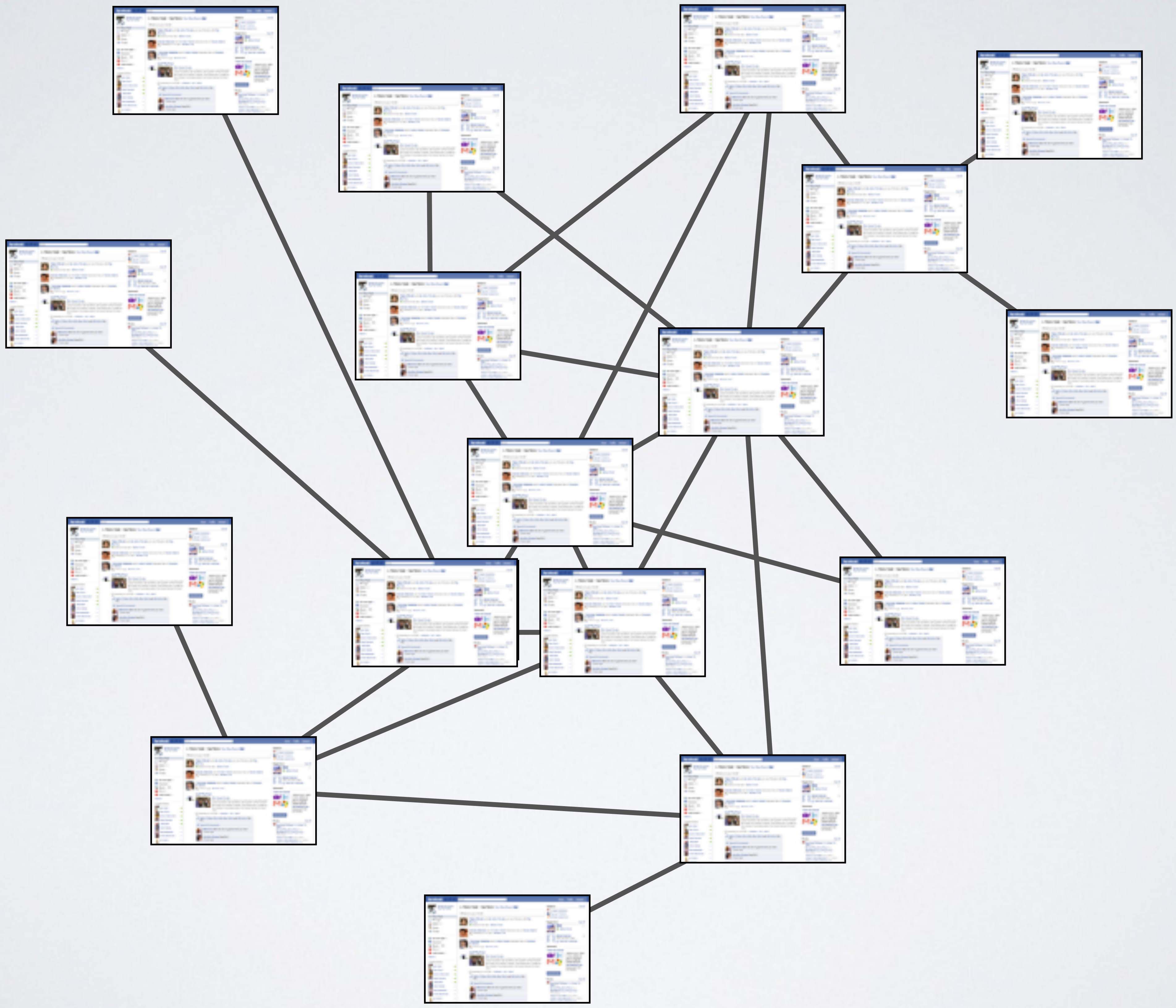
Design and Analysis for Experiments in Networks

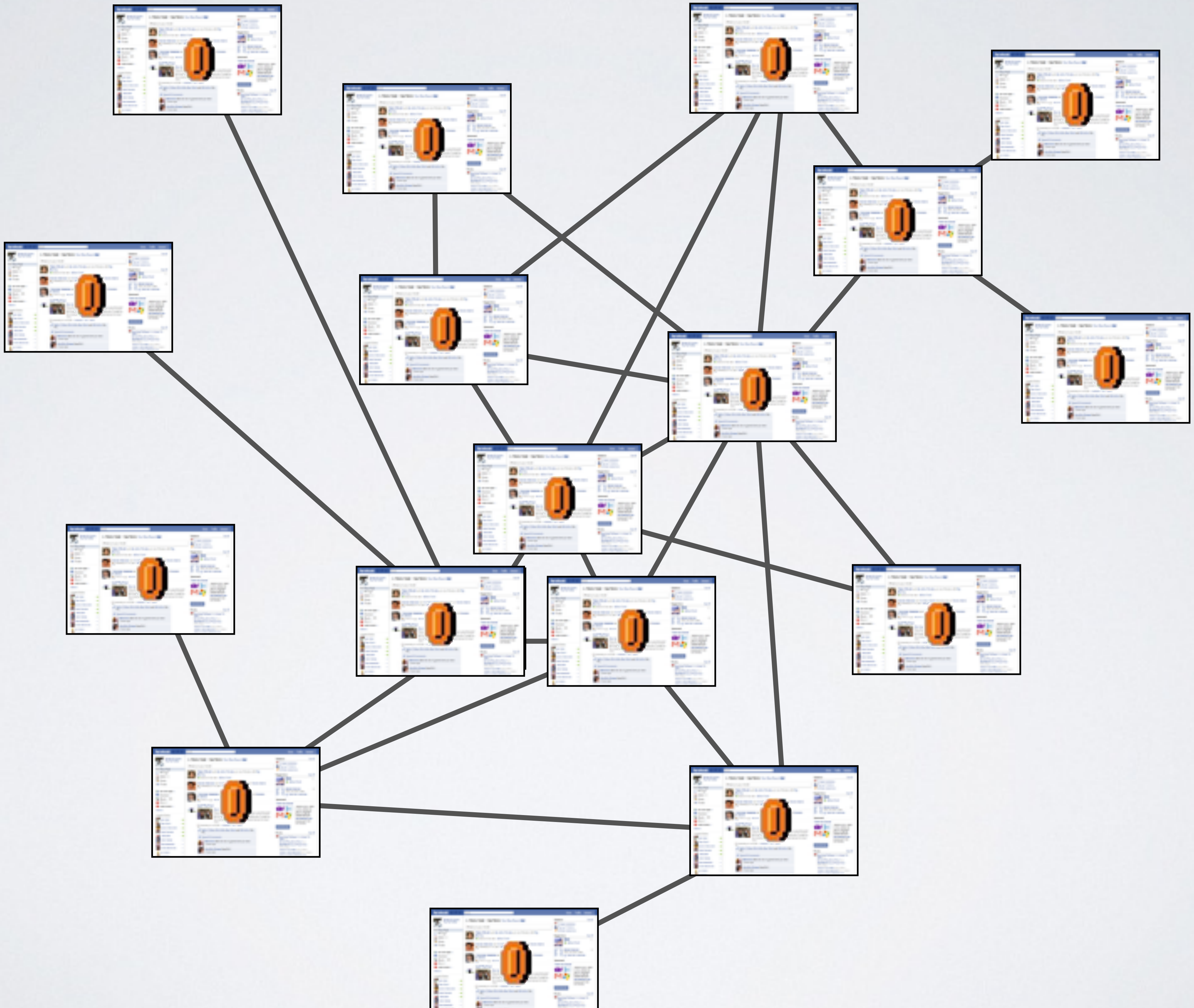
Sufficient Conditions for Reducing Bias from Interference

Johan Ugander (Cornell, Microsoft Research)
joint work with Dean Eckles and Brian Karrer (Facebook)
CODE@MIT
October 10, 2014

