Design of Randomized Experiments on Networks When Treatment Propagates: An Exploration

Bruce Desmarais
University of Massachusetts Amherst

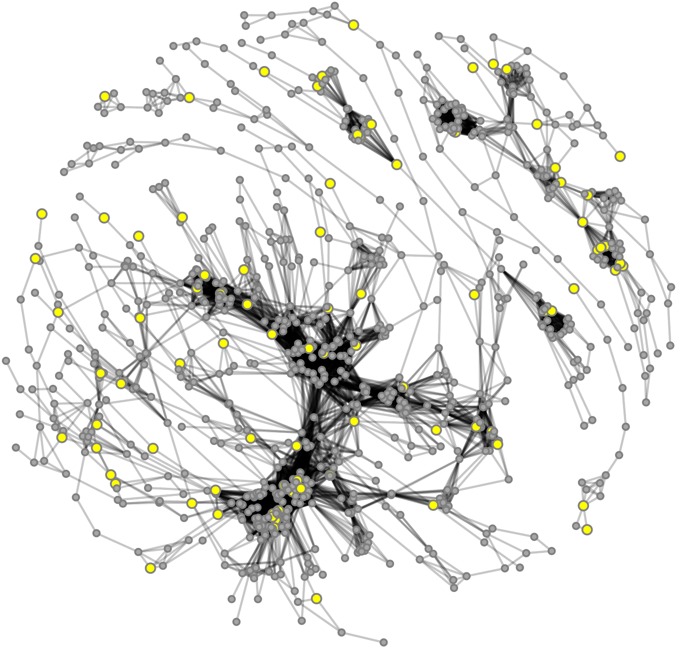
Collaborators:
Jake Bowers, Wayne Lee, Simi Wang, Mark Frederickson, Nahomi Ichino

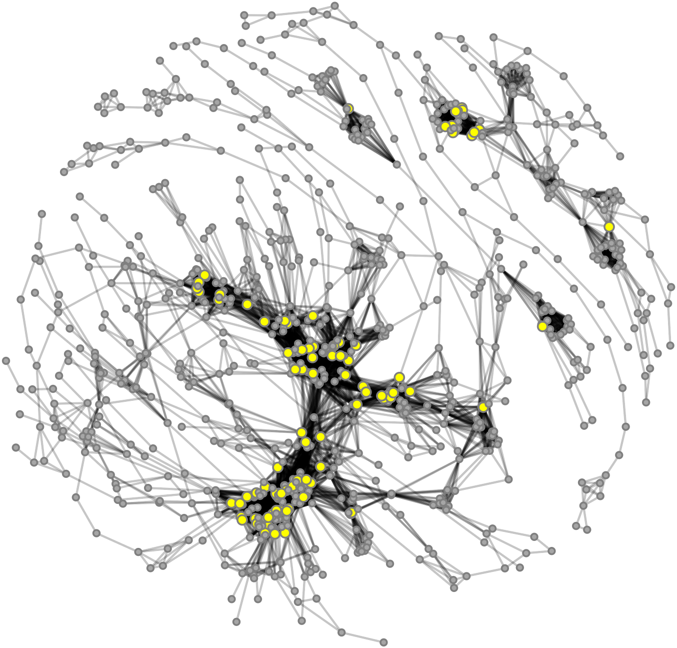
Prepared for the 2014 Political Networks Conference, McGill University

