Interventions on Network Formation

Robert Ward
Marc Santolini

Georgia Tech
Stanford University
Learning Planet Institute

Research

Social networks affect many outcomes

Practice

Interventions to change social networks often fail

Research

Social networks affect many outcomes

Practice

Interventions to change social networks often fail

Example 1: Education

Carrell, Sacerdote and West, 2013, *Econometrica*

UNITED STATES AIR FORCE ACADEMY

Example 2: Healthcare

Funk and Park, 2022, *Working Paper*

Example 3: Finance

Banerjee, et al. 2021, *Working Paper*

SPANDANA

SPANDANA SPHOORTY FINANCIAL LIMITED

And more,

Corporate Retreats Kneeland and Kleinbaum, 2022, *Working Paper*

Job Seeking Online Rajkumar et al., 2022, *Science*

Conferences Zajdela et al., 2022, *PhysRevRes*

Organizational Structure Kleinbaum et al., 2021, *OrgSci*

Firm Innovation Argyres et al., 2020, *SMJ*

Startup Bootcamps Hasan and Koning, 2019, *SMJ*

Interfirm Relationships Cai and Sziedl, 2018, *QJE*

Grant Applications Boudreau et al., 2018, *Rev Econ Stat*

Natural Disasters Phan and Airoldi, 2015, *PNAS*

Business Mixers Ingram and Morris, 2007, *ASQ*

How can we design better interventions?

Formalize the problem and prove how hard it is to solve

Use an agent-based model to understand

1. How do current designs perform for different *networks*, *outcomes*?
2. When and why do they fail?
3. How sensitive are those answers to the information and authority of planners have at their disposal?
4. How can we design around the problems we find?

Formalization

Consider a social planner

Who controls the institutions of network formation.

They observe a network G and an outcome $Y(G)$ that depends on it.

Their objective is to maximize outcomes

They need to solve two problems

1. Identify which relationships should be added or removed
2. Specify inducements to create those changes

Consider a social planner

Who controls the institutions of network formation.

They observe a network G and an outcome $Y(G)$ that depends on it.

Their objective is to maximize outcomes

They need to solve two problems

1. Identify which relationships should be added or removed
2. Specify inducements to create those changes

The Network Alteration Problem

Given a network $G = (V, E, X)$ and outcome $Y(G)$

Nodes Edges Attributes

Find the altered network G' that maximizes $Y(G)$ without adding or removing more than B edges.

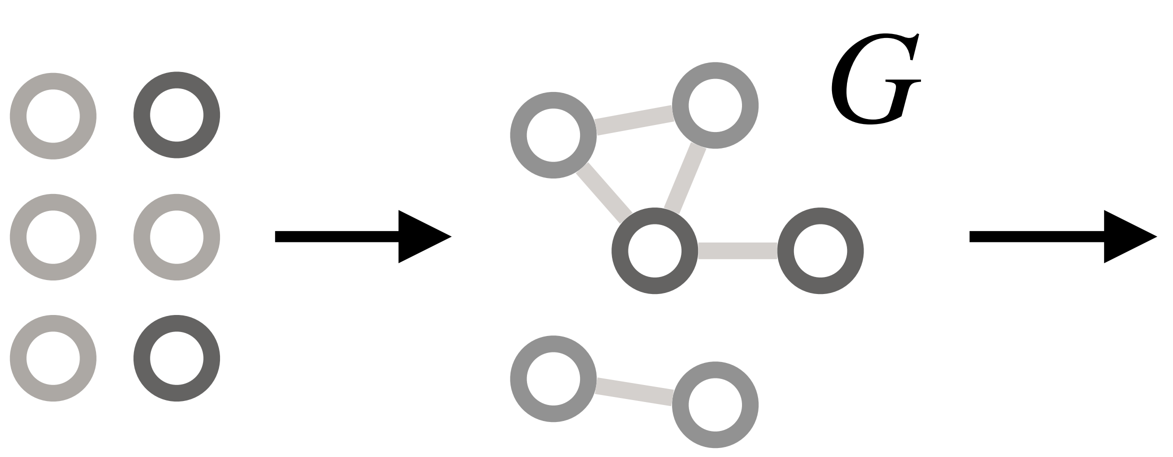
Theorems: this problem is NP-Hard and APX-Hard

There is no efficient algorithm that we can use to solve the problem.

