Estimating network-mediated causal effects via spectral embeddings

Alex Hayes

NetSci 2024

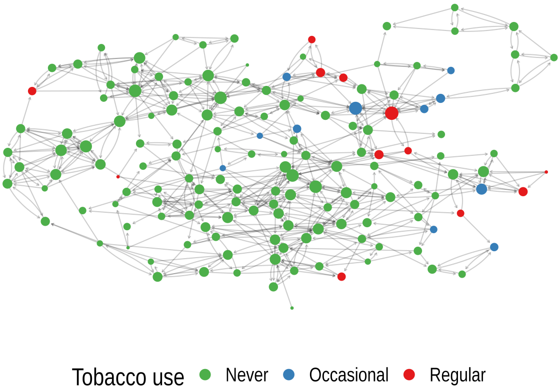
Department of Statistics
University of Wisconsin-Madison

This is joint work

Mark Fredrickson
University of Michigan

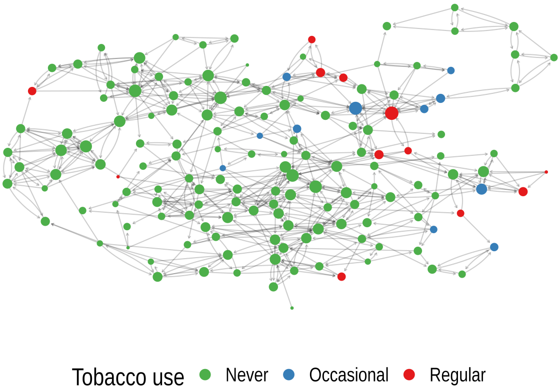
Keith Levin
UW-Madison

Motivation: understand causes of teenage smoking (Michell and West, 1996)



- 129 middle-schoolers in Glasgow
- **Social network** based on self-reported friendships
- **Auxiliary data:** spending money, leisure activities
- **Demographics:** sex, age
- **Behavior:** alcohol, cannabis and tobacco use

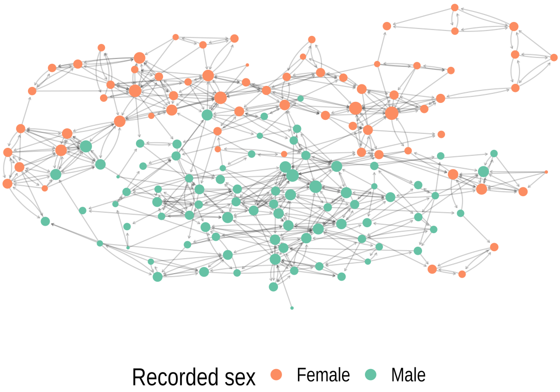
Motivation: understand causes of teenage smoking (Michell and West, 1996)



- 129 middle-schoolers in Glasgow
- **Social network** based on self-reported friendships
- **Auxiliary data:** spending money, leisure activities
- **Demographics:** sex, age
- **Behavior:** alcohol, cannabis and tobacco use

Question: how does sex influence tobacco use?

Smoking is a sexually differentiated behavior and also a social behavior



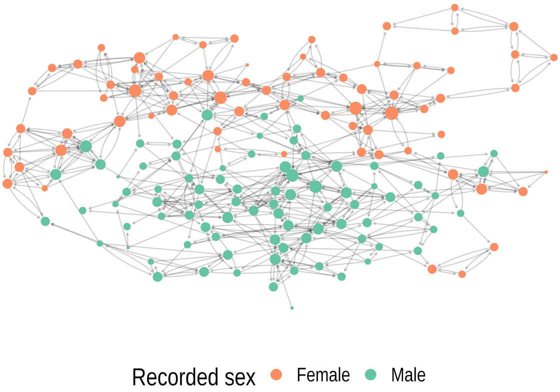
Direct effect of sex on smoking

- Social and cultural expectations could lead to more or less smoking

Indirect effect of sex on smoking

- Students to be friends with other students of same sex…
- …friends induce one another to smoke

Smoking is a sexually differentiated behavior and also a social behavior



Direct effect of sex on smoking

- Social and cultural expectations could lead to more or less smoking

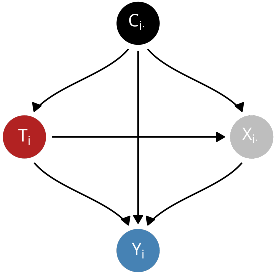
Indirect effect of sex on smoking

- Students to be friends with other students of same sex...
- ...friends induce one another to smoke

Should anti-smoking interventions target the direct or the social mechanism?

Defining causal effects is straightforward when there is no social network

Treatment	T_i	$\in \{0, 1\}$
Outcome	Y_i	$\in \mathbb{R}$
Mediators	$X_{i\cdot}$	$\in \mathbb{R}^{1\times d}$
Confounders	$C_{i\cdot}$	$\in \mathbb{R}^{1\times p}$



Definition (Average treatment effect)

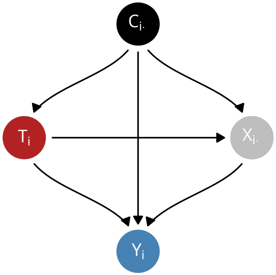
$$\Psi_{\mathrm{ate}} = \mathbb{E}[Y(1) - Y(0)]$$

$Y(t)$ is the counterfactual outcome under treatment value $t \in \{0, 1\}$.