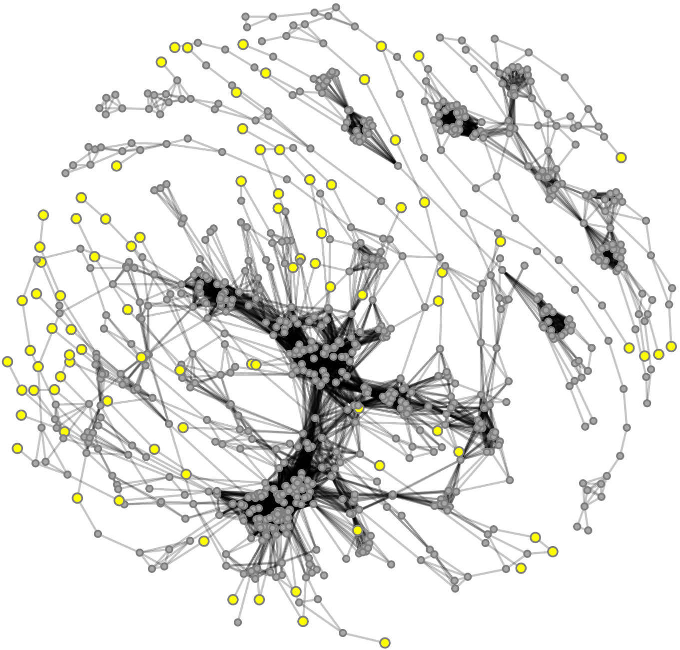
How should experiments be designed to learn about interference in networks?

Example: Ghana 2008 Voter Registration Fraud Experiment (Ichino)

- 868 Electoral Registration Stations (vertices) connected by roads (edges)
- Density of 2.2%
- Graph Transitivity of 70.3%







Notation for Causal Questions

- **Treatment assigned to** i: $Z_i \in \{0, 1\}$
- **Vector of all treatment assignments**: $\mathbf{Z} = \{Z_1, \dots, Z_n\}$
- **Sharp Null of No Treatment Effect**: all potential outcomes are equal

$$Y_i(\mathbf{Z}) = Y_i(\mathbf{Z}') = Y_i(Z_i = 1, \mathbf{Z}_{-i}) = Y_i(Z_i = 0, \mathbf{Z}'_{-i}) \;\; \forall \mathbf{Z} \neq \mathbf{Z}'$$

- **Graph-conditioned exposure**: $Y_{it}(D_{it})$, D_{it} records exposure condition

Example of treatment propagation: The Ising Model

Initial treatment status drawn from $Z_i \sim \text{Bernoulli}(\alpha)$. "Infection" / "Exposure" probability at each iteration is

$$pr(E_i = 1) = \frac{1}{1 + \exp(\frac{2}{F}(k_i - 2m_i))}$$

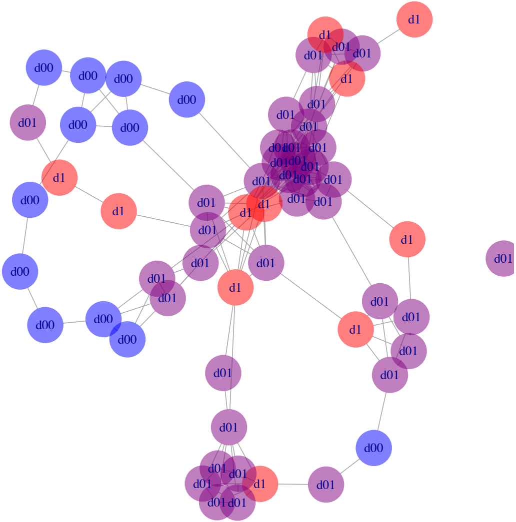
- k is number of (directly adjacent) neighbors ($0 \leq k_i \geq K$)
- m is number of already infected neighbors ($0 \leq m_i \geq M$)
- F is "temperature" or "propensity to be infected"

This example: Only two time periods ($t \in \{0, 1\}$). The Ising model controls actual infection after an experimenter assigns Z_i at $t = 0$. Record infection after the first period.

