To Treat or not to Treat, That is the Question

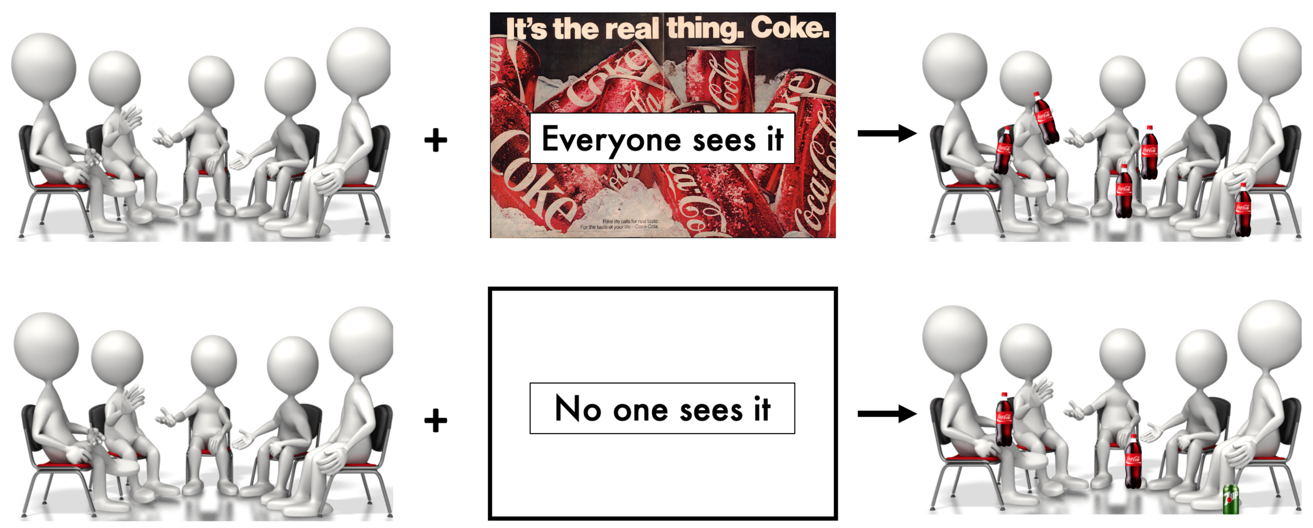
Cornell University

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

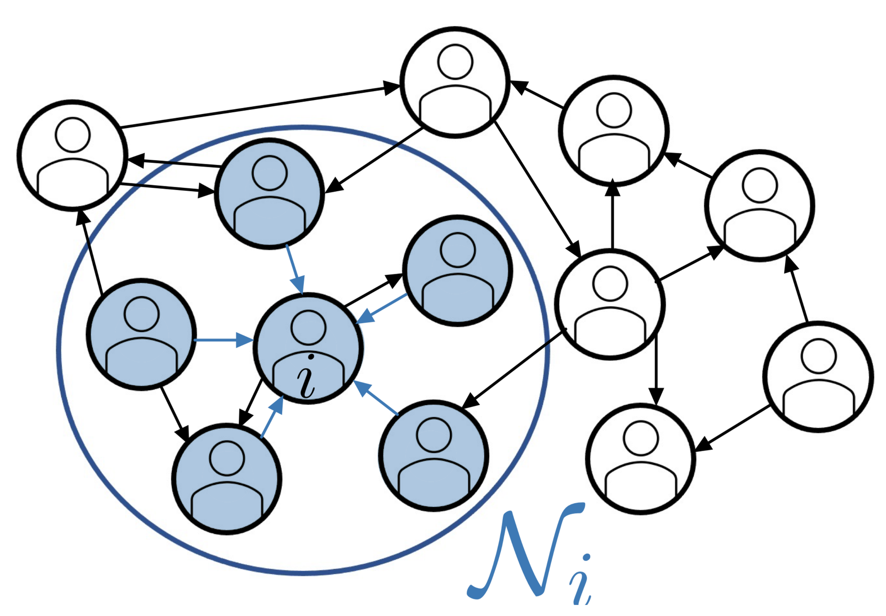


- **Network Interference:** Word-of-mouth spreads advertiser's message beyond direct viewers
- Interference violates the SUTVA assumption and introduces bias to classic estimators

