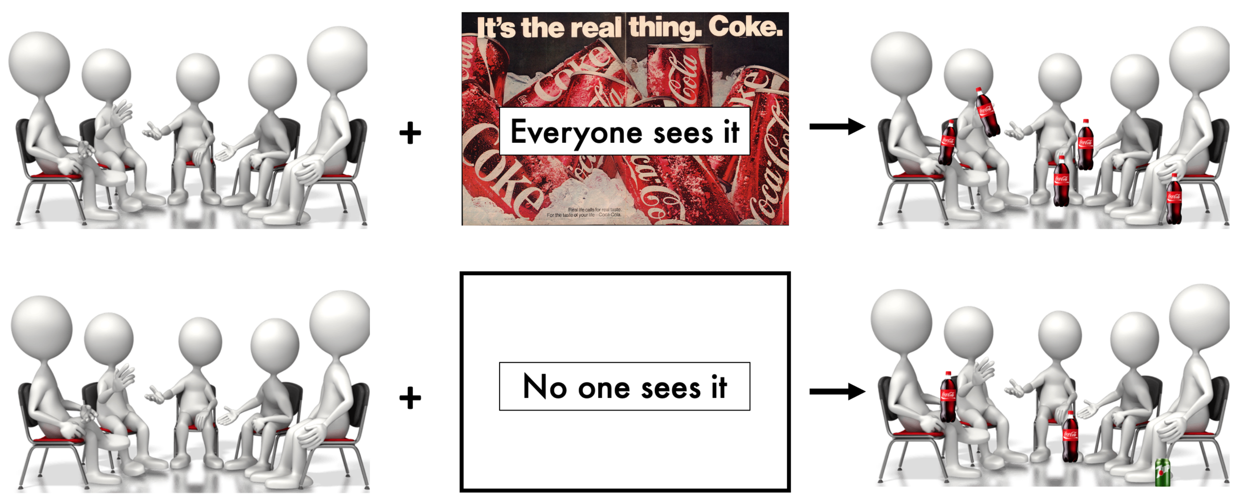
Casual Inference with Neighborhood Interference and Low-Order Interactions

Mayleen Cortez-Rodriguez [1] Matthew Eichhorn [1] Christina Lee Yu [2]

[1]Center for Applied Mathematics [2]Operations Research and Information Engineering

The Problem

- A company runs an experiment to estimate the effectiveness of a national ad campaign
- The **Total Treatment Effect (TTE)** estimand measures the change in the average individual's behavior when everyone sees the ad versus when no one does

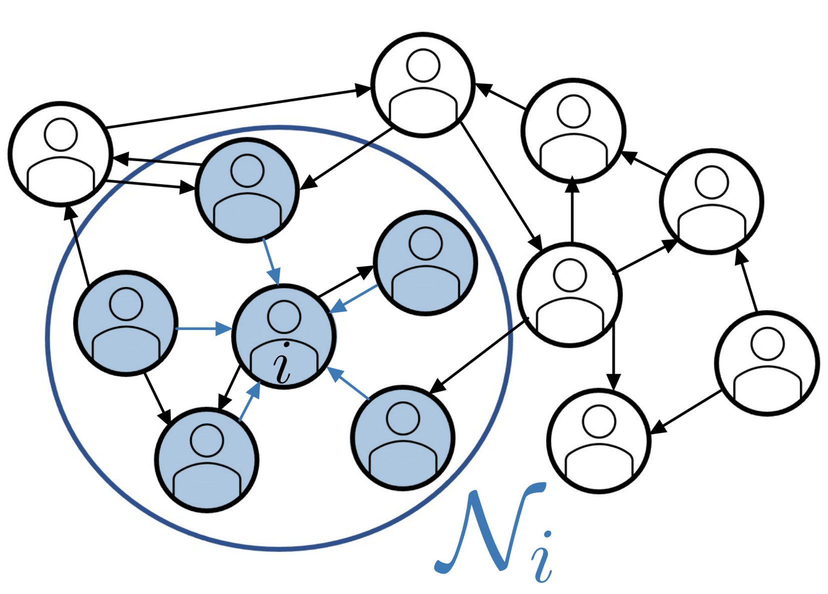


- **Network Interference:** Word-of-mouth spreads advertiser's message beyond direct viewers
- Interference violates the SUTVA assumption and introduces bias to classic estimators