Model

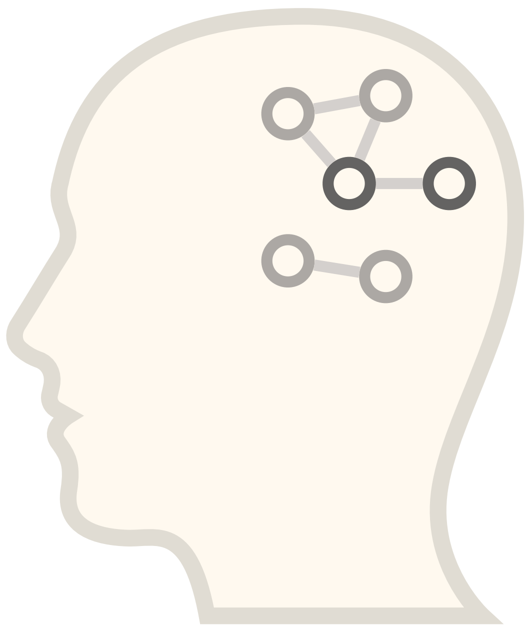
Flow of the agent-based model

Formation

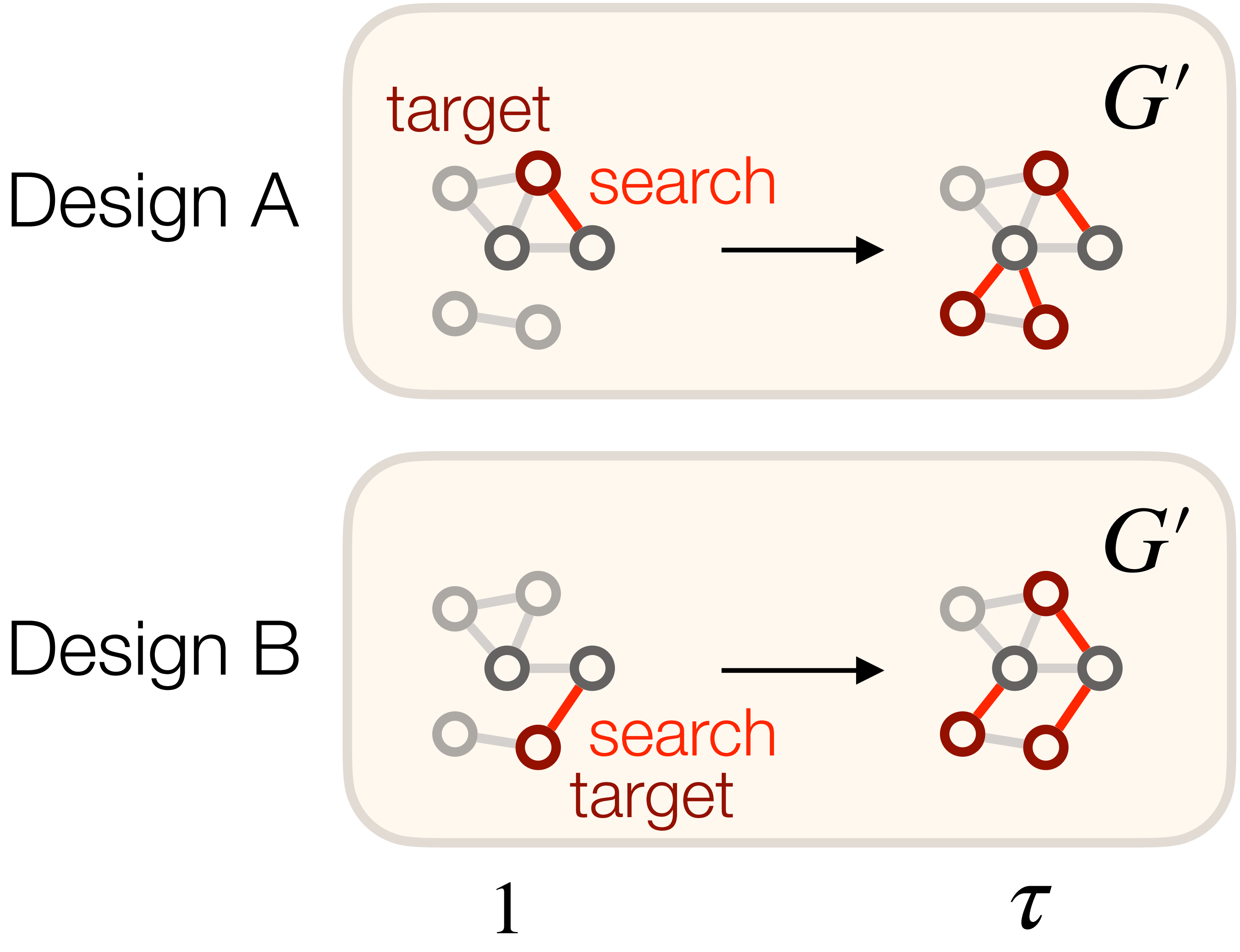


Generate an empirically plausible network

Observation



Intervention



Design and run an intervention

Evaluation

$$\text{ATE} = \frac{1}{\tau}\sum_{u}^{\tau} y_u(G') - y_u(G)$$

Estimate the change in expected outcomes

Monte Carlo experiments

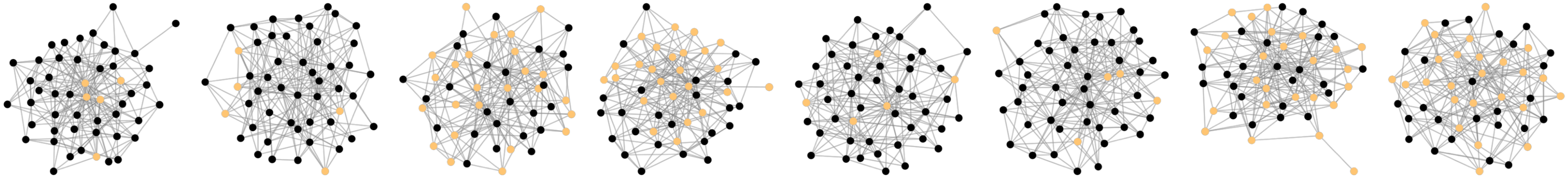
Generate an instance of an empirical class of networks (G1-G8)

Preferential attachment: *weakly scale-free* degree distribution and *degree assortative*

High/low *triadic closure* (clustering)

Balanced or unbalanced attribute *composition* (10% or 50% of nodes have desired attribute)

Choice preferences: *homophily* or *heterophily* on an attribute of interest



G1 G2 G3 G4 G5 G6 G7 G8

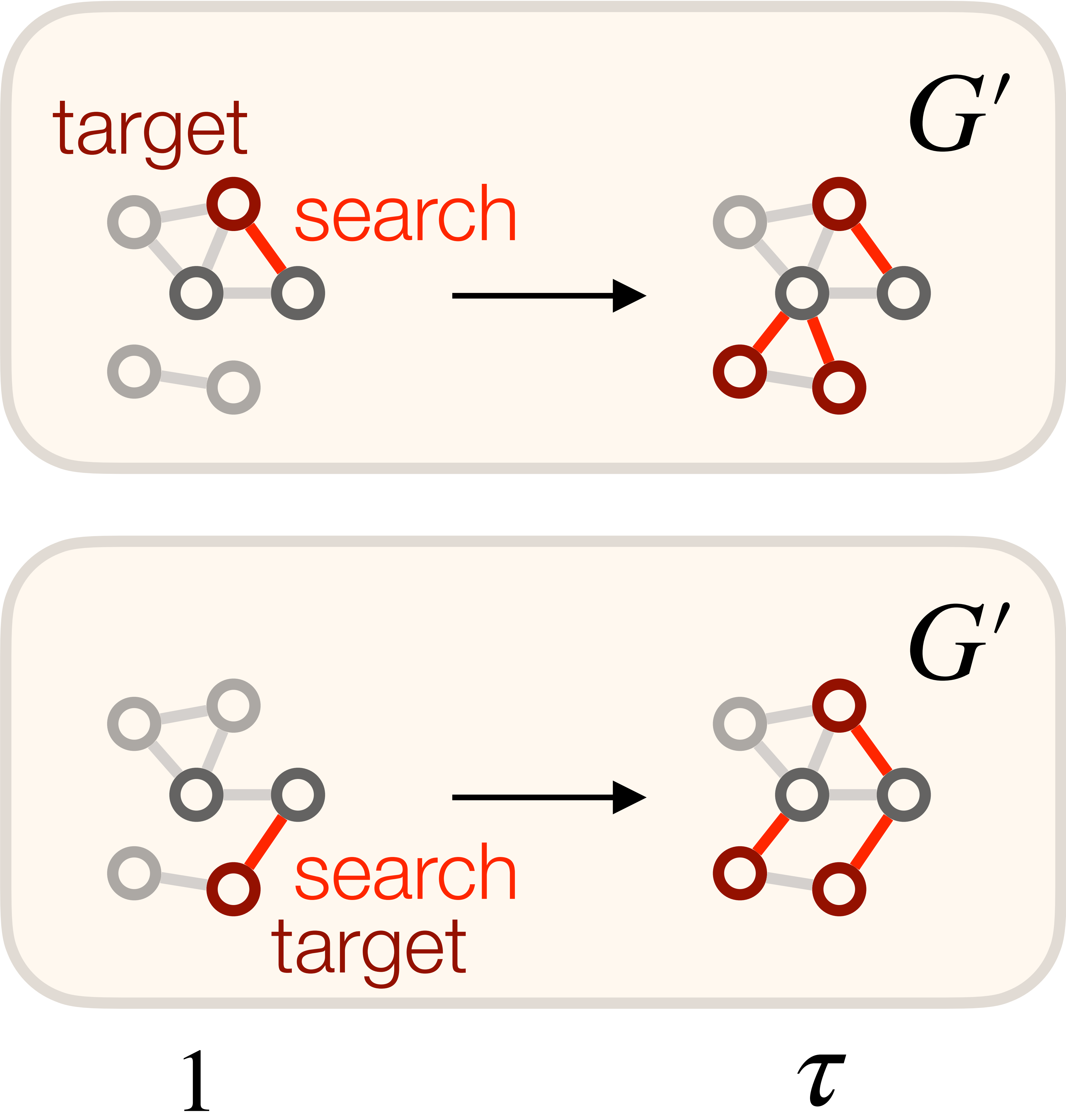
We do this by extending the model of Holme and Kim (2001) and statistically validating that our artificial social networks are consistent with empirical observations.