Example of exposure types on the Ghana nodes

One draw from the Ising propagations with $\text{pr}(Z_i = 1) \sim \text{Bernoulli}(.15)$ for $t \in \{0, 1\}$ and Temperature$= 10$.

$$d_1 \equiv D_i(Z_i = 1, 0 \leq m_{it} \geq M)$$
$$d_{(0,0)} \equiv D_i(Z_i = 0, m_{it} = 0)$$
$$d_{(0,1)} \equiv D_i(Z_i = 0, m_{it} \geq 1)$$



ATE: Aronow and Samii Estimator

Definitions:

- D_i indexes the exposure condition of observation i
- $\pi_i(d_k)$ is the probability i ends up in condition d_k
- We consider
  - $d_{(0,1)}$, untreated with at least one treated neighbor
  - $d_{(0,0)}$, untreated with no treated neighbor

Estimand:

$$\hat{\mu}(d_k) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{I}(D_i = d_k) \frac{Y_i(d_k)}{\pi_i(d_k)}$$

$$\hat{\tau}(d_{(0,1)}, d_{(0,0)}) = \hat{\mu}(d_{(0,1)}) - \hat{\mu}(d_{(0,0)})$$

Testing:

Following Horvitz and Thompson (1952) and CLT:
$\text{Var}(\hat{\tau}(d_{(0,1)}, d_{(0,0)})) \leq 1/N^2 \left(\text{Var}(\hat{\mu}(d_{(0,1)})) + \text{Var}(\hat{\mu}(d_{(0,0)}))\right)$

Simulations to assess power and guide design

- Baseline (pre-treatment) response

$$Y(0,0) \sim U(0,1)$$

- Treatment allowed to propagate for one period and true (constant, multiplicative) model of causal effects computed:

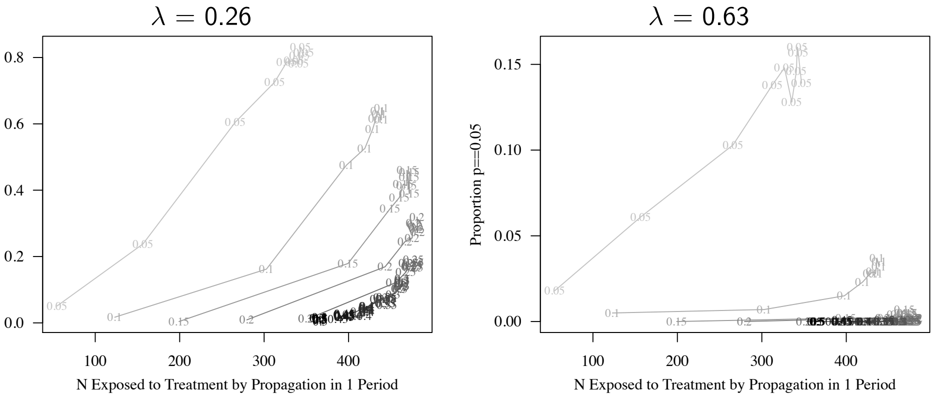
$$Y(1,0) = Y(0,1) = \lambda Y(0,0)$$

- Consider parameters
  - Treatment assignment probability $\alpha \in \{0.05, 0.10, \ldots, 0.50\}$
  - 'Temperature' $T \in \{0, 10, \ldots, 100\}$
  - $\lambda \in \{0.26, 0.63\}$

- Simulate 1,000 propagations at each combination of T and α.

- For each realized propagation, assess power to reject $H_0 : \tau = 0$ at significance level .05.

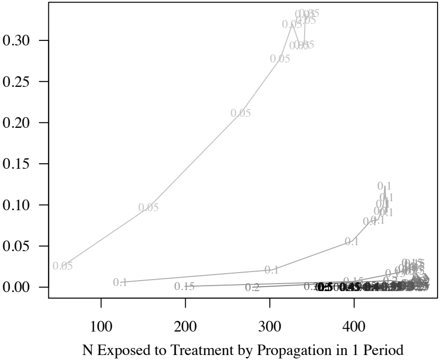
Power as a Function of N Indirectly Treated