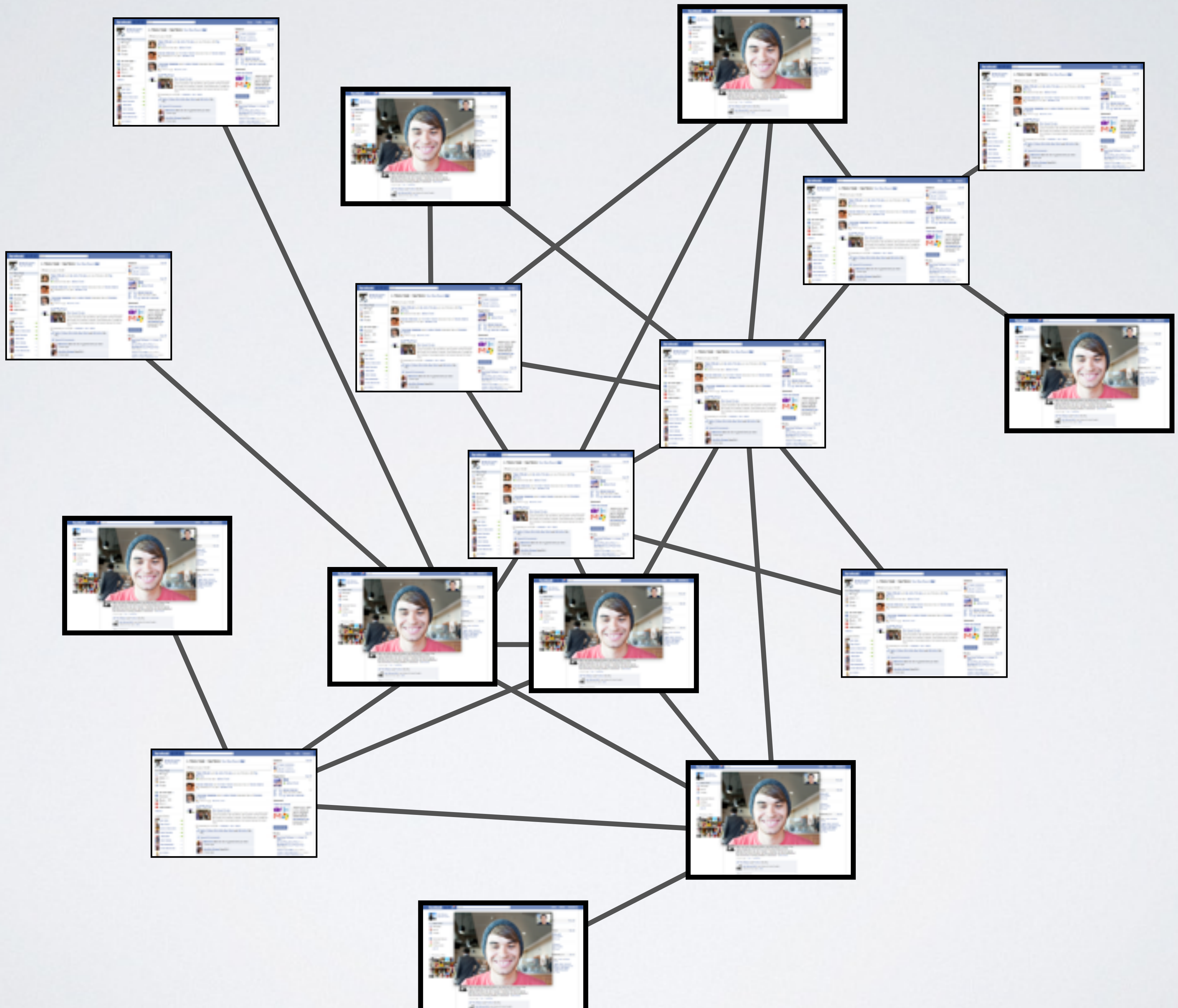
A/B testing on a social network



A/B testing on a social network

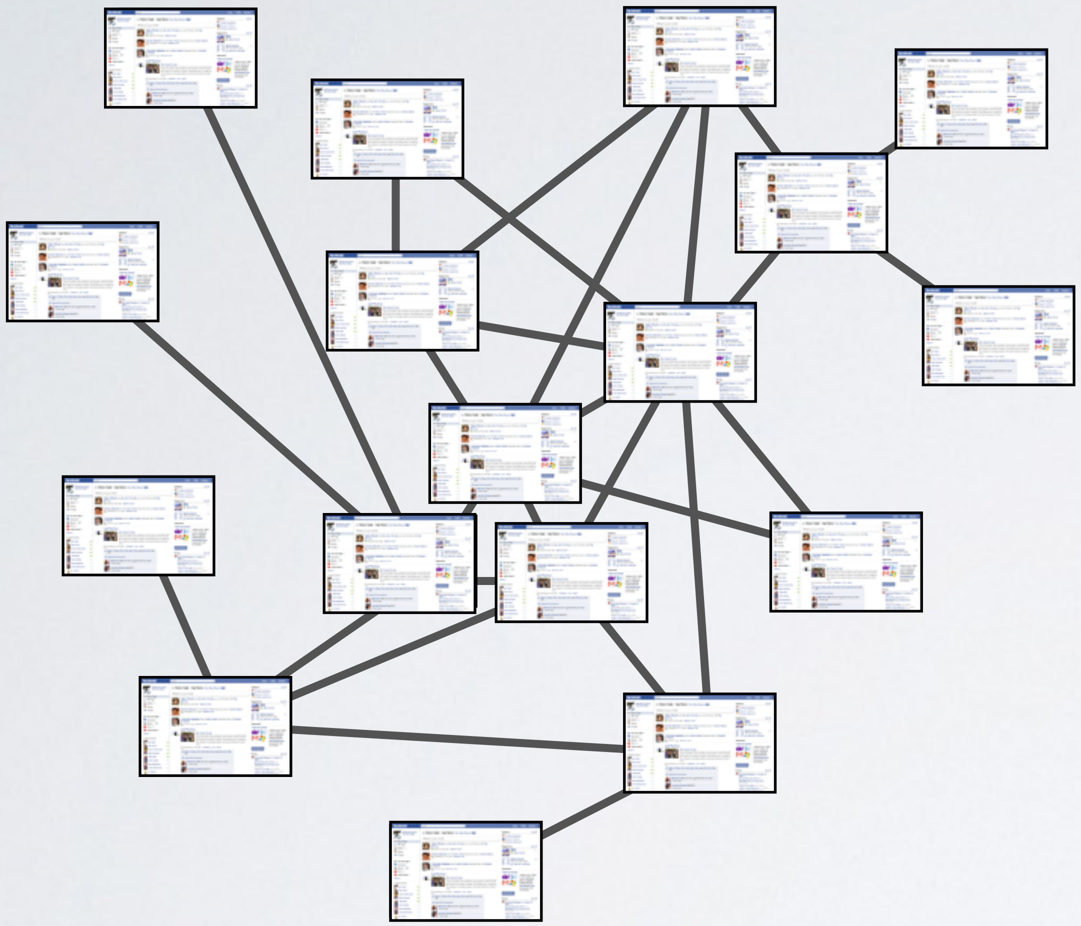


A/B testing on a social network

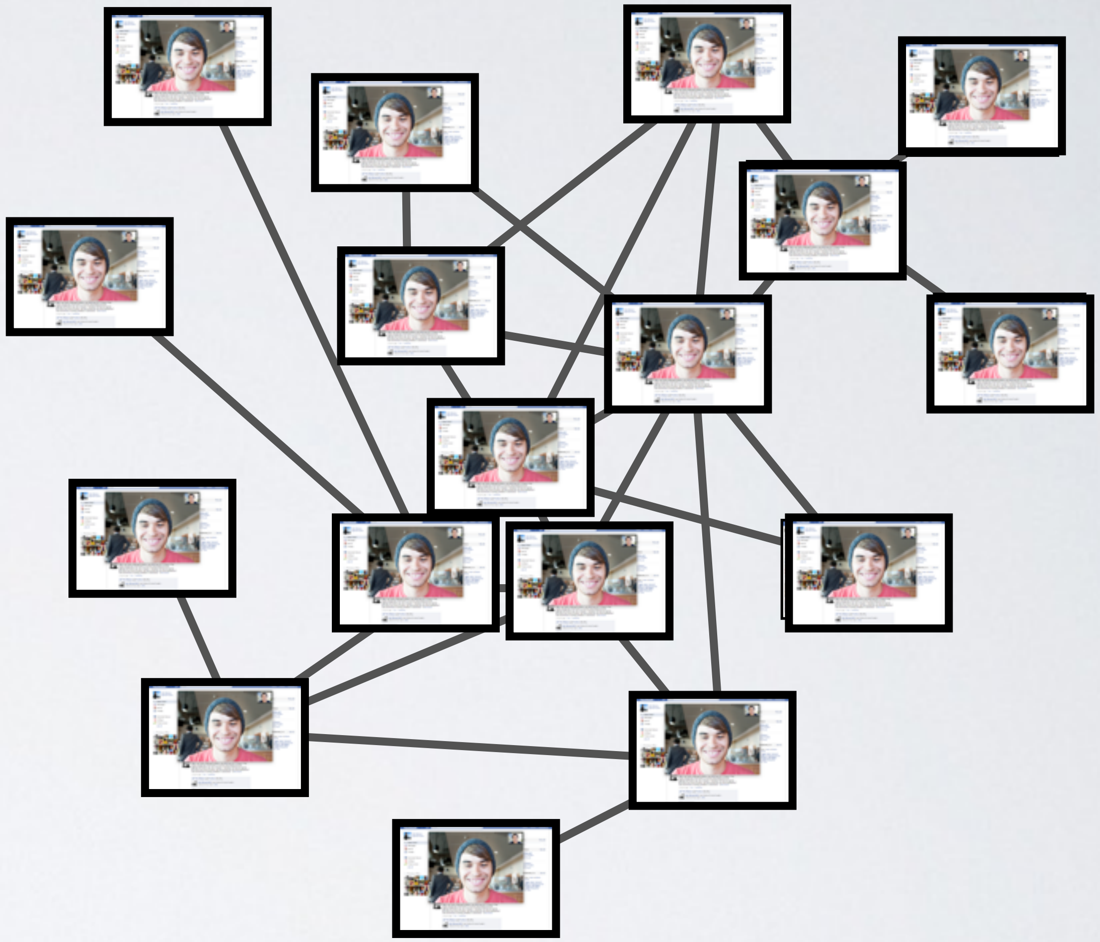


Causal inference on networks

Universe A



Universe B



Fundamental problem: want to compare (average treatment effect, ATE), but can't observe network in both states at once.