Monte Carlo experiments

Design an intervention on the network

1. Targeting mechanism that tells us which nodes to treat. *e.g. find the most clustered node*

2. Search mechanism that tells us who to connect them to. *e.g. friends-of-friends (local search) or structurally distant peers (distant search)*



Three algorithms are commonly used in practice

1. <u>Random</u> e.g. Ingram and Morris, 2007; Boudreau et al., 2017
2. <u>Intuition</u> e.g. Carrell et al., 2013
3. <u>Simple planning</u> e.g. Gupta et al 2013, Rajkumar et al., 2022

Monte Carlo experiments

Consider one of four types of outcomes

Y1: threshhold peer effect, (-) clustering effect

Y2: threshhold peer effect, (+) clustering effect

Y3: linear-in-means peer effect, (-) clustering effect

Y4: linear-in-means peer effect, (+) clustering effect

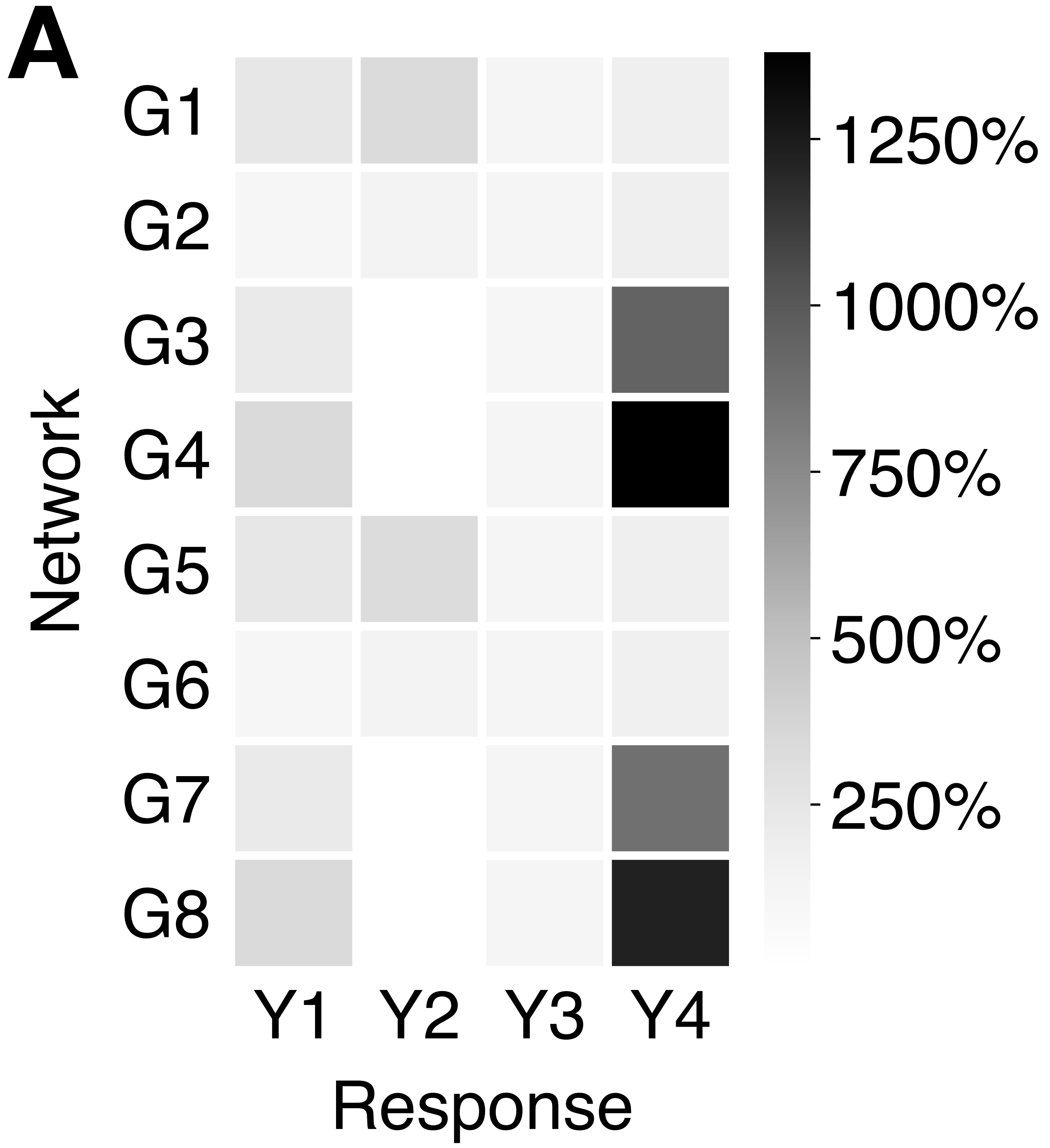
10,000 runs for each *network, outcome, design*