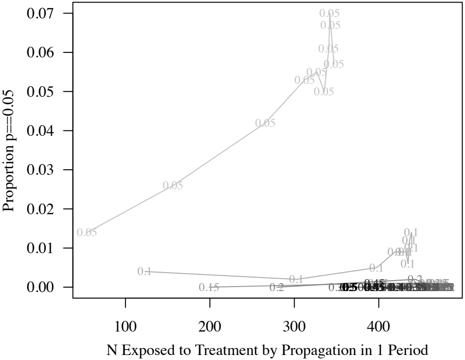
Power With Additive Effects Model

$$Y(0,0) = 1 + U(0,1)$$
$$Y(0,1) = \lambda + U(0,1)$$

$\lambda = 0.26$



$\lambda = 0.63$

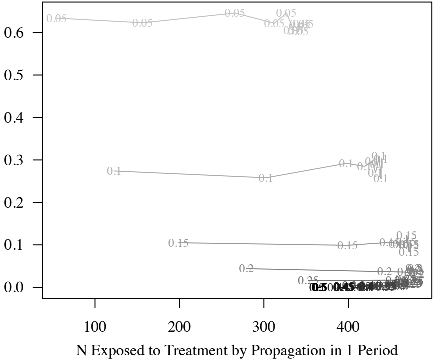


Power With Certain Propagation

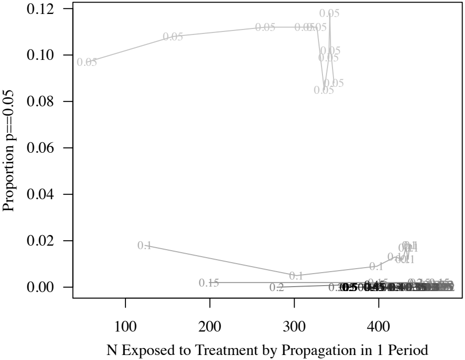
$$Y(d_{00}) = 1 + U(0,1)$$

$$Y(d_{01}) = \lambda + U(0,1)$$

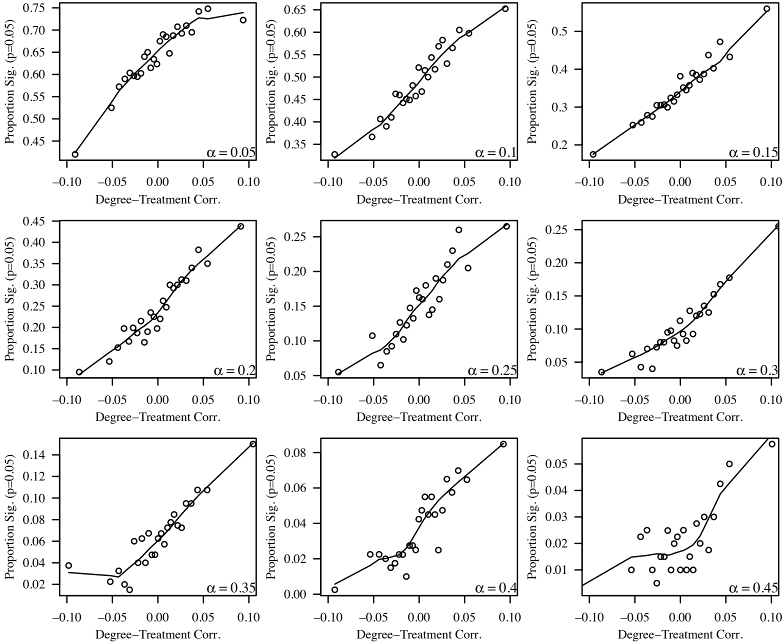
$\lambda = 0.26$



$\lambda = 0.63$



Correlating Treatment with Degree

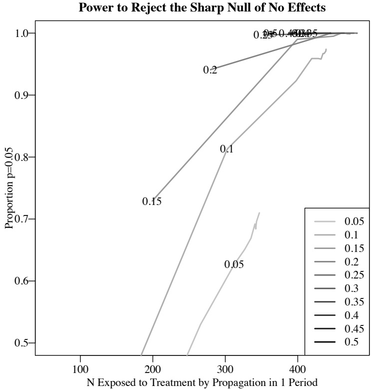


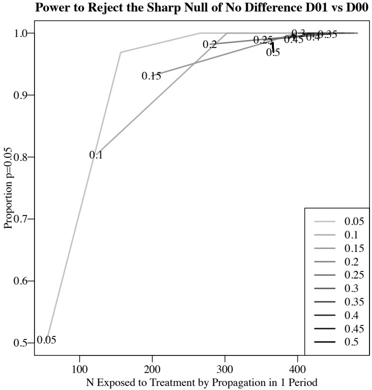
Tests of a Sharp Null: Distributional Differences