Formalizing the Problem

- **Population:** Directed graph on n nodes, edges encode interference
- **Treatment:** Indicated by $\mathbf{z} \in \{0,1\}^n$
- **Outcomes:** $Y_i(\mathbf{z})$ for each individual i

$$\mathrm{TTE} \triangleq \frac{1}{n}\sum_i \left(Y_i(\mathbf{1}) - Y_i(\mathbf{0})\right)$$



Assumptions

1. **Neighborhood Interference:** Individual i's outcome Y_i is a function of the treatment assignments of in-neighbors $\{z_j\}_{j\in\mathcal{N}_i}$

2. **β-Order Interactions:** Only small subsets of *treated* neighbors affect i's outcome

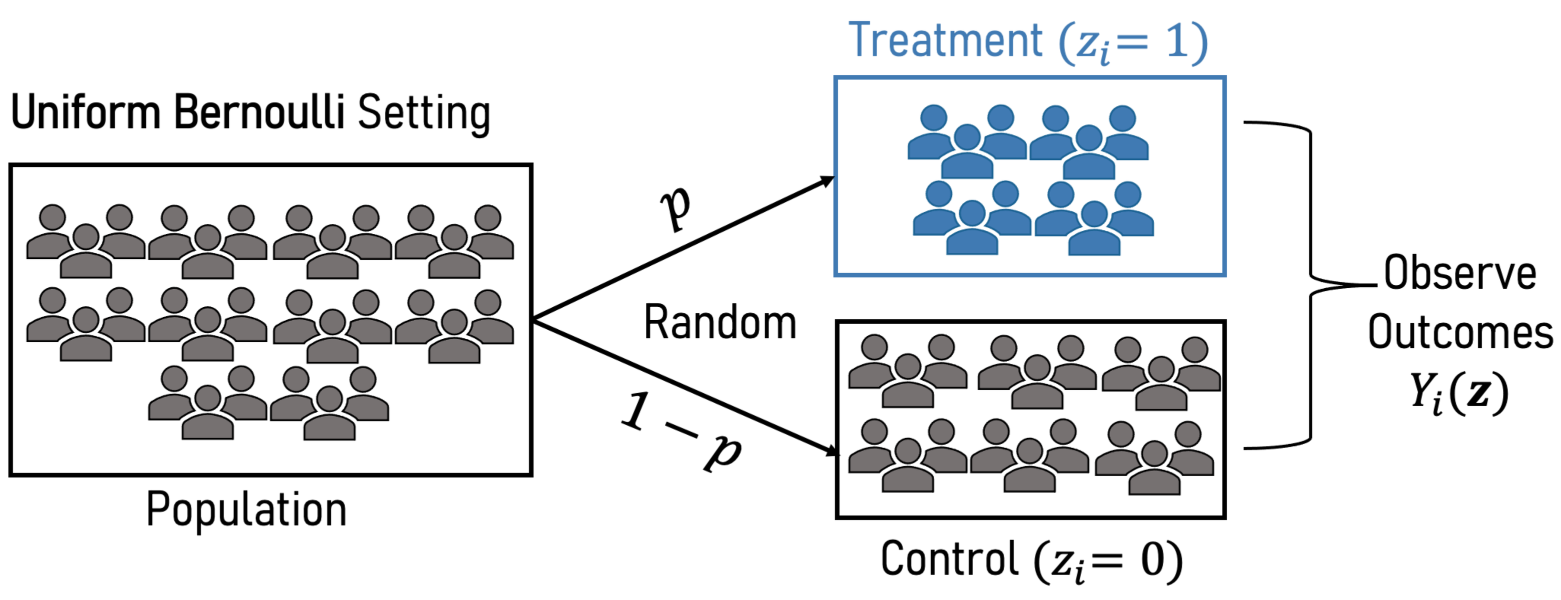
$$Y_i(\mathbf{z}) = \sum_{\substack{\mathcal{S}\subseteq\mathcal{N}_i \\ |\mathcal{S}|\le\beta}} c_{i,\mathcal{S}} \prod_{j\in\mathcal{S}} z_j$$