Results

Perfect compliance

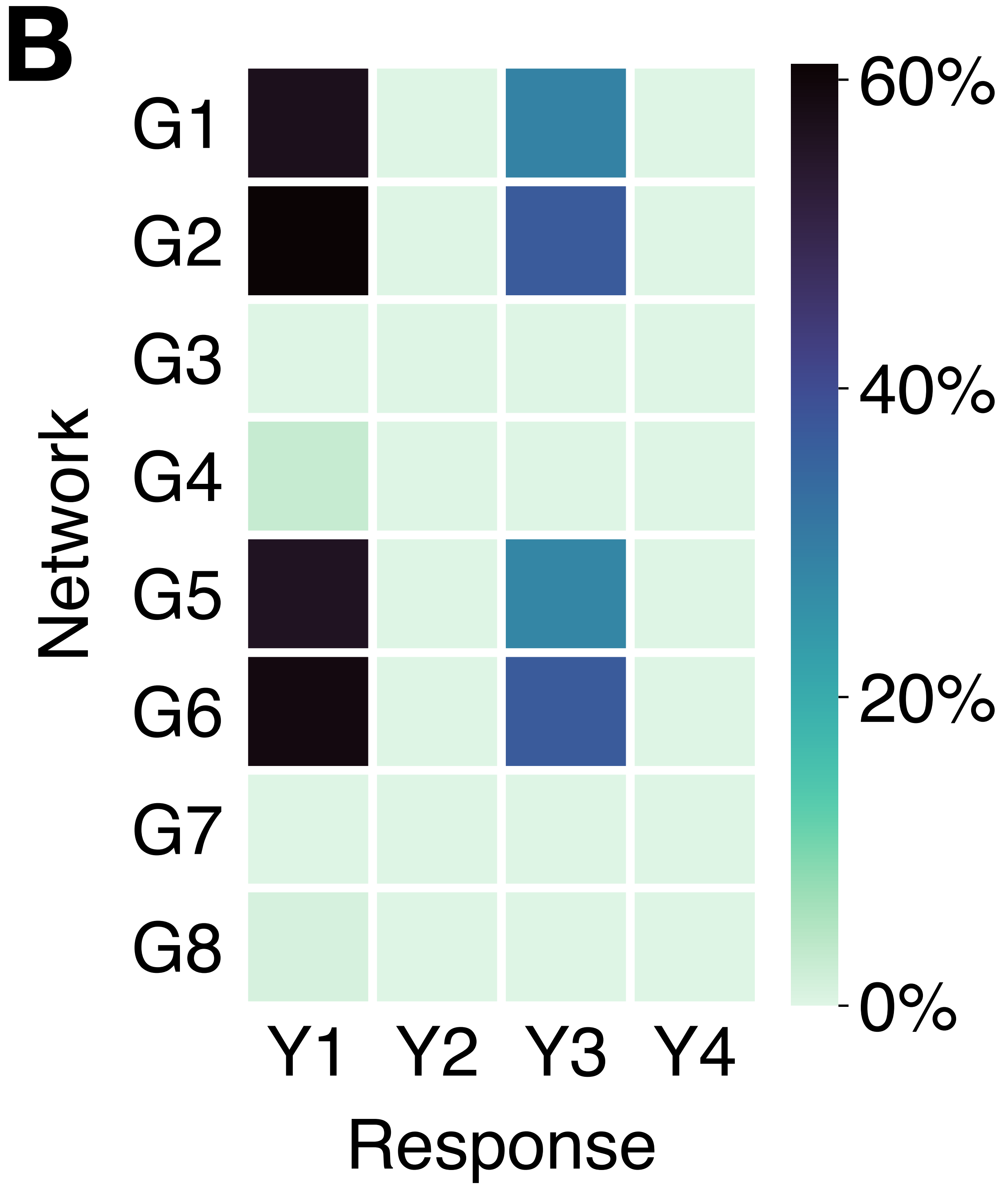
How would different designs perform if there were no problems with implementation?