Bowers, Jake, Mark M. Fredrickson and Costas Panagopoulos. 2013. "Reasoning about Interference Between Units: A General Framework." *Political Analysis* 21(1):97–124.

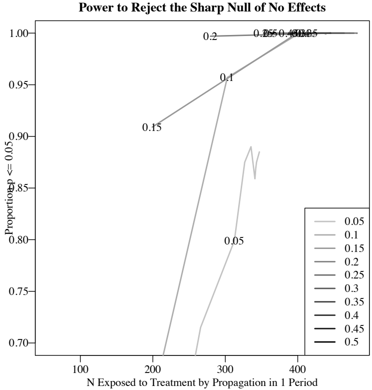
Details: Anderson-Darling *k*-sample test statistic (Scholz and Stephens, 1987); randomization distributions via RItools (Fredrickson, Bowers, Hansen 2014).

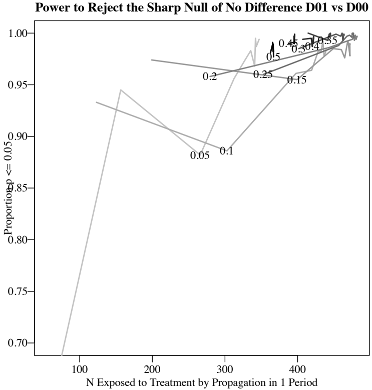
Power as a Function of N Indirectly Treated



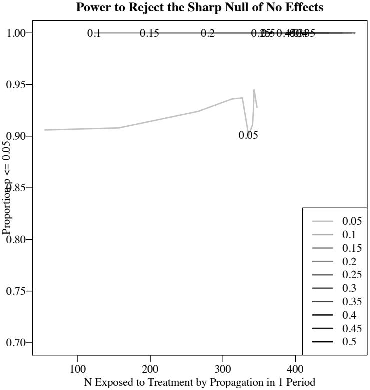


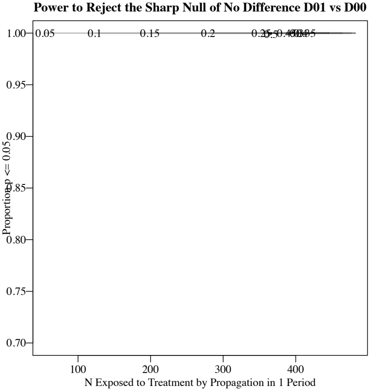
Power With Additive Effects Model





Power With Certain Propagation





Summary

- *Under propagation, power optimizing proportion assigned to treatment may be much less than 0.5.*
- *Higher order parameters of treatment assignment distribution can dramatically affect statistical power*
- Trade-off between power to detect network-moderated and non-network-moderated effects (see also Bowers, Fredrickson, Panagopoulos 2013).
- Test statistics matter: even if interest focuses on ATE, the ATE may provide little power when the true model is not a simple distributional shift.
- Limitation on power driven by the need for strong (i.e., isolated) controls.

Next Steps

- More measures of design adequacy: RMSE, Type I error, Power.
- Make $pr(Z_i = 1)$ depend on topology if not also covariates.
  - degree-correlated assignment
  - community-wise assignment schemes
- Consider higher-order spreading
- Re-parameterize Ising for substantive interpretation.
- Provide more general simulation system that accommodates different:
  - models of propagation
  - models of effects
  - statistical/causal inferential focus