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

$$\text{TTE} \triangleq \frac{1}{n}\sum_i \Big(Y_i(\mathbf{1}) - Y_i(\mathbf{0})\Big)$$



Assumptions

1. **Neighborhood Interference:** Individual i's outcome Y_i is a function only of $\{z_j\}_{j\in\mathcal{N}_i}$
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. **Bernoulli Randomized Design:** $z_i \sim \text{Bernoulli}(p_i)$ independently for $p_i \in [0,1]$

Research Question

Can we design estimators for the total treatment effect under the assumptions listed above that are unbiased and have reasonable bounds on their variance?

Estimating the TTE without Network Knowledge: PI(β)

- Sometimes, we may posit interference but lack access to the causal network
- Social network companies may not reveal network structure to their advertisers

Using a staggered-rollout experimental design, we compensate for the lack of network information by taking multiple outcome measurements.

Consider the function $F(p) = \mathop{\mathbb{E}}_{\mathbf{z}\sim\text{Bern}(p)}\left[\frac{1}{n}\sum_{i=1}^n Y_i(\mathbf{z})\right]$ and note the following:

- $\text{TTE} = F(1) - F(0)$
- F is a polynomial in p with degree $\le \beta$
- Computing the average of $\{Y_i(\mathbf{z})\}_{i=1}^n$ with $\mathbf{z} \sim \text{Bern}(p)$ gives an unbiased estimate of $F(p)$

Recast TTE estimation as polynomial extrapolation after $\beta+1$ rounds of treatment rollout:

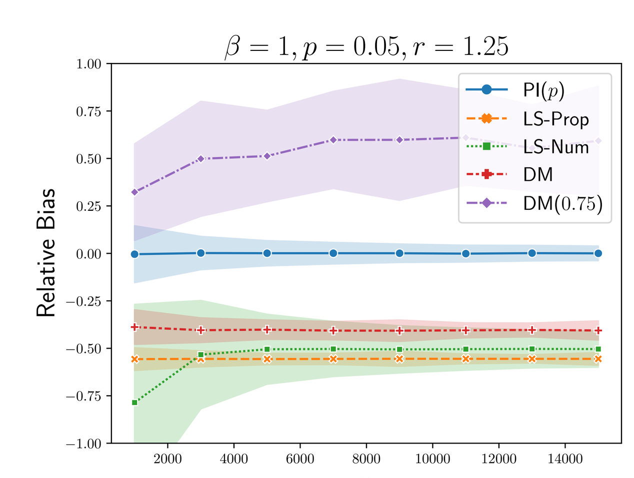
Sample independent $u_i \sim \text{Unif}(0,1)$ for each i and define $\{\mathbf{z}^t\}_{t=0}^{\beta}$ with $z_i^t = \mathbb{I}(u_i \le \frac{tp}{\beta})$. Then,

$$\widehat{\text{TTE}}_{\text{PI}(\beta)} := \frac{1}{n}\sum_{i=1}^{n}\sum_{t=0}^{\beta}\Big(\ell_t(1) - \ell_t(0)\Big)\cdot Y_i(\mathbf{z}^t), \qquad \ell_t(x) = \prod_{\substack{s=0 \\ s\neq t}}^{\beta}\frac{x-x_s}{x_t-x_s}$$

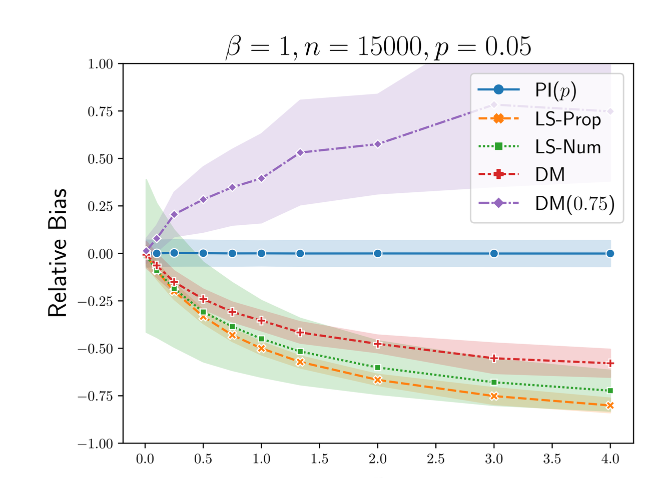
is an unbiased estimator for TTE with variance $\mathcal{O}\Big(\frac{d^2\beta^2}{n}\cdot\big(\frac{\beta}{p}\big)^{2\beta}\Big)$.