% increase in ATE from using simple planning



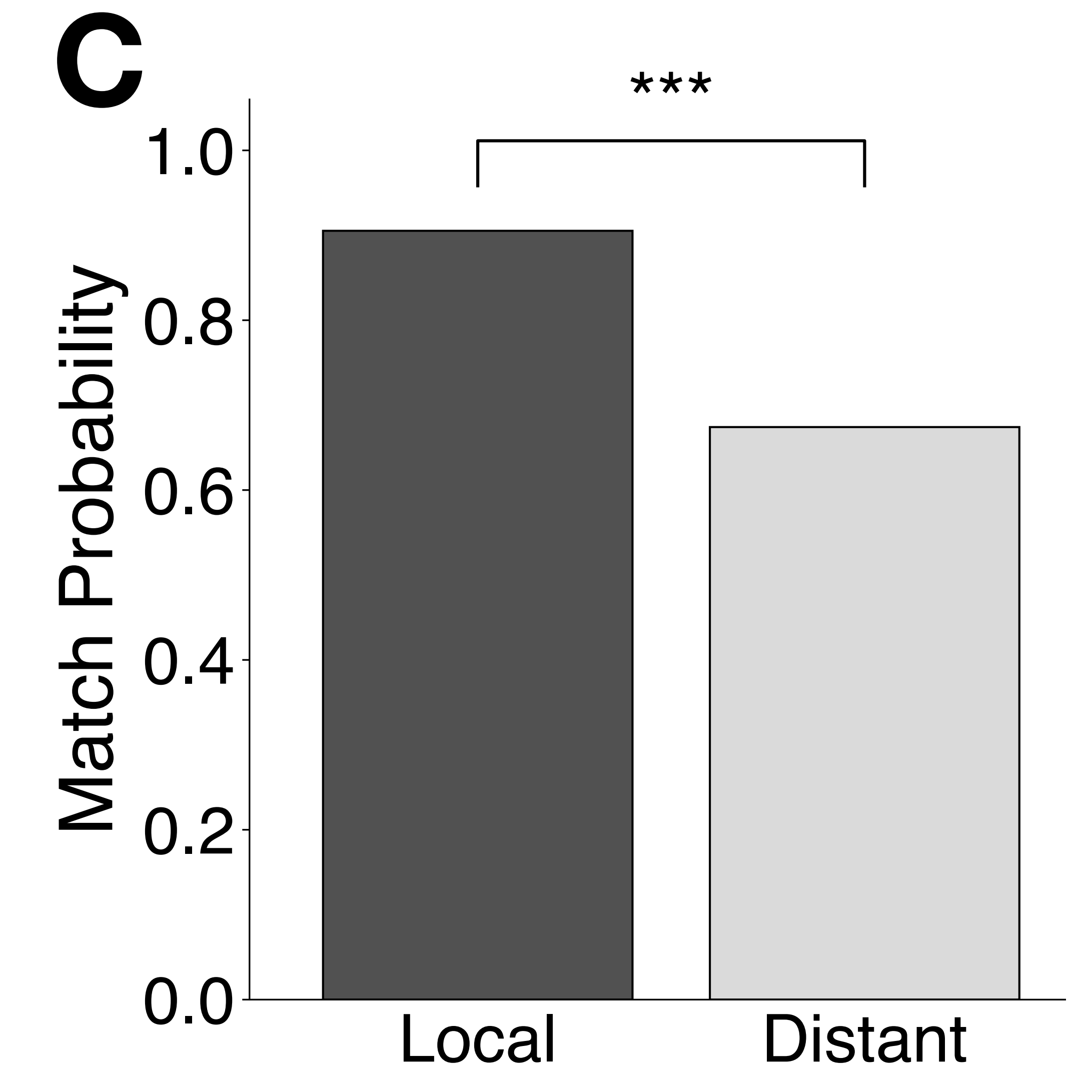
Random Assignment

% increase in ATE from using simple planning



Intuition

Mechanism explaining **B**



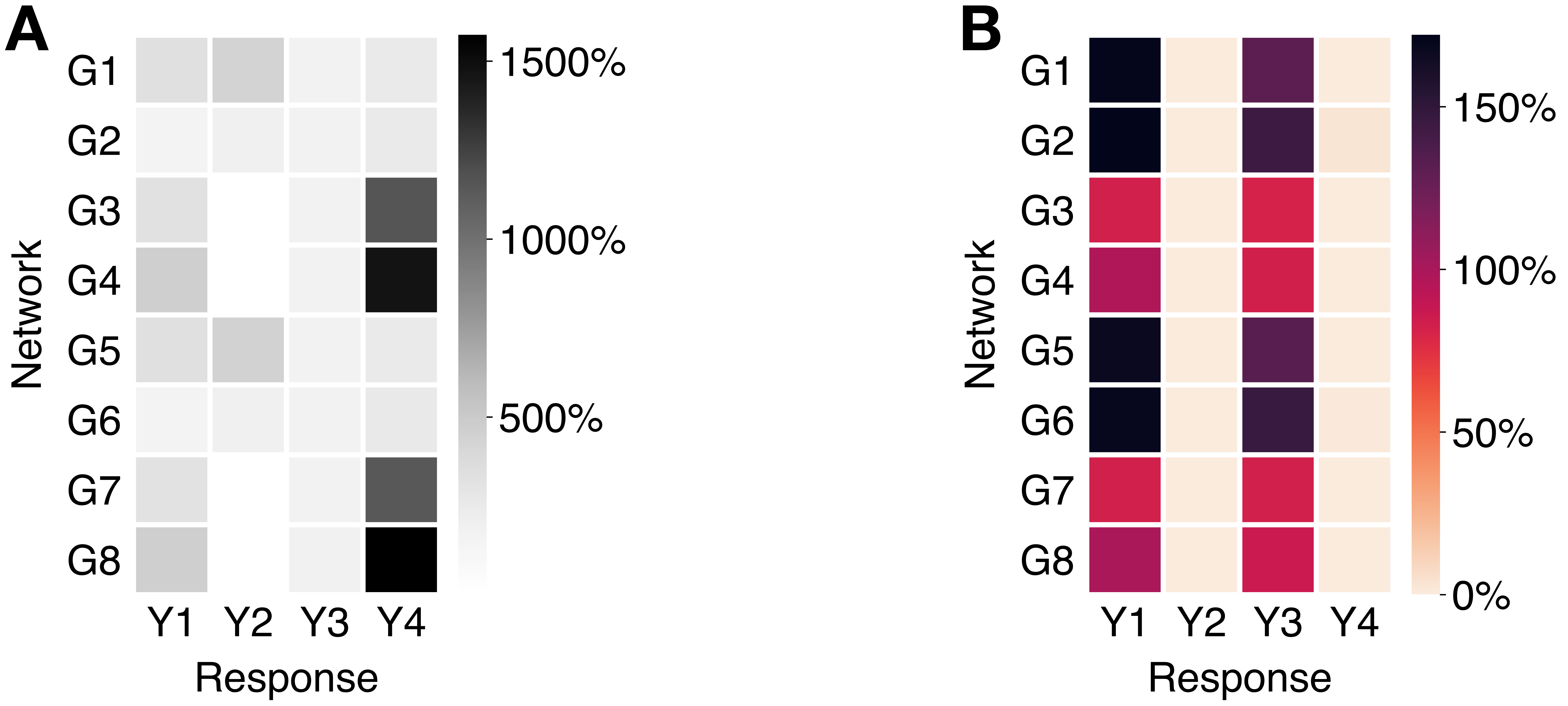
Intuition fails when clustering is detrimental, because we need to induce distant connections, but

Imperfect compliance

What happens when not every node complies with the intervention?

