within-movie/across-viewer spillovers.
  - $\overline{Y}_T - \overline{Y}_C - \overline{Y}_C + \overline{Y}_C$ tells us about direct effect.

- Variance under randomization can be estimated (little messy)

# 6. MULTIPLE RANDOMIZATION DESIGNS

- More complex Multiple Randomization Design:

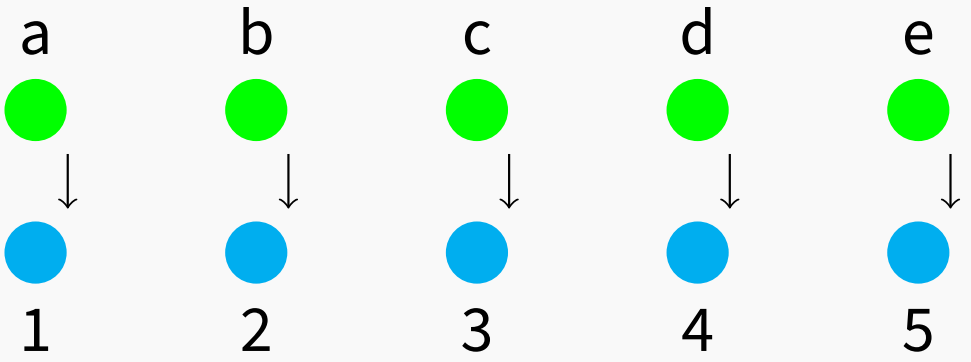
$$\mathbf{W} = \begin{pmatrix}
 & \textit{Movie Ex periment} & & & & & & & & \textit{Viewer Ex periment} & & & & \\
\textit{viewers} \rightarrow & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\
\textit{movies} & A & A & A & A & A & A & A & A & B & B & B & B & B \\
\downarrow & & & & & & & & & & & & & \\
1 & C & C & C & C & C & C & C & C & T & C & T & C & C \\
2 & C & C & C & C & C & C & C & C & T & C & T & C & C \\
3 & T & T & T & T & T & T & T & T & T & C & T & C & C \\
4 & C & C & C & C & C & C & C & C & T & C & T & C & C \\
5 & T & T & T & T & T & T & T & T & T & C & T & C & C \\
6 & T & T & T & T & T & T & T & T & T & C & T & C & C \\
7 & C & C & C & C & C & C & C & C & T & C & T & C & C
\end{pmatrix}$$

# 7. BIPARTITE EXPERIMENTS

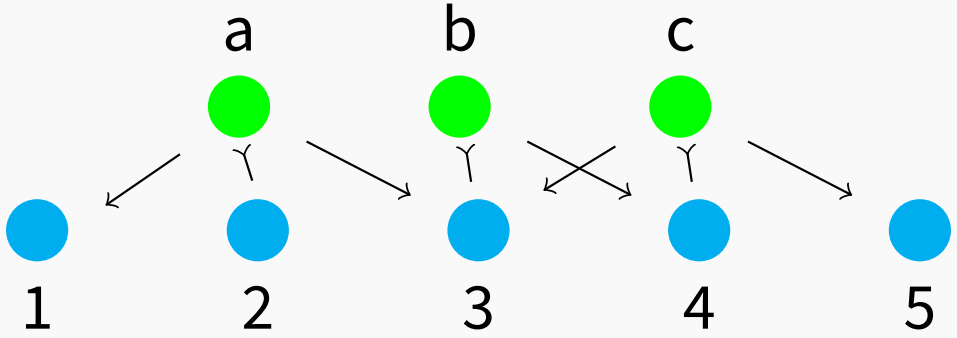
- Sometimes we intervene on units from one population but measure outcomes on units from a different population, without one-to-one correspondence.
  - Set of $J$ intervention units that are exposed to intervention: e.g., power plants can have polution abatement or not.
  - Set of $N$ outcome units for whom we measure outcomes: e.g., individuals have health outcomes affected by abatement in nearby plants.
- Potential outcomes: $Y_i(W_1, \ldots, W_J)$ defined for outcome units, indexed by treatments for intervention units
- Each intervention unit affects some subset of outcome units.

# 7. BIPARTITE EXPERIMENTS

Intervention Units Outcome Units




Conventional Experiment, one-to-one




Bipartite Experiment, one-to-many

# BIPARTITE EXPERIMENTS

- Examples of bipartite settings:
  - Zigler and Papadogeorgou (2021):
    - ⋆ intervention units: power plants,
    - ⋆ outcome units: hospitals.
  - Stock (1989):
    - ⋆ intervention units: pollution sites,
    - ⋆ outcome units: houses.
  - Borusyak & Hull (2023):
    - ⋆ intervention units: railway lines,
    - ⋆ outcome units: regions.