3. **Bounded Effects:** For each individual i, $\sum_{\substack{\mathcal{S}\subseteq\mathcal{N}_i \\ |\mathcal{S}|\le\beta}} |c_{i,\mathcal{S}}| = \mathcal{O}(1)$

4. **Known Network Structure:** We have knowledge of each $\mathcal{N}_i$

Bernoulli Randomized Design

Treatments sampled independently: $z_i \sim \mathrm{Bernoulli}(p_i)$ with $p_i \in (0,1)$



Research Question

In our potential outcomes framework, can we design an unbiased TTE estimator under Bernoulli randomized design that has a reasonable bound on its variance?

Our Estimator: PI(β)

Unbiased estimator for the TTE with variance $\mathcal{O}\left(\frac{d^2}{n}\cdot\left(\frac{d^2}{p(1-p)}\right)^\beta\right)$ given by

$$\widehat{TTE}_{\mathrm{PI}(\beta)} = \frac{1}{n}\sum_{i=1}^n Y_i(\mathbf{z}) \sum_{\substack{\mathcal{S}\subseteq\mathcal{N}_i \\ |\mathcal{S}|\le\beta}} g(\mathcal{S}) \prod_{j\in\mathcal{S}}\left(\frac{z_j}{p_j} - \frac{1-z_j}{1-p_j}\right)$$

with g, a deterministic, real-valued function chosen for unbiasedness

Special case of the "psuedoinverse" (PI) estimator proposed in [2]

The Horvitz-Thompson Estimator

Contrast with unbiased network Horvitz-Thompson estimator, whose variance scales as $\Theta(1/p^d)$ [3]:

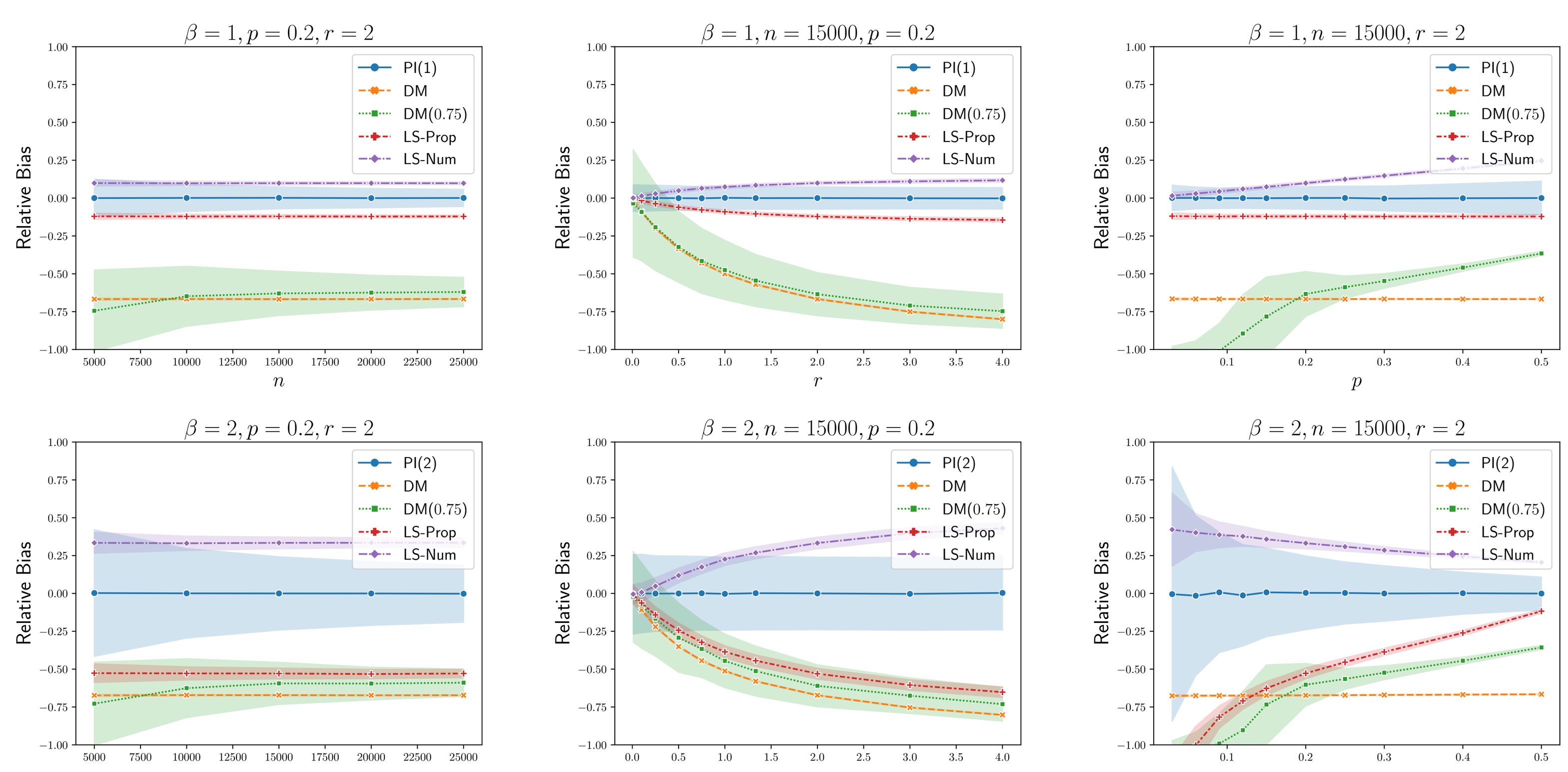
$$\widehat{\mathrm{TTE}}_{\mathrm{HT}} = \frac{1}{n}\sum_{i=1}^n Y_i(\mathbf{z})\left(\frac{\mathbb{I}(\mathbf{z}\text{ treats all of }\mathcal{N}_i)}{\Pr(\mathbf{z}\text{ treats all of }\mathcal{N}_i)} - \frac{\mathbb{I}(\mathbf{z}\text{ doesn't treat all of }\mathcal{N}_i)}{\Pr(\mathbf{z}\text{ doesn't treat all of }\mathcal{N}_i)}\right)$$

$$= \frac{1}{n}\sum_{i=1}^n Y_i(\mathbf{z})\left(\prod_{j\in\mathcal{N}_i}\frac{z_j}{p_j} - \prod_{j\in\mathcal{N}_i}\frac{1-z_j}{1-p_j}\right)$$

Our estimator scales polynomially in d and exponentially in β, a clear improvement when $\beta \ll d$.

Experiments

- Erdős-Rényi network of n nodes with edge probability $p_{\text{edge}} = 10/n$
- Parameter r governs the strength of interference effects
- Parameter p is the treatment budget
- Compare against difference-in-means (DM) and adjusted least-squares (LS) estimators
- **Observation**: Under a β-order outcomes model, our estimator PI(β) generally outperforms other estimators w.r.t. MSE



(a) Varying size of the population (b) Varying interference strength (c) Varying treatment budget

Ongoing Work

- Extend to other randomized designs (e.g. clustering)
- Bias-variance trade off when β is unknown
- Central Limit Theorem result to construct confidence intervals

References

[1] Mayleen Cortez, Matthew Eichhorn, and Christina Lee Yu. Exploiting neighborhood interference with low order interactions under unit randomized design. *arXiv preprint arXiv:2208.05553*, 2022.

[2] Adith Swaminathan, Akshay Krishnamurthy, Alekh Agarwal, Miro Dudik, John Langford, Damien Jose, and Imed Zitouni. Off-policy evaluation for slate recommendation. *Advances in Neural Information Processing Systems*, 30, 2017.

[3] Johan Ugander, Brian Karrer, Lars Backstrom, and Jon Kleinberg. Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 329–337. ACM, 2013.