50% if local search, 25% if distant search

% increase in ATE from using simple planning

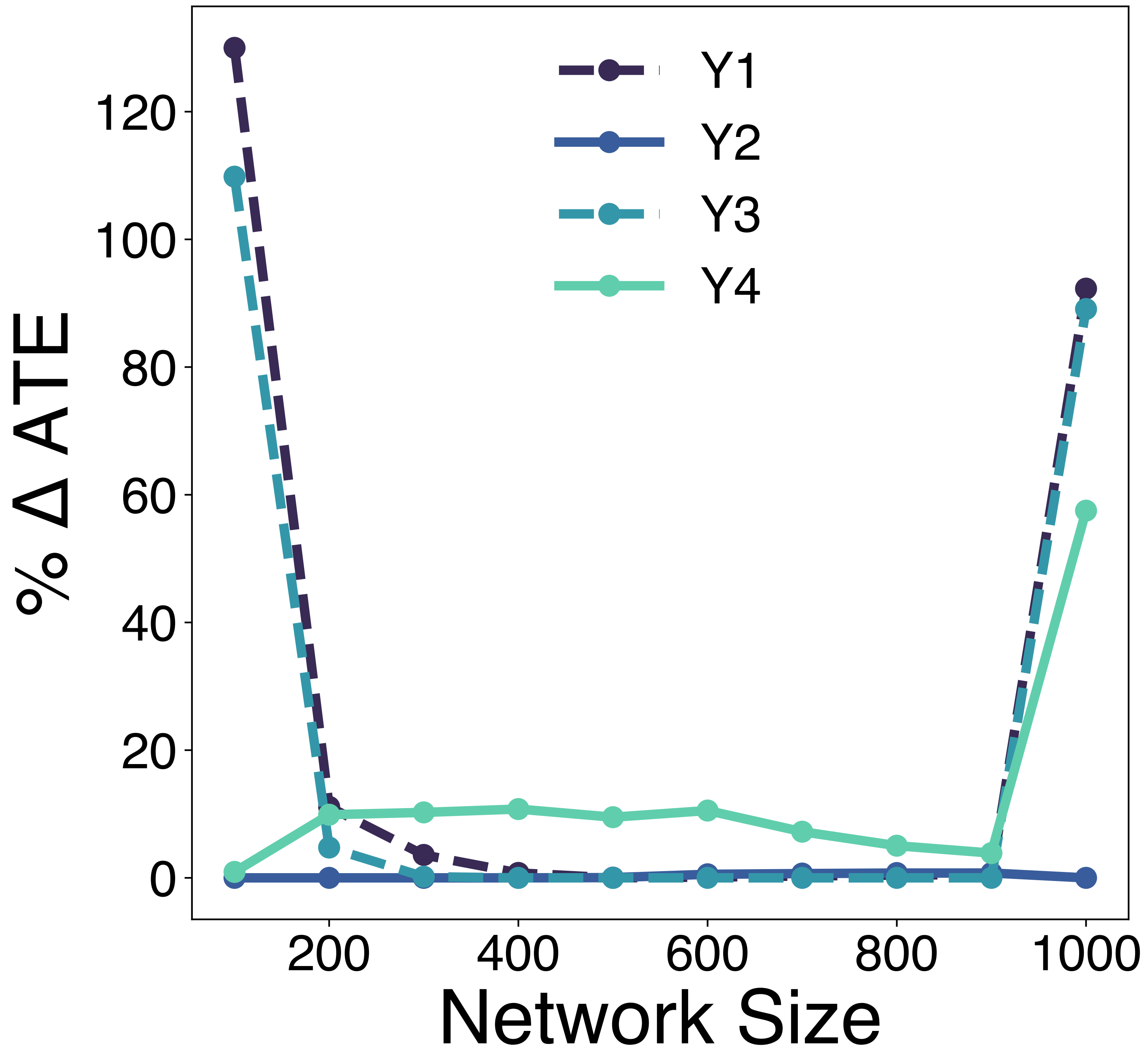


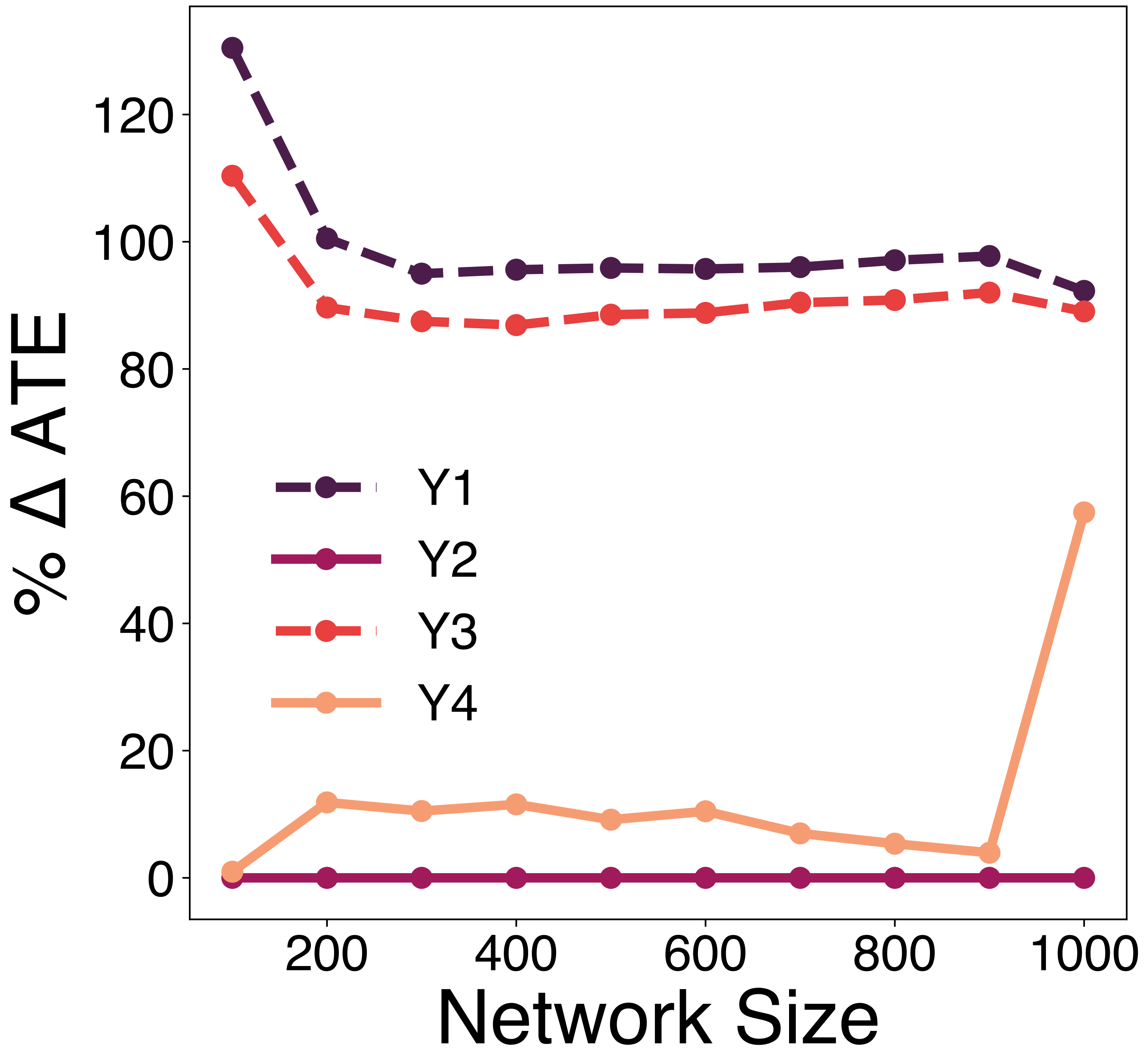
Random Assignment

Intuition

Sensitivity to network size

Percent difference in ATE by using simple planning instead of intuition

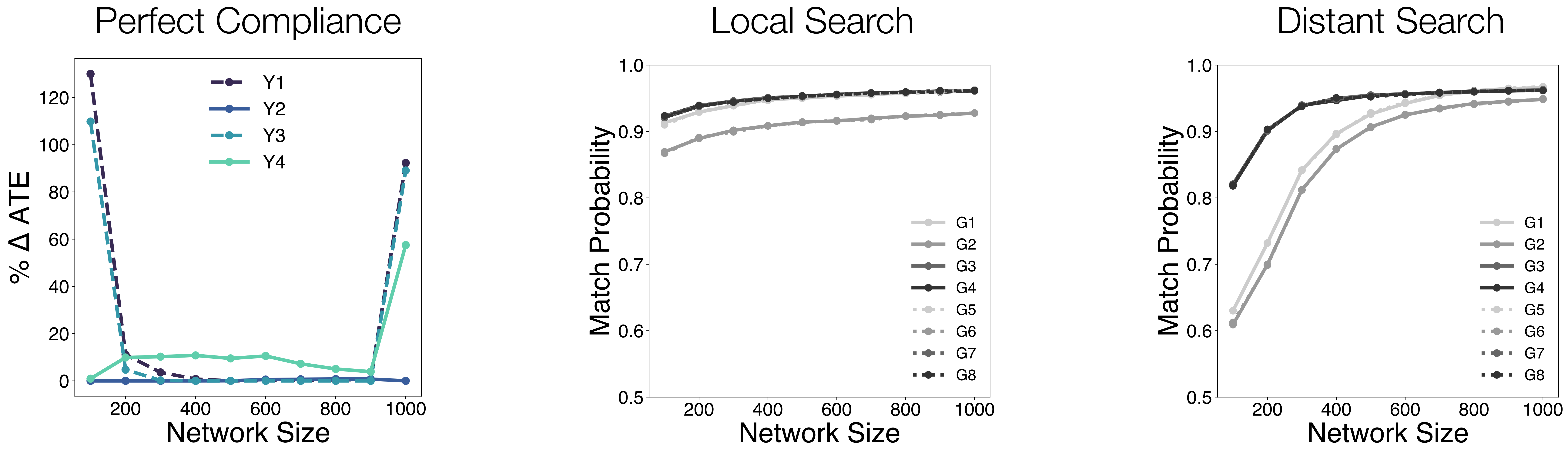




The problem gets better with scale, unless we make realistic assumptions about compliance