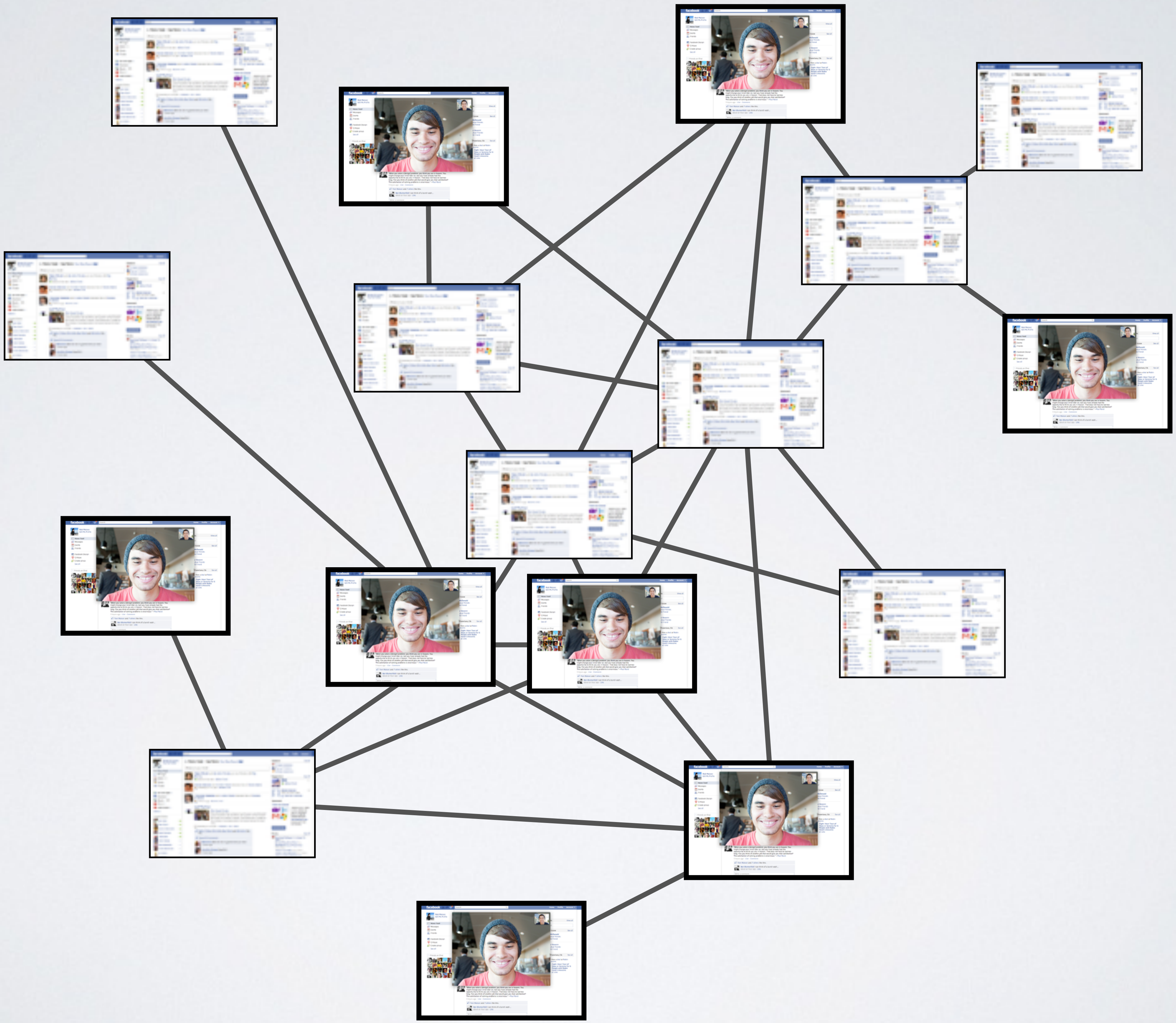
Experiments with interference

Chat/communication features

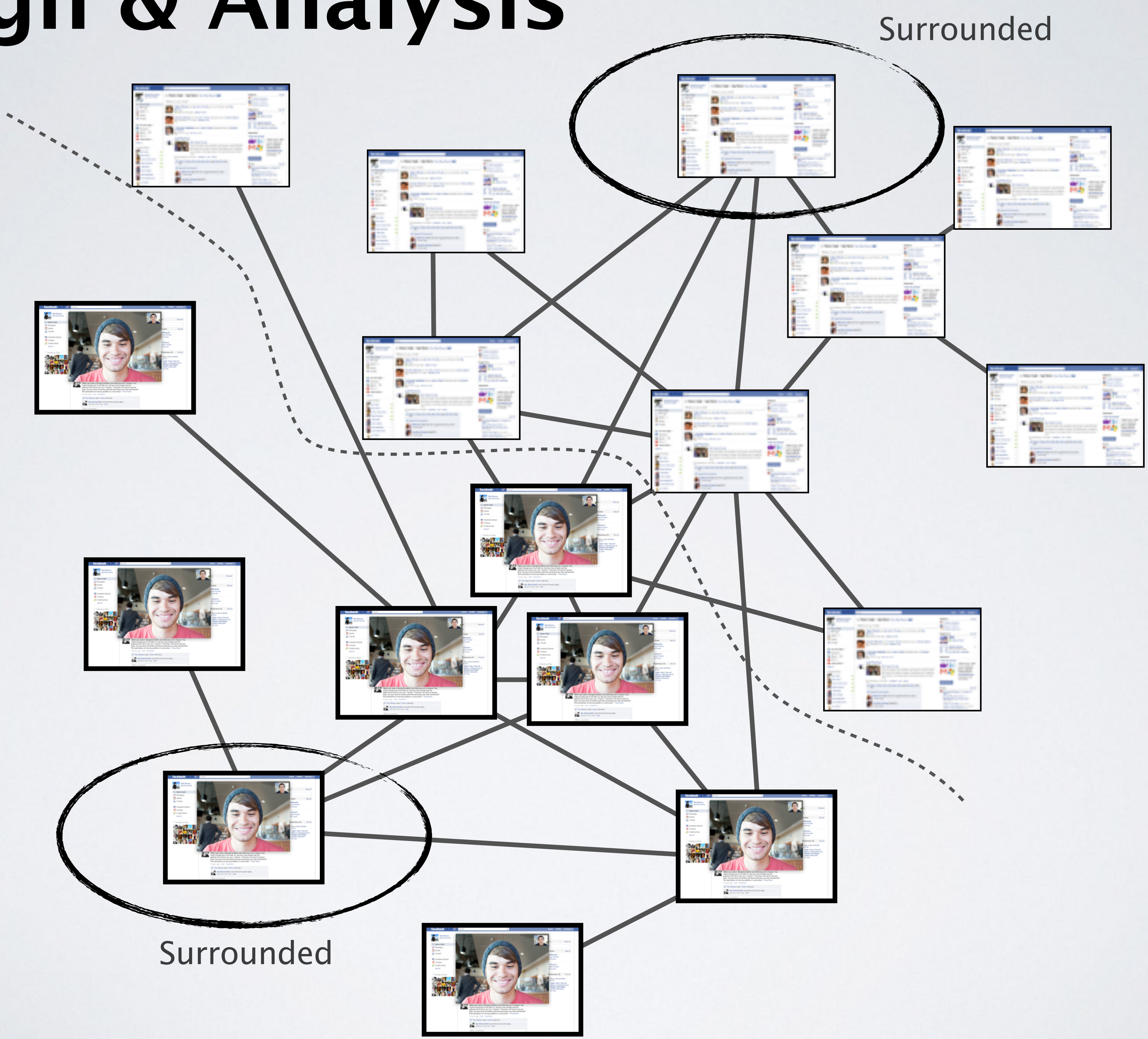
Product design

Content ranking models

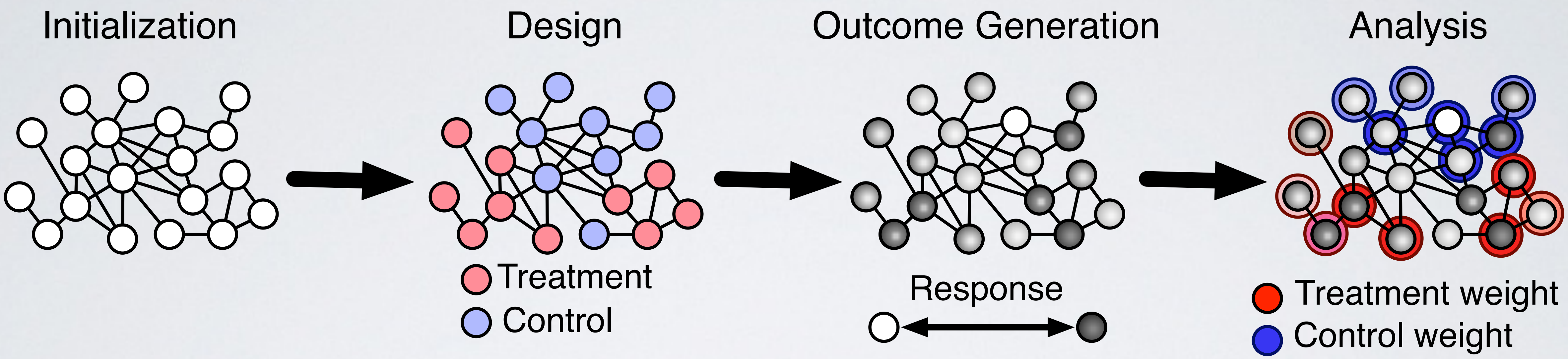
Design & Analysis



Design & Analysis



Network Experimentation Process



