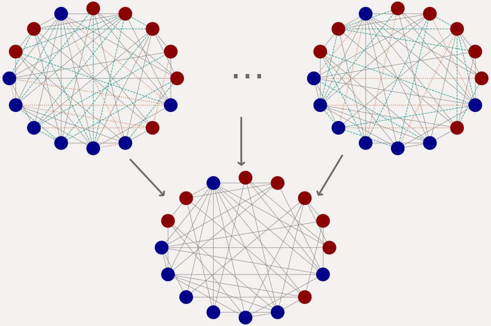
Estimating Causal Effects Using Proxy Interference Networks

Bar Weinstein
Daniel Nevo

Statistics & OR
Tel Aviv University

IDSAI 2025



Field experiement in 56 middle-schools.

Illustrative Example – Paluck et al. (2016)

Field experiement in 56 middle-schools.

1. *Goal:* reduce student conflicts through education.

Illustrative Example – Paluck et al. (2016)

Field experiement in 56 middle-schools.

1. *Goal:* reduce student conflicts through education.
2. *Challenge:* students are influenced by their peers.

Illustrative Example – Paluck et al. (2016)

Field experiement in 56 middle-schools.

1. *Goal:* reduce student conflicts through education.
2. *Challenge:* students are influenced by their peers.
3. *Data:* Multiple friendship networks.
   - Who do you spend time with?
   - Who is your best friend?
   - Measured at pre and post.
   - Self-reported.

Illustrative Example – Paluck et al. (2016)

Field experiement in 56 middle-schools.

1. *Goal:* reduce student conflicts through education.
2. *Challenge:* students are influenced by their peers.
3. *Data:* Multiple friendship networks.
   - Who do you spend time with?
   - Who is your best friend?
   - Measured at pre and post.
   - Self-reported.

- Which of the networks, if any, is the true network? which to choose?

Illustrative Example – Paluck et al. (2016)

Field experiement in 56 middle-schools.

1. *Goal:* reduce student conflicts through education.
2. *Challenge:* students are influenced by their peers.
3. *Data:* Multiple friendship networks.
   - Who do you spend time with?
   - Who is your best friend?
   - Measured at pre and post.
   - Self-reported.

- Which of the networks, if any, is the true network? which to choose?
- **Objective:** Estimate the intervention effects using the proxy networks.

Background

- **Causal Inference.** Effect of treatment on an outcome.

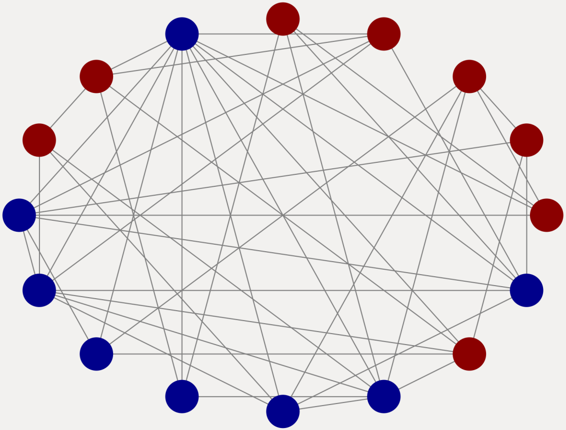
Background

- **Causal Inference.** Effect of treatment on an outcome.
- **Interference.** Treatment of one unit affect the outcomes of others.
  - Spillovers, peer effects, contagion, etc.

Background

- **Causal Inference**. Effect of treatment on an outcome.
- **Interference**. Treatment of one unit affect the outcomes of others.
  - Spillovers, peer effects, contagion, etc.
- Treatments spreads through a **network**.
  - Nodes: units. Edges: pairwise interactions.

- **Causal Inference.** Effect of treatment on an outcome.
- **Interference.** Treatment of one unit affect the outcomes of others.
    - Spillovers, peer effects, contagion, etc.
- Treatments spreads through a **network**.
    - Nodes: units. Edges: pairwise interactions.
- Examples:
    - *Social networks.* Information & behavior spread.
    - *Public health.* transmisstion of infectious diseases or addictive drugs.
    - A/B testing in marketplaces.



The Challenge

- Randomized experiment in networked population.

- Randomized experiment in networked population.
- Accurately measuring social networks is challenging.