Experiments

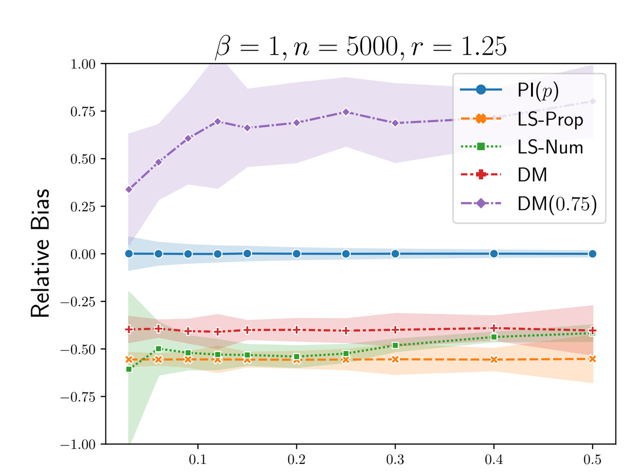
- Configuration model network with in-degrees distributed as a power law with exponent 2.5
- Parameter r governs the strength of interference effects and p is the treatment budget
- Compare against difference-in-means (DM) and least-squares (LS) estimators
- **Observation:** Our estimator PI(p) is unbiased with lower variance than the other estimators



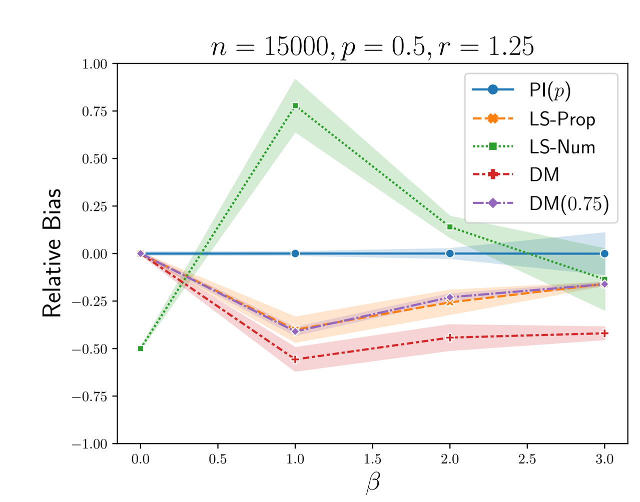
(a) Varying size of the population



(b) Varying interference strength



(c) Varying treatment budget



(d) Varying the model degree

Estimating the TTE with Network Knowledge: SNIPE(β)

- Knowledge of neighborhood sets $\mathcal{N}_i$ with max neighborhood size represented by d
- We have $p_i \in [p, 1-p]$ for some $p \in (0, 0.5]$