- **Initialization:** An empirical graph or graph model
- **Design:** Graph cluster randomization
- **Outcome generation:** Observe behavior (or simulate)
- **Analysis:** Discerning effective treatment

Initialization

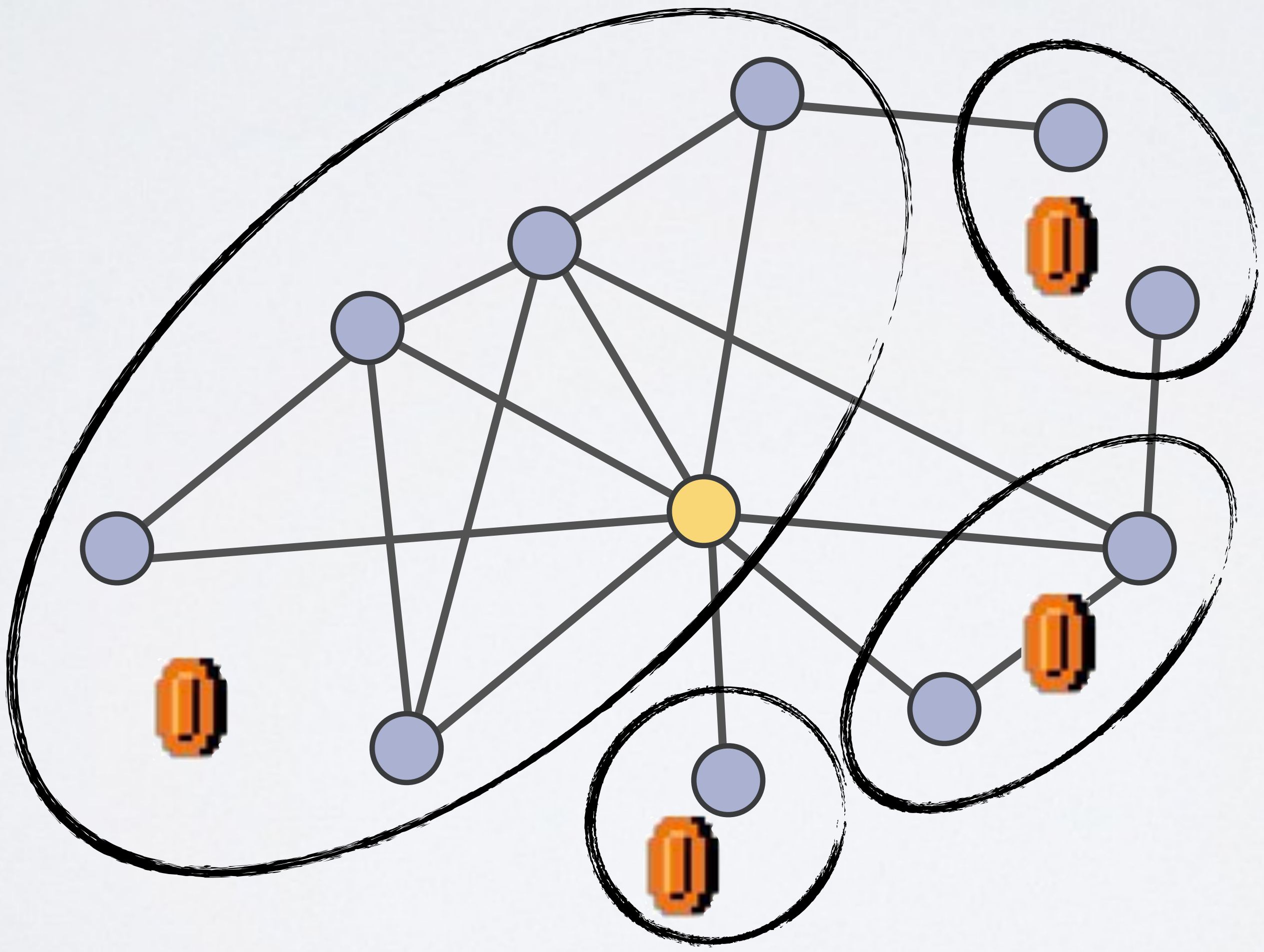
Design: how to assign?

Design: how to assign?

Design: how to assign?