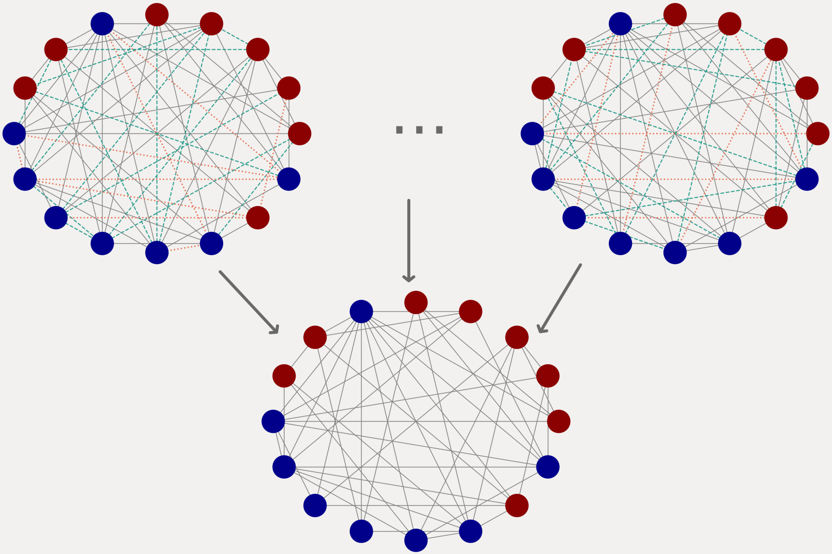
The Challenge

- Randomized experiment in networked population.
- Accurately measuring social networks is challenging.
- We observe only proxy measurements of the true network.
  - Measurements error.
  - Multiple sources of data.
  - Multilayer networks.
- True network remains latent.

The Challenge

- Randomized experiment in networked population.
- Accurately measuring social networks is challenging.
- We observe only proxy measurements of the true network.
  - Measurements error.
  - Multiple sources of data.
  - Multilayer networks.
- True network remains latent.

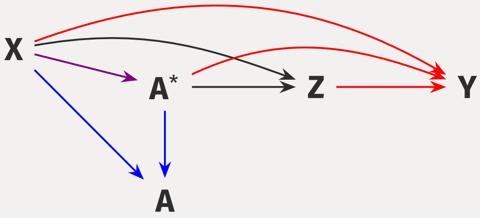
How can we estimate causal effects using proxy networks?



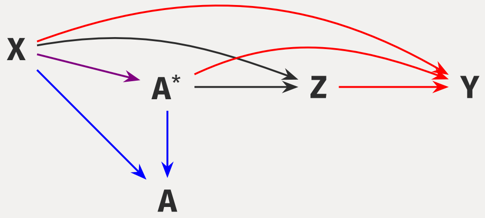
Formal Setup

- Finite population $i \in \{1, \dots, N\}$.
- Treatments: $\mathbf{Z} \in \{0, 1\}^N$.
- Outcomes: $\mathbf{Y} \in \mathbb{R}^N$.
- Covariates/features: $\mathbf{X}$.
- True interference network: $\mathbf{A}^* \in \{0, 1\}^{N \times N}$.
- Proxy networks: $\mathbf{A} = \left(\mathbf{A}^1, \dots, \mathbf{A}^B\right)$.

- Population-level directed acyclic graph (DAG):



- Population-level directed acyclic graph (DAG):



- Requires probabilistic models:
  1. *True network.* $p(\mathbf{A}^*|\mathbf{X}, \theta)$.
  2. *Proxy networks.* $p(\mathbf{A}|\mathbf{A}^*, \mathbf{X}, \gamma)$.
  3. *Outcomes.* $p(\mathbf{Y}|\mathbf{Z}, \mathbf{A}^*, \mathbf{X}, \eta)$.

Causal Estimands

Impact of hypothetical treatments assignment policies.

Impact of hypothetical treatments assignment policies.

1. **Static.** $\mathbb{E}\left[Y_i | do(\mathbf{Z} = \mathbf{z}), \mathbf{X}, \mathbf{A}^*\right]$.
   - Treating all ($\mathbf{z} = \mathbf{1}$) vs none ($\mathbf{z} = \mathbf{0}$).

Impact of hypothetical treatments assignment policies.

1. **Static.** $\mathbb{E}\left[Y_i | do(\mathbf{Z} = \mathbf{z}), \mathbf{X}, \mathbf{A}^*\right]$.
    - Treating all ($\mathbf{z} = \mathbf{1}$) vs none ($\mathbf{z} = \mathbf{0}$).
2. **Dynamic.** $\mathbb{E}\left[Y_i | do(\mathbf{Z} = h(\mathbf{X}, \mathbf{A}^*)), \mathbf{X}, \mathbf{A}^*\right]$.
    - Treating units with specific features, e.g., above certain age.