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

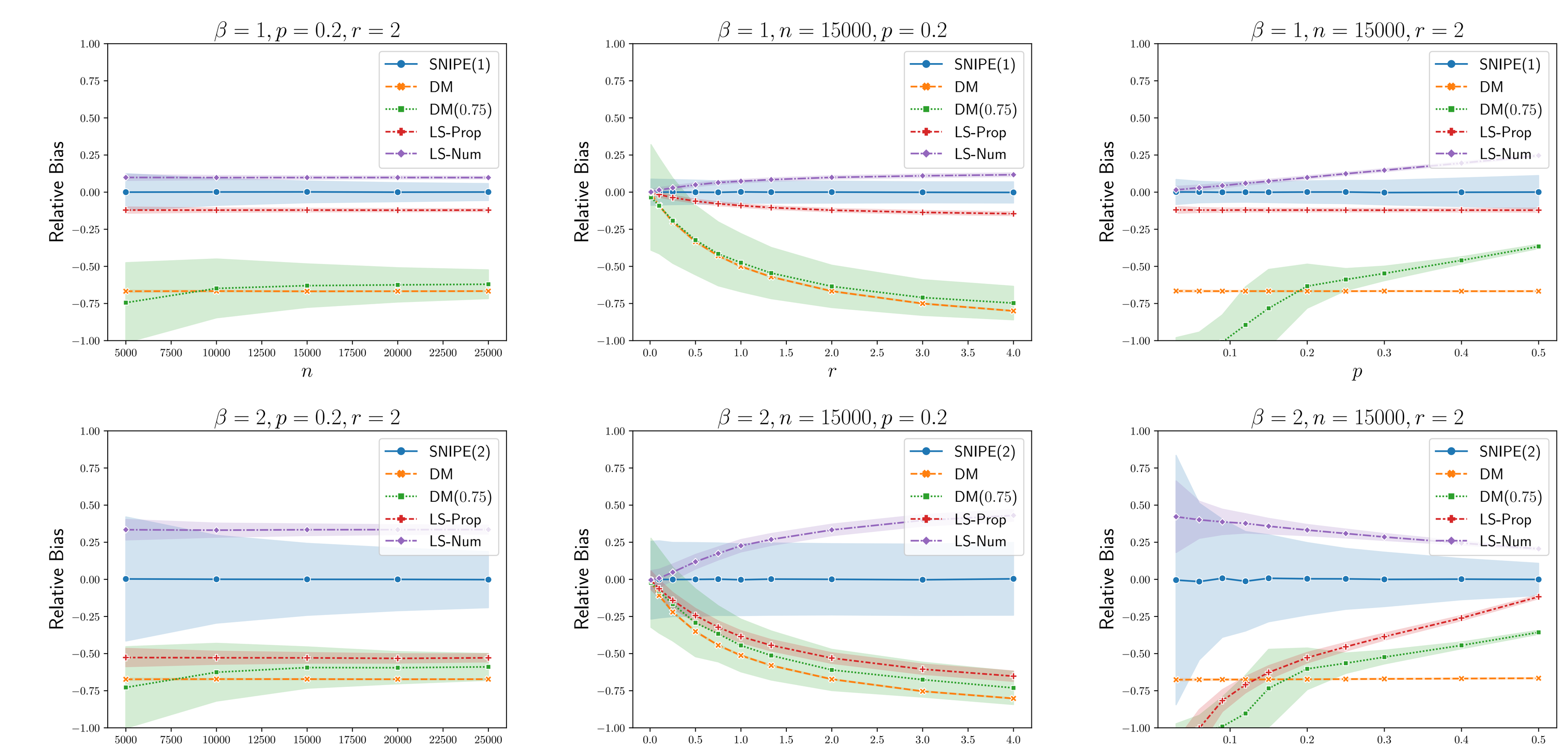
$$\widehat{TTE}_{\text{SNIPE}(\beta)} = \frac{1}{n}\sum_{i=1}^{n} Y_i(\mathbf{z}) \sum_{\substack{\mathcal{S}\subseteq\mathcal{N}_i \\ |\mathcal{S}|\le\beta}} g(\mathcal{S}) \prod_{j\in\mathcal{S}}\left(\frac{z_j}{p_j} - \frac{1-z_j}{1-p_j}\right)$$

where g is a deterministic, real-valued function chosen to ensure unbiasedness

The Horvitz-Thompson estimator is also unbiased, but its variance scales as $\Theta(1/p^d)$. 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 $p_{\text{edge}} = 10/n$
- **Observation**: Our estimator SNIPE(β) generally outperforms other estimators w.r.t. MSE



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

Future Work

- Graph-Agnostic Setting: Generalize to dynamic situations (time-varying effects or networks)
- SNIPE Setting: Extend to other randomized designs (e.g. clustering)
- Both Settings: How to optimally choose β

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] Mayleen Cortez, Matthew Eichhorn, and Christina Lee Yu. Staggered rollout designs enable causal inference under interference without network knowledge. *arXiv preprint arXiv:2205.14552*, 2022.