Design: Graph cluster randomization

- Partition graph into clusters (small, balanced clusters preferable)
- Assign each cluster to treatment with probability **q**
- Assign all vertices to their cluster's treatment



More information: Ugander–Karrer–Backstrom–Kleinberg, KDD 2013

Outcome generation

- **Nonparametric structural equation model** for observed outcomes, where outcomes are a function of vertex i's k_i neighbors' prior behavior:

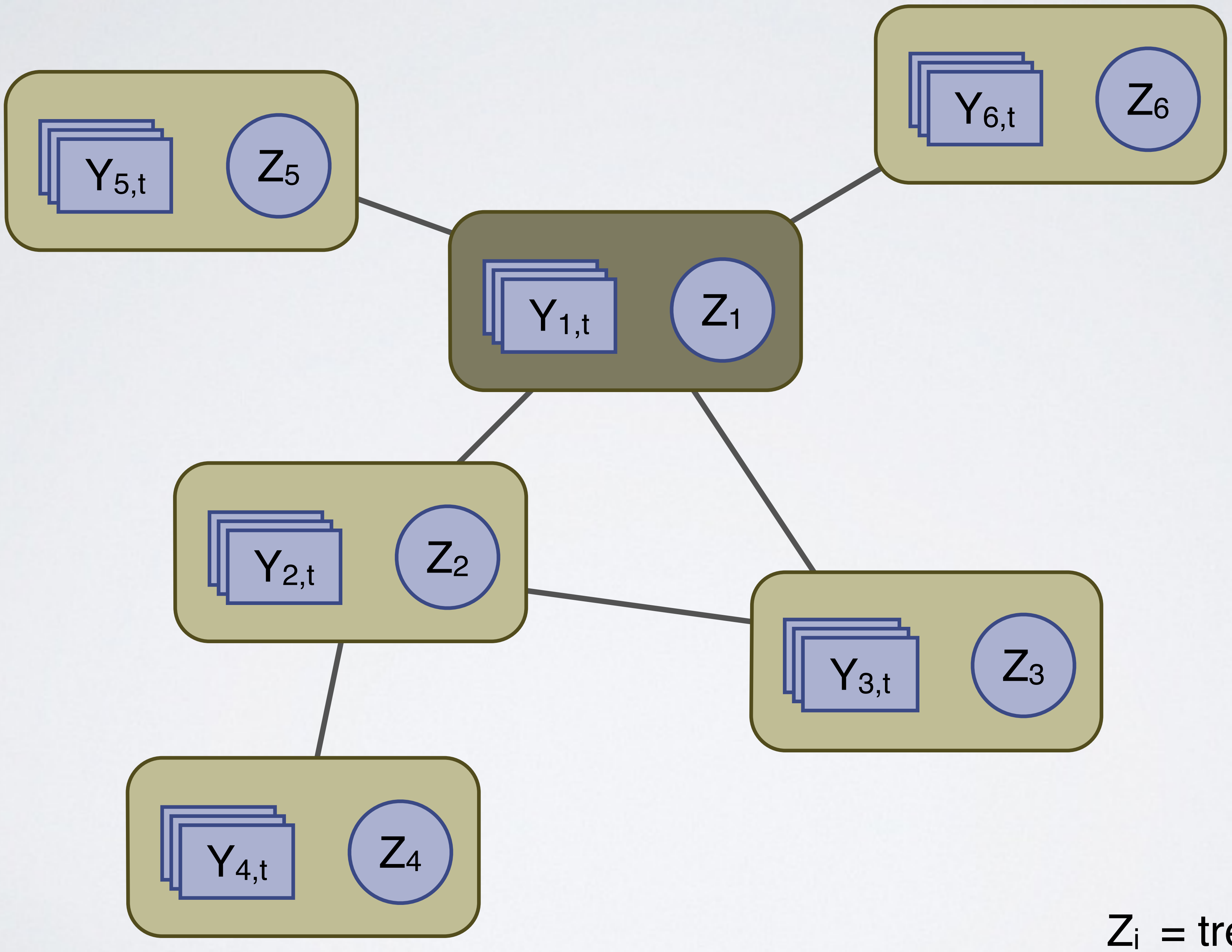
$$h_{i,t}(\cdot) : \mathbb{Z} \times \mathbb{Y}^{k_i} \times \mathbb{U}^N \to \mathbb{Y}$$