Impact of hypothetical treatments assignment policies.

1. **Static.** $\mathbb{E}\left[Y_i | do(\mathbf{Z} = \mathbf{z}), \mathbf{X}, \mathbf{A}^*\right]$.
   - Treating all ($\mathbf{z} = \mathbf{1}$) vs none ($\mathbf{z} = \mathbf{0}$).
2. **Dynamic.** $\mathbb{E}\left[Y_i | do(\mathbf{Z} = h(\mathbf{X}, \mathbf{A}^*)), \mathbf{X}, \mathbf{A}^*\right]$.
   - Treating units with specific features, e.g., above certain age.
3. **Stochastic.** $\mathbb{E}_{\pi_\alpha(\mathbf{Z})}\mathbb{E}\left[Y_i | do(\mathbf{Z}), \mathbf{X}, \mathbf{A}^*\right]$.
   - Randomly treating α_1 vs α_0 percent of units.

- *Data:* $\mathbf{O} = (\mathbf{Y}, \mathbf{Z}, \mathbf{X}, \mathbf{A})$. *Latent:* $\mathbf{A}^*$ and models' parameters (η, γ, θ).

- *Data:* $\mathbf{O} = (\mathbf{Y}, \mathbf{Z}, \mathbf{X}, \mathbf{A})$. *Latent:* $\mathbf{A}^*$ and models' parameters (η, γ, θ).
- Posterior distribution:

$$\begin{aligned} p(\mathbf{A}^*, \eta, \gamma, \theta | \mathbf{O}) \propto\ & p(\mathbf{Y}|\mathbf{Z}, \mathbf{A}^*, \mathbf{X}, \eta) p(\eta) \\ & \times p(\mathbf{A}|\mathbf{A}^*, \mathbf{X}, \gamma) p(\gamma) \\ & \times p(\mathbf{A}^*|\mathbf{X}, \theta) p(\theta). \end{aligned}$$

Bayesian Inference

- *Data:* $\mathbf{O} = (\mathbf{Y}, \mathbf{Z}, \mathbf{X}, \mathbf{A})$. *Latent:* $\mathbf{A}^*$ and models' parameters (η, γ, θ).
- Posterior distribution:

$$
\begin{aligned}
p(\mathbf{A}^*, \eta, \gamma, \theta | \mathbf{O}) \propto\ & p(\mathbf{Y}|\mathbf{Z}, \mathbf{A}^*, \mathbf{X}, \eta)p(\eta) \\
& \times p(\mathbf{A}|\mathbf{A}^*, \mathbf{X}, \gamma)p(\gamma) \\
& \times p(\mathbf{A}^*|\mathbf{X}, \theta)p(\theta).
\end{aligned}
$$

- Mixed space. Discrete $(\mathbf{A}^*)$ have $O(N^2)$ terms!

- *Data:* $\mathbf{O} = (\mathbf{Y}, \mathbf{Z}, \mathbf{X}, \mathbf{A})$. *Latent:* $\mathbf{A}^*$ and models' parameters (η, γ, θ).
- Posterior distribution:

$$\begin{aligned} p(\mathbf{A}^*, \eta, \gamma, \theta | \mathbf{O}) \propto\ & p(\mathbf{Y}|\mathbf{Z}, \mathbf{A}^*, \mathbf{X}, \eta) p(\eta) \\ & \times p(\mathbf{A}|\mathbf{A}^*, \mathbf{X}, \gamma) p(\gamma) \\ & \times p(\mathbf{A}^*|\mathbf{X}, \theta) p(\theta). \end{aligned}$$

- Mixed space. Discrete $(\mathbf{A}^*)$ have $O(N^2)$ terms!
- Propose two sampling schemes:
  1. *Modularization.* "break" the posterior into smaller, more manageable parts.
  2. *Gibbs sampling.* Sample discrete with *Local Informed Proposals*.

Simulations

- $N = 500$ units. True network is measured with error. Outcomes from MRF.

Simulations

- N = 500 units. True network is measured with error. Outcomes from MRF.