# 8. PROOF OF CONCEPT EXPERIMENTS (NOT QUITE ABOUT INTERFERENCE)

- Suppose we want to establish that a treatment is having some effect, at least for some subpopulation.
- How should we design an experiment that is well-powered for that (rather than for estimating the average effect)?
- In many experiments the power of testing the null of of a zero average effect against the alternative of a non-zero average effect is low.
- How should we change the design to optimize the chances of showing that the treatment has some effect, for some subpopulation?

# 8. PROOF OF CONCEPT EXPERIMENTS

- Suppose a researcher is carrying out a sequential experiment with a binary treatment, and has a discrete pre-treatment variable, say a binary indicator, young/old.
- At each stage the researcher can select a pair of individuals, either a pair of young or a pair of old individuals.
- The researcher than assigns one of the pair to the treatment and one to the control. In the end the resaercher wants to test whether the average effect in the sample is zero or positive.
- The question is how to select the pair at each stage to optimize for power. This is similar to a multi-armed bandit, but with a different objective.

# 8. PROOF OF CONCEPT EXPERIMENTS

- What if we see that so far the treatment effect for young is bigger than that for the old,
  - We should oversample young individuals
  - But not sample proportional to the probability that the effect for young is bigger than that for old (as in multi-armed bandit).
  - Need to take into account outcome variance for young and old.

# 8. PROOF OF CONCEPT EXPERIMENTS

- Suppose a researcher is carrying out an experiment with a treatment taking on $K$ (say $K = 3$) values and a binary outcome.
- As new units come in they can be assigned to any of the three treatments.
- What is the optimal strategy for assigning units given current information in order to maximize the power of testing the null of no effect of the treatment?

# REFERENCES


- Aronow, Peter M., and Cyrus Samii. "Estimating average causal effects under general interference, with application to a social network experiment." *Annals of Applied Statistics* 11, no. 4 (2017): 1912-1947.
- Athey, Susan, Peter J. Bickel, Aiyou Chen, Guido Imbens, and Michael Pollmann. Semiparametric estimation of treatment effects in randomized experiments. No. w29242. National Bureau of Economic Research, 2021.
- Athey, Susan, Raj Chetty, Guido W. Imbens, and Hyunseung Kang. The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely. No. w26463. National Bureau of Economic Research, 2019.

- Bajari, Patrick, Brian Burdick, Guido W. Imbens, Lorenzo Masoero, James McQueen, Thomas S. Richardson, and Ido M. Rosen. "Experimental Design in Marketplaces." *Statistical Science* 1, no. 1 (2023): 1-19.
- Bond, Robert M., Christopher J. Fariss, Jason J. Jones, Adam DI Kramer, Cameron Marlow, Jaime E. Settle, and James H. Fowler. "A 61-million-person experiment in social influence and political mobilization." *Nature* 489, no. 7415 (2012): 295-298.
- Borusyak, Kirill, and Peter Hull. Non-random exposure to exogenous shocks: Theory and applications. No. w27845. National Bureau of Economic Research, 2020.
- Hudgens, Michael G., and M. Elizabeth Halloran. "Toward causal inference with interference." *Journal of the American Statistical*

- Kasy, Maximilian, and Anja Sautmann. "Adaptive treatment assignment in experiments for policy choice." *Econometrica* 89, no. 1 (2021): 113-132.
- Lewis, Randall A., and Justin M. Rao. "The unfavorable economics of measuring the returns to advertising." *The Quarterly Journal of Economics* 130, no. 4 (2015): 1941-1973.
- Manski, Charles F. "Identification of treatment response with social interactions." *The Econometrics Journal* 16, no. 1 (2013): S1-S23.
- Munro, Evan, Stefan Wager, and Kuang Xu. "Treatment effects in market equilibrium." arXiv preprint arXiv:2109.11647 (2021).

- Papadogeorgou, Georgia, Fabrizia Mealli, and Corwin M. Zigler. "Causal inference with interfering units for cluster and population level treatment allocation programs." Biometrics 75, no. 3 (2019): 778-787.
- Stock, James H. "Nonparametric policy analysis." *Journal of the American Statistical Association* 84, no. 406 (1989): 567-575.
- Viviano, Davide, Lihua Lei, Guido Imbens, Brian Karrer, Okke Schrijvers, and Liang Shi. "Causal clustering: design of cluster experiments under network interference." arXiv preprint arXiv:2310.14983 (2023).
- Wager, Stefan, and Susan Athey. "Estimation and inference of heterogeneous treatment effects using random forests." *Journal of the American Statistical Association* 113, no. 523 (2018):

- Wager, Stefan, and Kuang Xu. "Experimenting in equilibrium." *Management Science* 67, no. 11 (2021): 6694-6715.
- Zigler, Corwin M., and Georgia Papadogeorgou. "Bipartite causal inference with interference." Statistical science: a review journal of the Institute of Mathematical Statistics 36, no. 1 (2021): 109.