Sensitivity to network size

Probability of finding a match through each search mechanism at different scales

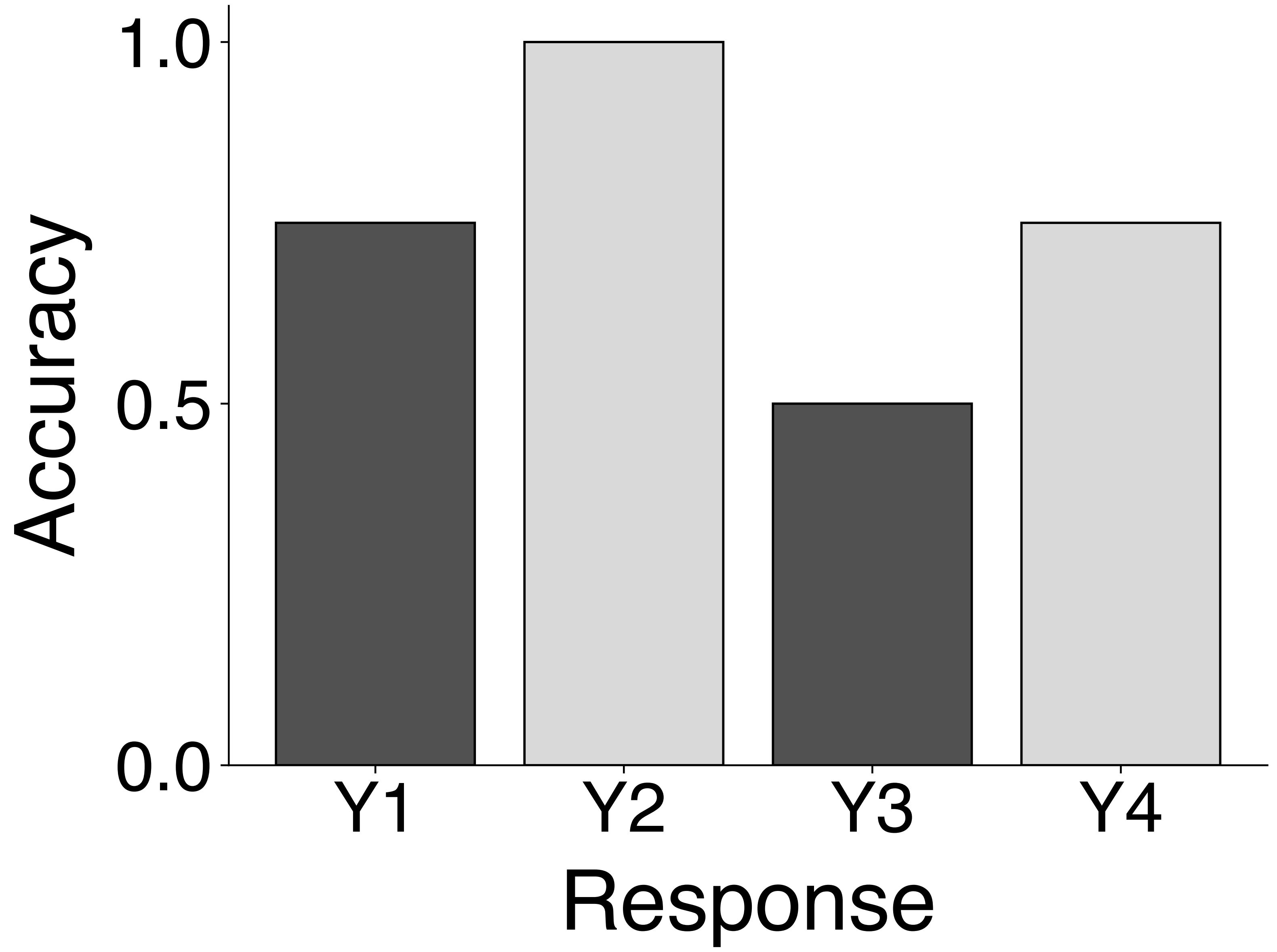


Low probabilities of matching conditional on compliance lead intuition to fail because we simply run out of distant connections to make. We use local search which has a lower marginal treatment effect, but has a higher match rate, giving it a larger average treatment effect than distant search.