- Example: Utility linear-in-means used in simulations

$$Y^*_{i,t} = \alpha + \beta Z_i + \gamma \frac{A'_i Y_{i,t-1}}{k_i} + U_{i,t}$$

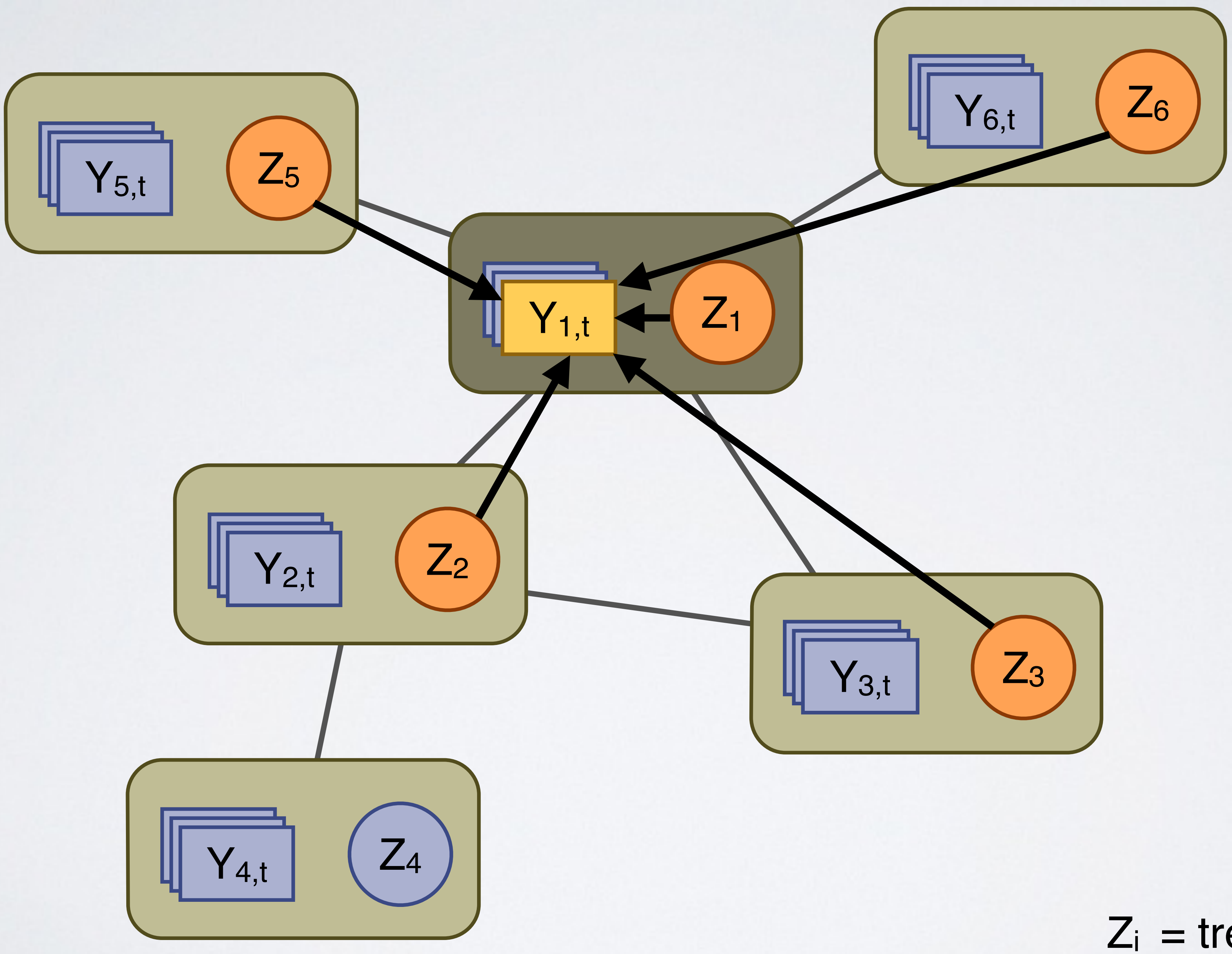
$$Y_{i,t} = \mathbf{1}[Y^*_{i,t} > 0]$$

Treatment vs. Behavior Mediation



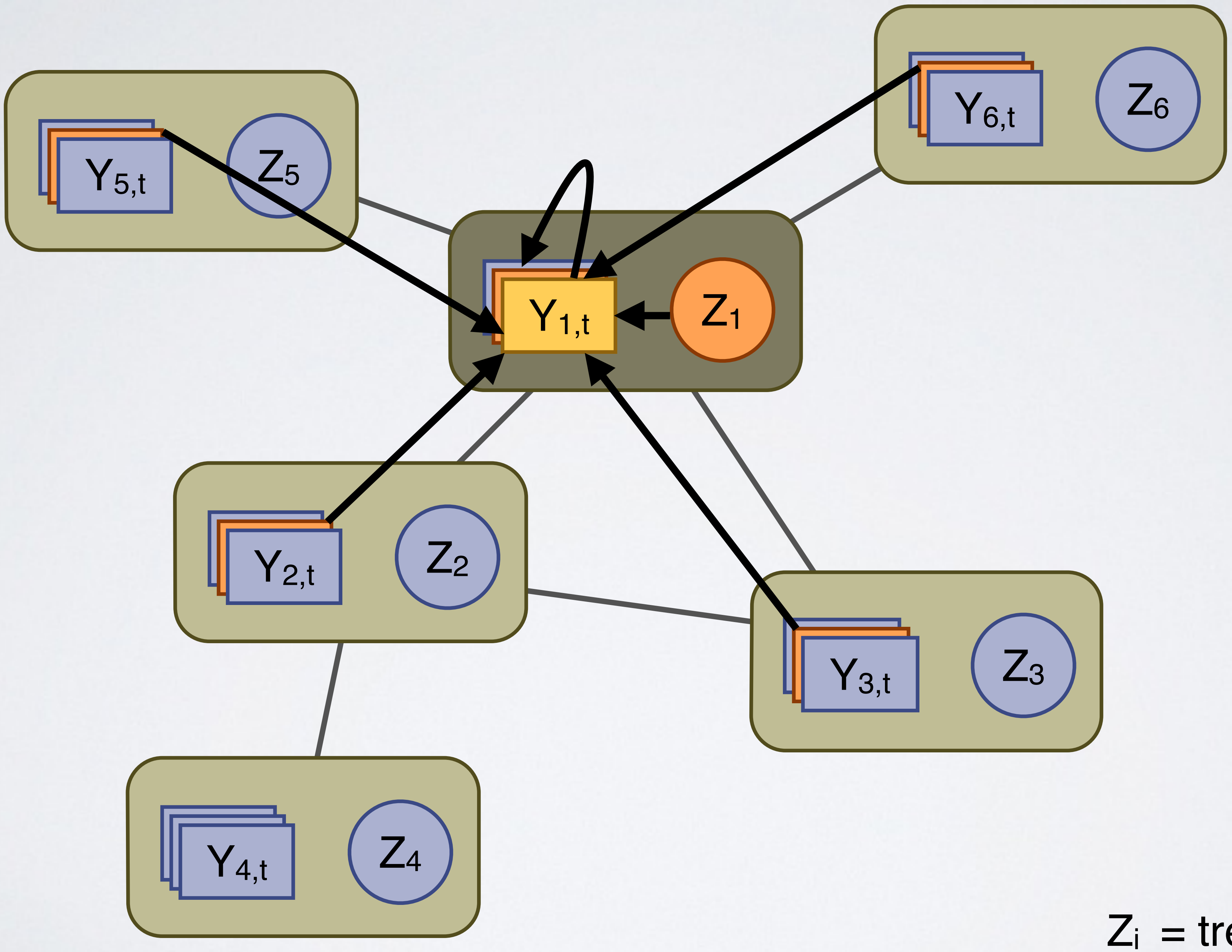
Z_i = treatment
$Y_{i,t}$ = response at time t

Treatment vs. Behavior Mediation



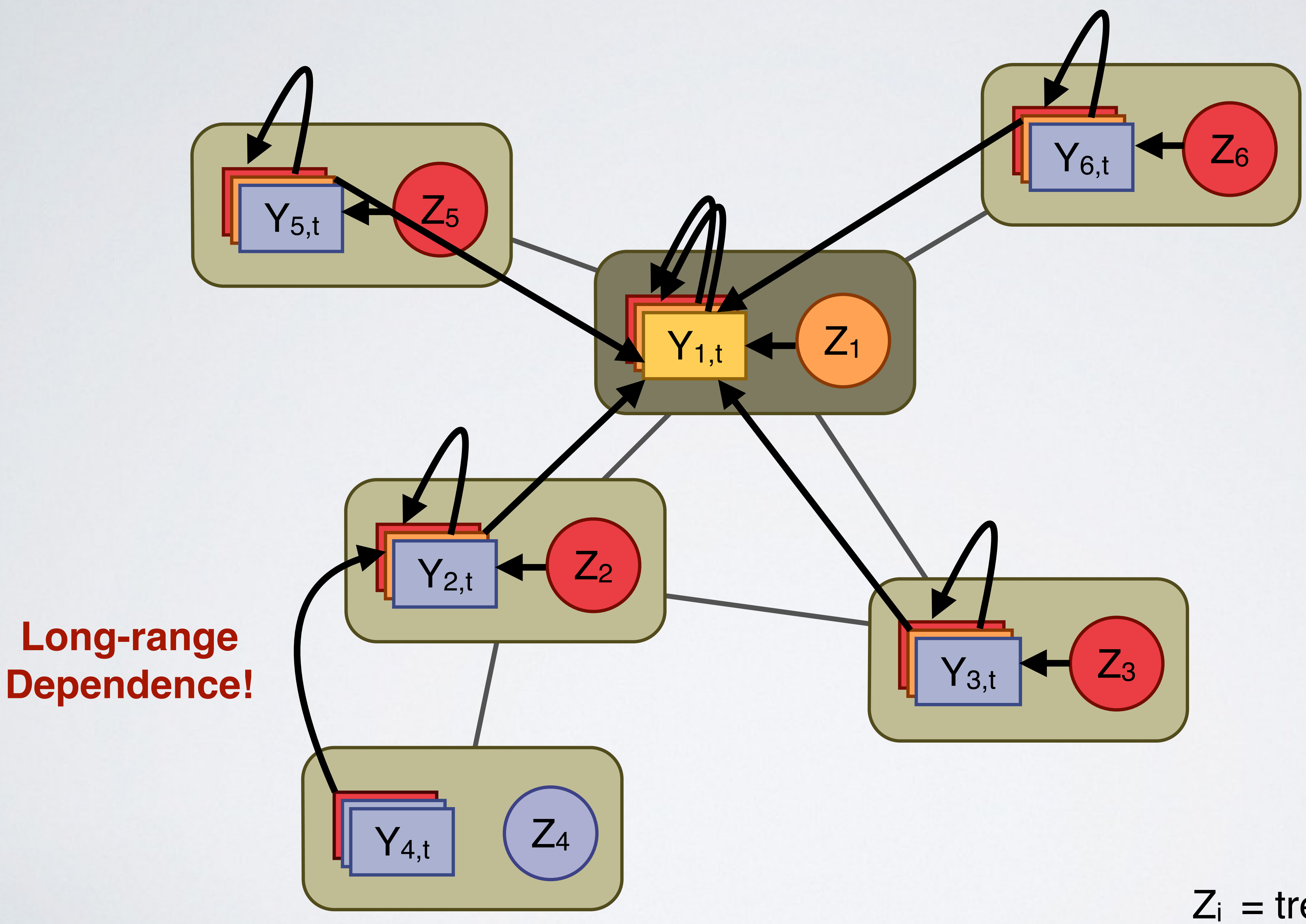
Z_i = treatment
$Y_{i,t}$ = response at time t

Treatment vs. Behavior Mediation



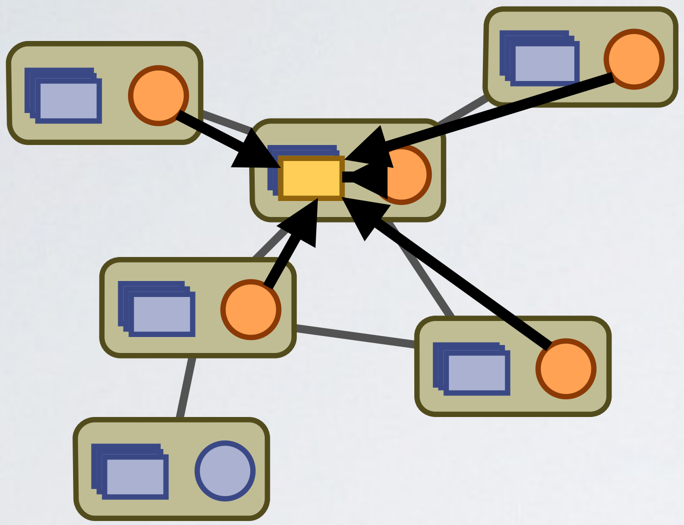
Z_i = treatment
$Y_{i,t}$ = response at time t