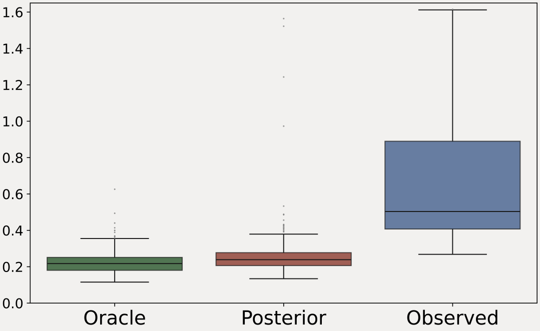


Figure 1: MAPE (↓) of stochastic estimand.

- N = 500 units. True network is measured with error. Outcomes from MRF.

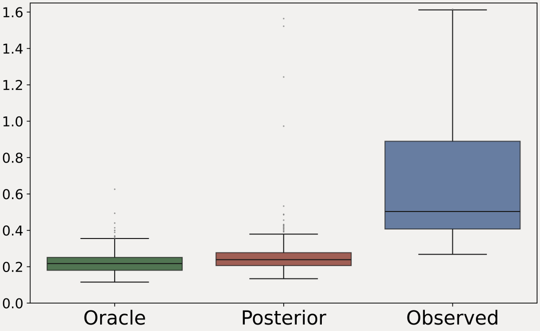


Figure 1: MAPE (↓) of stochastic estimand.

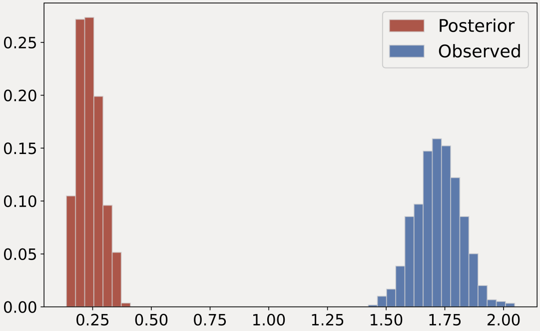
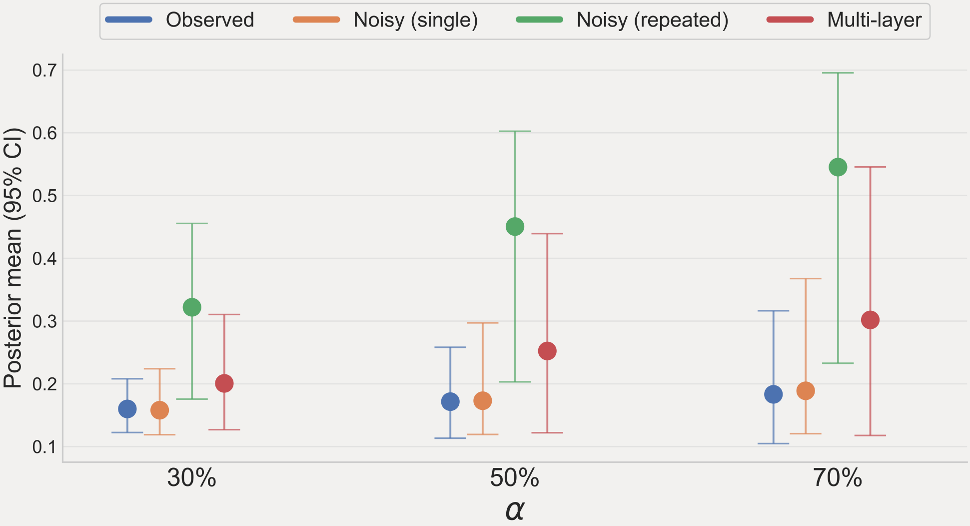


Figure 2: MAE (←) of network statistics.

Data Analysis - Paluck et al. (2016)

- Outcome is indicator of anti-conflict behavior.
- Stochastic estimand.
- Four available networks.
  - Use one as ”Observed” true network.
  - Analysis using different combination of the four proxy networks.

Data Analysis - Paluck et al. (2016)



Key Takeaways

- Network interference is common $\Rightarrow$ implications for many A/B tests.

Key Takeaways

- Network interference is common ⇒ implications for many A/B tests.
- Correctly measuring social relations is often impossible.
- Can estimate causal effects using proxy networks.

Key Takeaways

- Network interference is common $\Rightarrow$ implications for many A/B tests.
- Correctly measuring social relations is often impossible.
- Can estimate causal effects using proxy networks.
- Bayesian framework for inference. Computation is challenging.

Thank You!

Appendix

More math details.