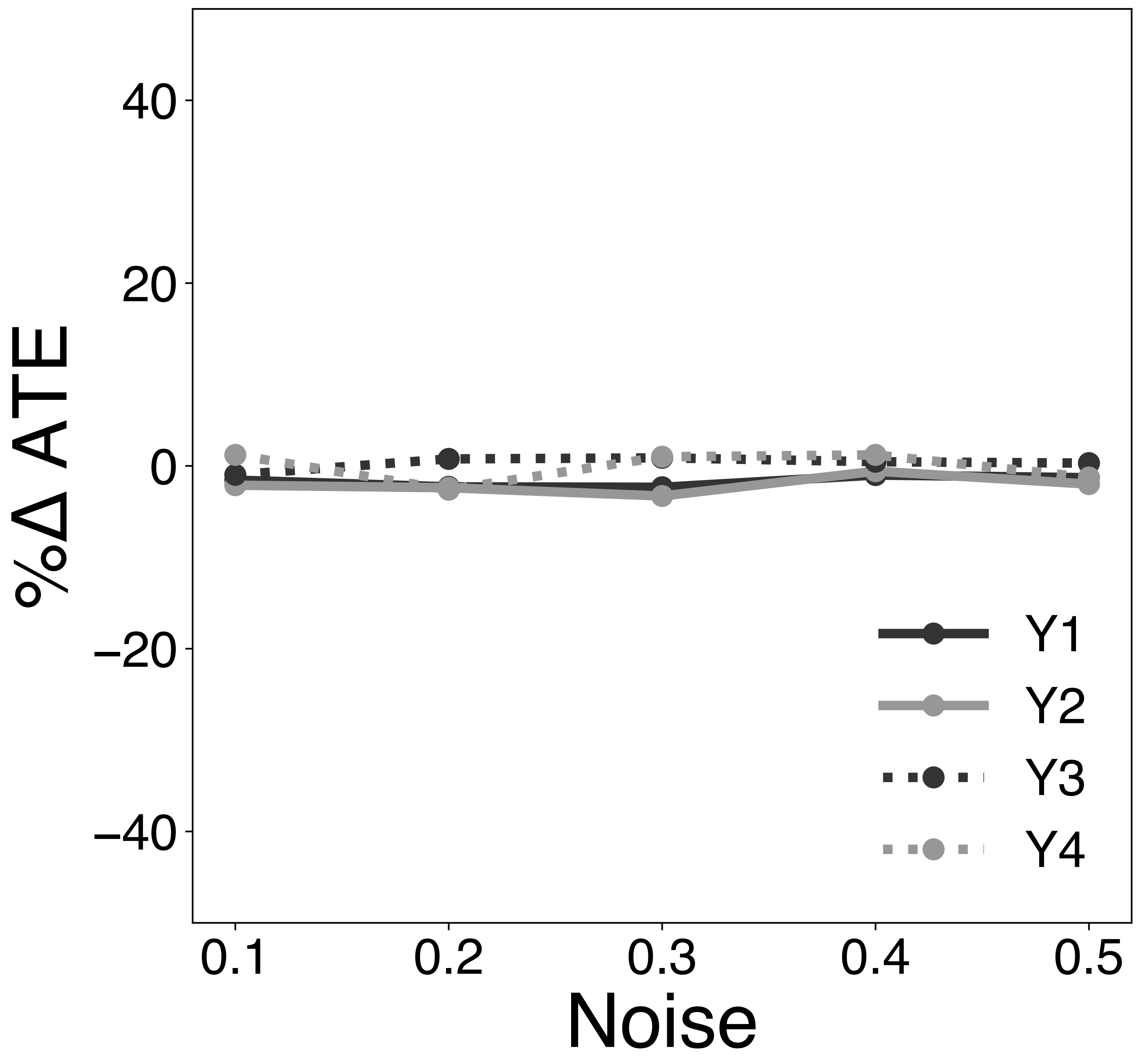
Robustness of the LATE

How do unobserved non-compliance and noise affect our ability to recover the ATE?

When does the LATE recover the same choices as ATE?



Noisy v.s. Clean ATE



The ATE maximizing intervention is not meaningfully affected by noise.

Principles for intervention design

Different network properties affect outcomes and are interdependent, causing a variety of problems that require unintuitive solutions.

1. Measure network composition *and* structure
2. Estimate their their effects on outcomes, LATE seems sufficient.
3. Tailor targeting and search mechanisms to outcomes.
4. Optimal interventions will mix search mechanisms to control global network properties and maximize the number of treated nodes. This is particularly true when clustering is bad, and we can only add edges.
5. Optimize sequence of interventions at once instead of stepwise

Computational design of network interventions

Find a sequence of interventions (targeting+search) that maximizes the ATE

Solve through simulated annealing

Begin with a randomly drawn sequence of interventions.

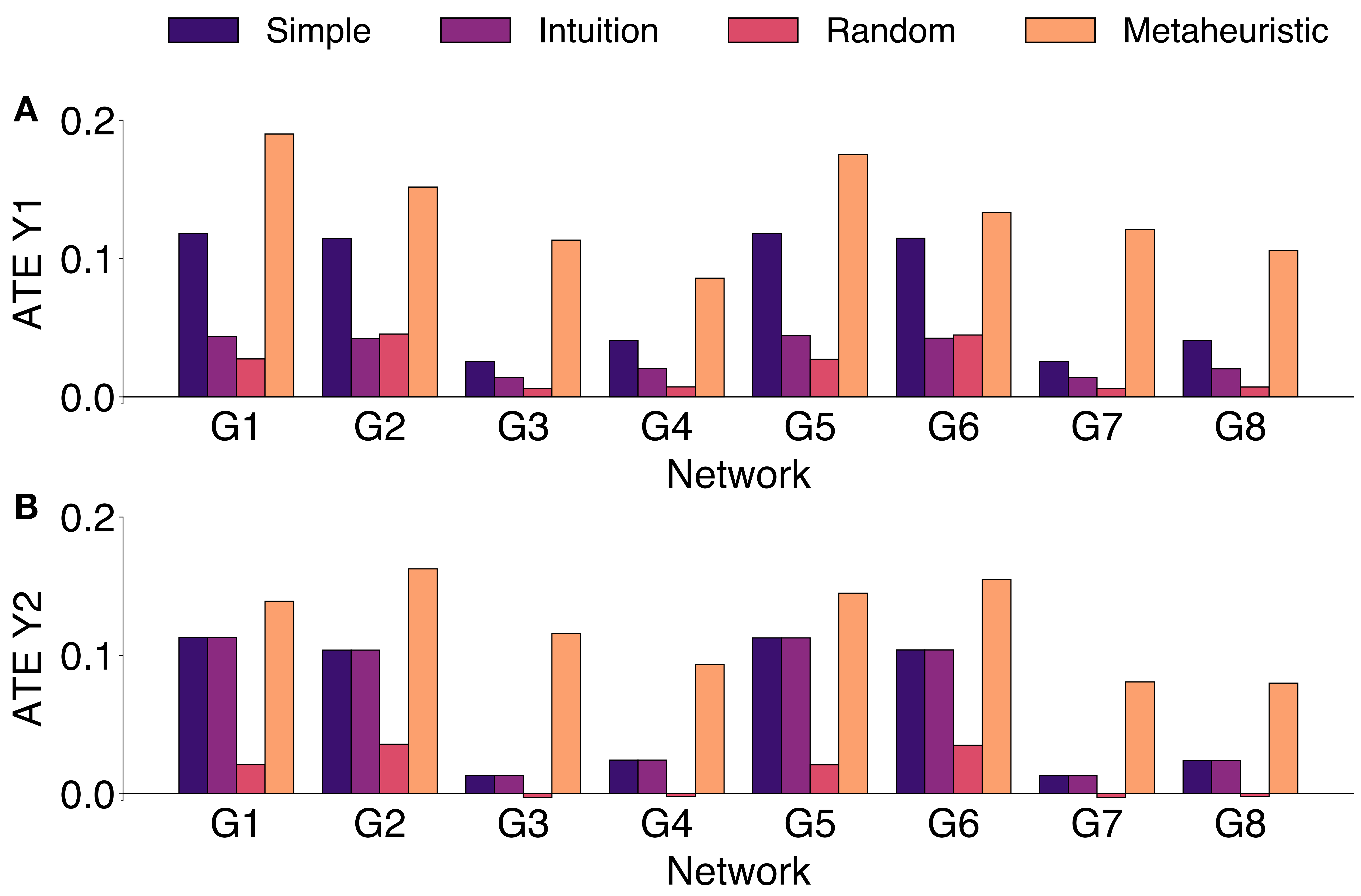
Point mutate the sequence and check the ATE

Always accept higher ATE solution.

Accept lower ATE according to Metropolis criterion.

Decrease temperature over time.

Metaheuristic intervention design



Conclusions

Current work on social networks

How do networks form?
How do they affect outcomes of their nodes?
How can we control processes running over given networks? *e.g. influence maximization*

This study

How can we control network formation to improve the outcomes of given processes?
Principles for minimal top-down control of bottom-up organization

We show that

1. Optimal designs are not tractable for most empirical settings
2. Common designs perform poorly even with perfect implementation
3. Using a simple metaheuristic, we can produce designs that perform orders of magnitude better

Follow on work

Tools *w/Marc Santolini, Rathin Jeyaram*

Credible estimation of local network effects

More reliable ATE, LATE

Reconstruction of social networks from unstructured text w/LLMs

Less missing data

Theory and Applications

Designing inducements for network formation

Recruitment in the presence of peer effects

Task allocation in self-managed teams

Thank you!

Funding
Google / Vermont Complex Systems Center: OCEAN Grant
Megan Palmer and Drew Endy, Stanford University
French National Research Agency: AAPG2021

robertnward.github.io