Treatment vs. Behavior Mediation



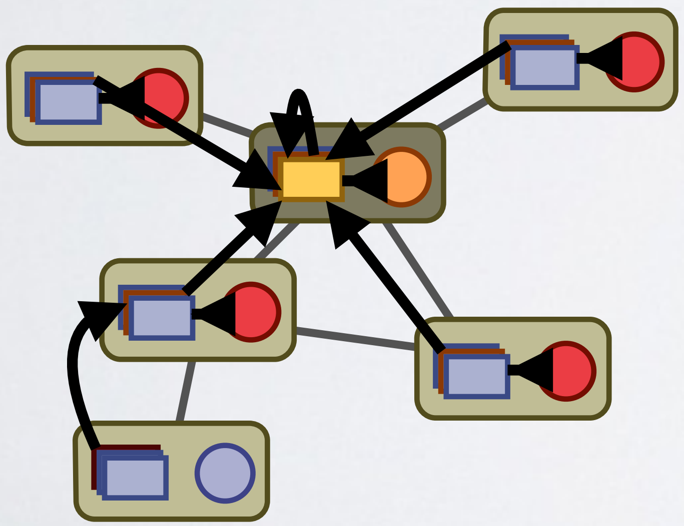
Z_i = treatment
$Y_{i,t}$ = response at time t

Treatment vs. Behavior Mediation



Only treatment matters

- No long-range dependence
- Unbiased estimators easy



Behavior matters

- Adds long-range dependence
- Estimator needs model of behavior
- Bias is tricky
- Realistic

Analysis

- Interested in average treatment effect (ATE)
- Population estimands:

$$\mu_g^d(z) = \frac{1}{N} \sum_i \mathbb{E}^d[Y_i | g_i(Z) = g_i(z)]$$

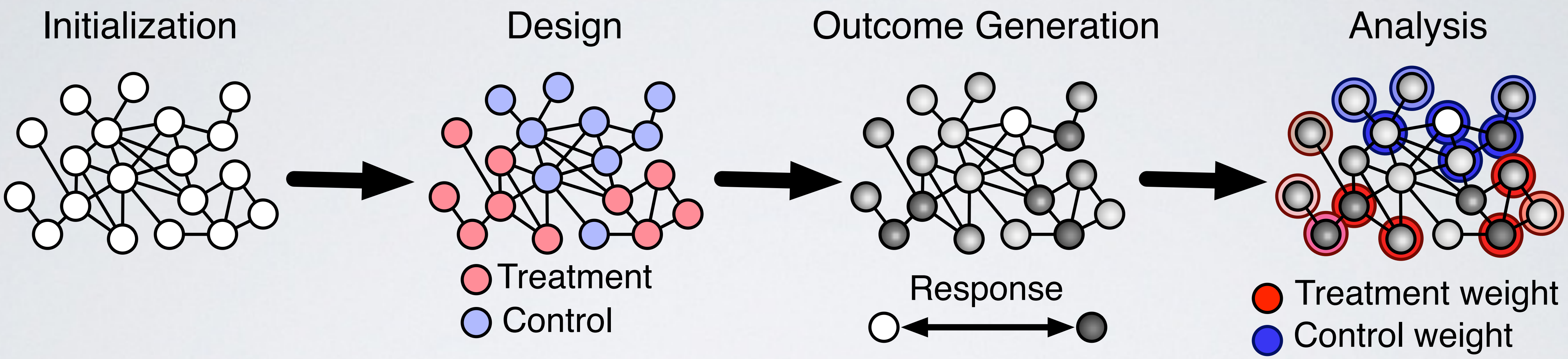
$$\tau_g^d(z_1, z_0) = \mu_g^d(z_1) - \mu_g^d(z_0)$$

  - indices: experimental design **d**, effective treatment **g**

- **Examples of g**: Individualistic treatment response (ITR),
  Neighborhood treatment response (NTR),
  Fractional neighborhood treatment response (FNTR)

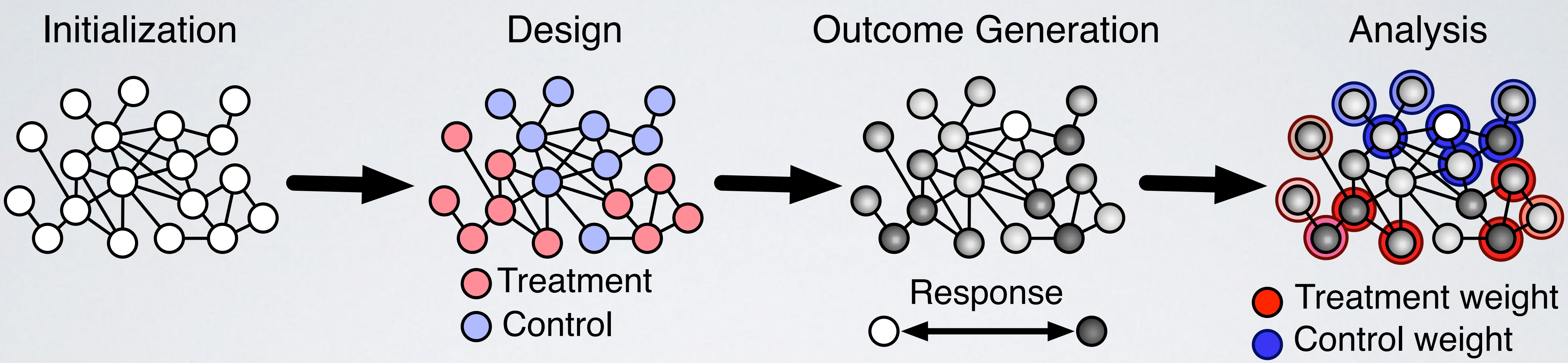
(see Manski, 2013)

Network Experimentation Process



- **Initialization:** An empirical graph or graph model
- **Design:** Graph cluster randomization
- **Outcome generation:** Observe behavior (or simulate)
- **Analysis:** Discerning effective treatment

Network Experimentation Process



- **Initialization:** An empirical graph or graph model
- **Design:** Graph cluster randomization
- **Outcome generation:** Observe behavior (or simulate)
- **Analysis:** Discerning effective treatment

Does design/analysis reduce bias? RMSE?

Bias reduction from design

- **Summary:** For linear outcomes model, if responses are monotonically increasing in treatment, can prove that graph cluster randomization reduces bias.

- **Theorem**:

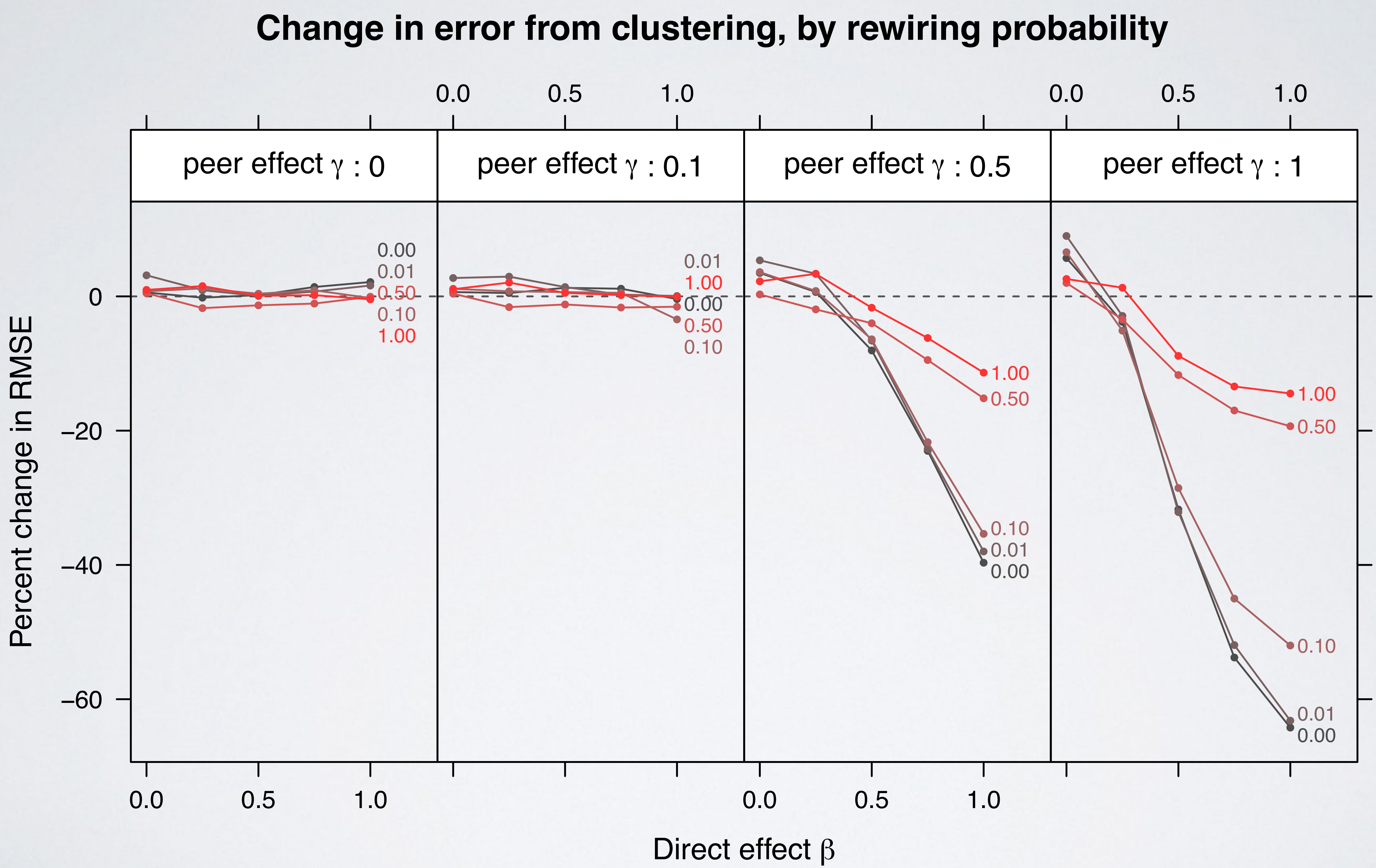
Assume we have a linear outcome model for all vertices $i \in V$ such that

$$\mathbb{E}_U[Y_i(z, U)] = a_i + \sum_{j \in V} B_{ij} z_j$$

and further assume that $Y_i(z, u)$ is monotonically increasing in z for every $u \in \mathbb{U}^N$ and vertex i such that $B_{ij} \geq 0$.

Then for some mapping of vertices to clusters, the absolute bias of $\tau^d_{\mathrm{ITR}}(1, 0)$ when d is graph cluster randomization is less than or equal to the absolute bias when d is independent assignment, with a fixed treatment probability p.

RMSE reduction from design



In simulations on SM/dcSBM networks, up to 60% reduction in RMSE

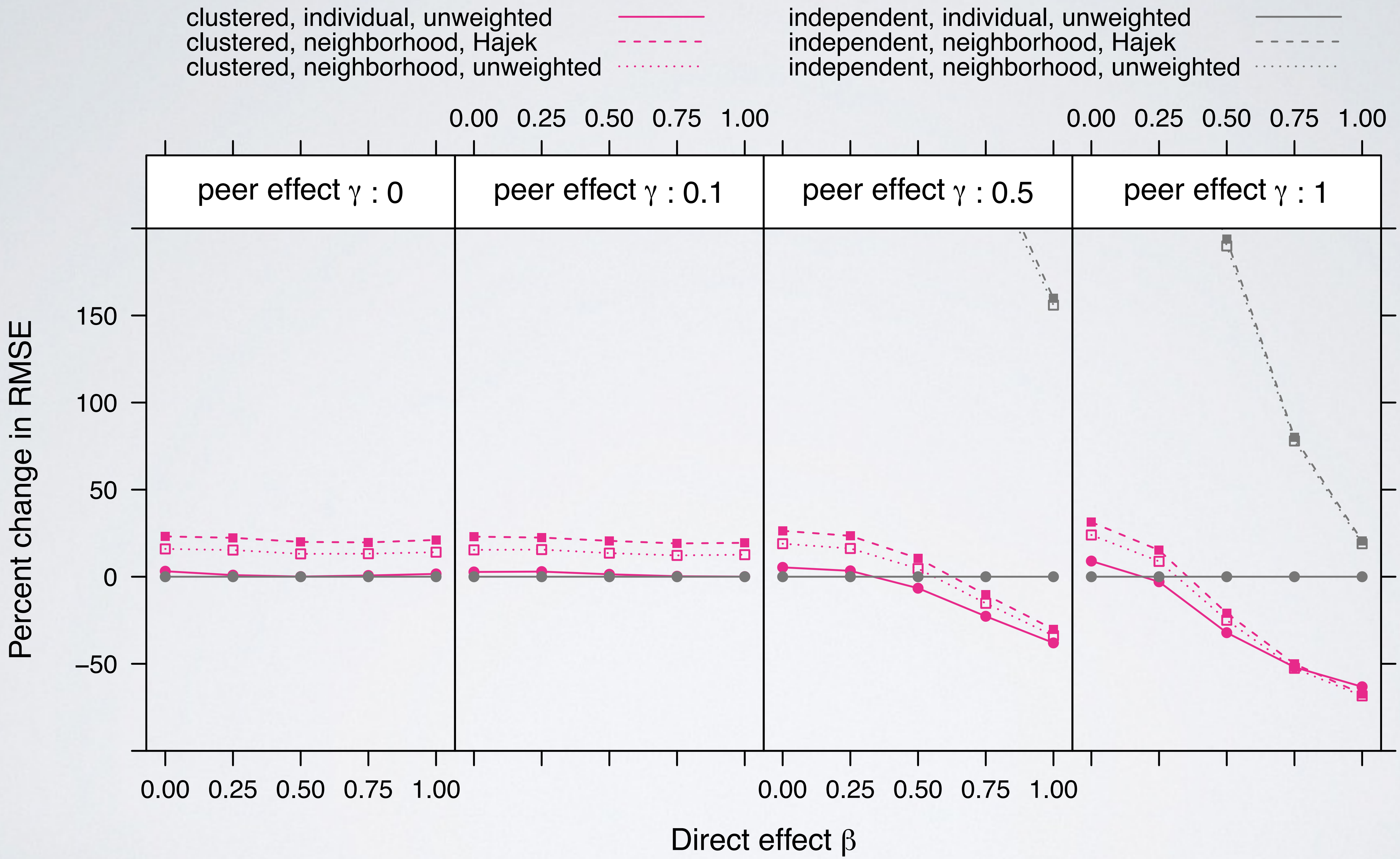
Bias reduction from analysis

- **Summary:** For independent random assignment design, if responses are monotonically increasing or decreasing in treatment, can prove that more restrictive specification of treatment has lower bias than less restrictive specification (e.g. NTR more restrictive ITR).

- **Theorem**:

Let $g^A(\cdot)$ and $g^B(\cdot)$ be vectors of such functions where $g_i^A(\cdot)$ is more restrictive than $g_i^B(\cdot)$ for every vertex i, and let independent random assignment be the experimental design. A sufficient condition for estimand $\tau_{g^A}^{\text{ind}}(1,0)$ to have less than or equal absolute bias than $\tau_{g^B}^{\text{ind}}(1,0)$, is that we have monotonically increasing responses or monotonically decreasing responses for every vertex with respect to z.

RMSE reduction from analysis



…but RMSE can go up considerably: +400%. Or in some regimes: −50%.

Conclusions

- Unbiased ATE estimation unlikely for network experiments
- **Bias**: Reduced a lot by Design/Analysis , under assumptions
- **RMSE**: Still reduced considerably in some regimes, be careful
- Papers:
  - J Ugander, B Karrer, L Backstrom, J Kleinberg. Graph Cluster Randomization: Network Exposure to Multiple Universes, KDD'13.
  - D Eckles, B Karrer, J Ugander. Design and analysis of experiments in networks: Reducing bias from interference (arXiv)

Co-authors

Brian Karrer, Dean Eckles