Decomposing causal effects is straightforward when there is no social network

Two distinct mechanisms:

1. Direct effect along $T_i \to Y_i$ path
2. Indirect effect along $T_i \to X_{i\cdot} \to Y_i$ path



Definition (Natural direct and indirect effects)

$$\begin{aligned}\Psi_{\text{ate}} &= \Psi_{\text{nde}} + \Psi_{\text{nie}} \\ &= \mathbb{E}[Y(1, M(0)) - Y(0, M(0))] + \mathbb{E}[Y(1, M(1)) - Y(1, M(0))]\end{aligned}$$

$Y(t, m)$ is the counterfactual outcome under treatment t and mediator m

We want to decompose node-level effects into direct and social components

Network $A \in \mathbb{R}^{n \times n}$

For each node i:

- Treatment $T_i \in \{0, 1\}$
- Outcome $Y_i \in \mathbb{R}$
- Confounders $C_{i\cdot} \in \mathbb{R}^p$

The network A is noisy and complex but often exhibits high levels of homophily

We want to decompose node-level effects into direct and social components

Network $A \in \mathbb{R}^{n \times n}$

For each node i:

- Treatment $T_i \in \{0, 1\}$
- Outcome $Y_i \in \mathbb{R}$
- Confounders $C_{i\cdot} \in \mathbb{R}^p$

The network A is noisy and complex but often exhibits high levels of homophily

Assume each node belongs to a social group $X_{i\cdot}$ and behaviors vary with social group

Stochastic blockmodels are the canonical model for social groups

Nodes colored by social group

$X_{i\cdot} \in \{0,1\}^d$ one-hot indicator of social group

$B \in [0,1]^{d\times d}$ between-group friend probabilities

Friendships depend on social group and B

$$\mathbb{P}\left(A_{ij} = 1 \,\middle|\, X\right) = X_{i\cdot} B X_{j\cdot}^T$$

<u>Social groups X are latent (i.e., unobserved)</u>

We have very rich and general models for latent social groups

Nodes colored by social group

- Degree-correction
- Mixed-membership social groups
- Overlapping social groups
- Group-specific popularity parameters

Definition (Random dot product graph)

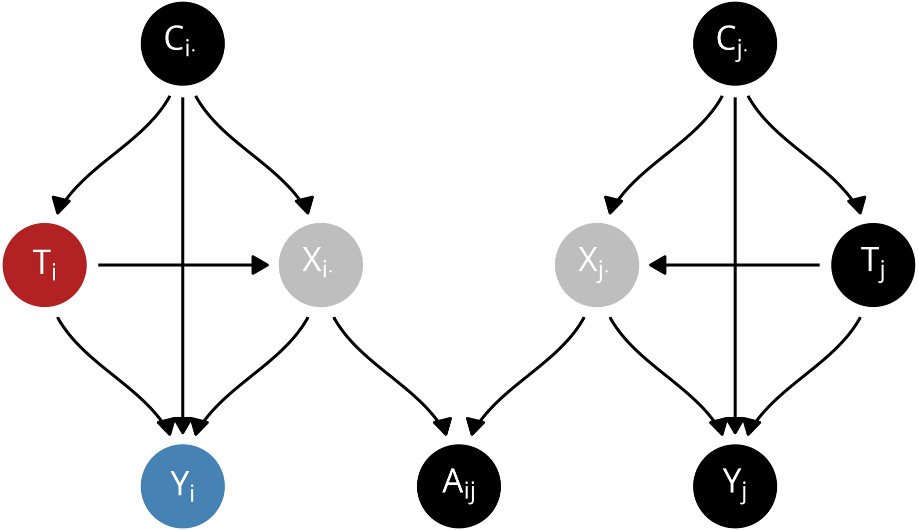
A symmetric, weighted adjacency matrix

$A = XX^T + E$

$X_{1\cdot}, \ldots, X_{n\cdot} \in \mathbb{R}^{1 \times d}$ i.i.d. latent positions

E independent sub-gamma noise

Goal: estimate direct and indirect effects as mediated by latent social groups



We want to use a regression estimator for direct and indirect effects

Assumption

$$\mathbb{E}[Y_i \mid T_i, C_{i\cdot}, X_{i\cdot}] = \beta_0 + T_i \beta_{\mathsf{t}} + C_{i\cdot}\beta_{\mathsf{c}} + X_{i\cdot}\beta_{\mathsf{x}} \quad \text{outcome model}$$

$$\mathbb{E}[X_{i\cdot} \mid T_i, C_{i\cdot}] = \theta_0 + T_i \theta_{\mathsf{t}} + C_{i\cdot}\Theta_{\mathsf{c}} \quad \text{mediator model}$$

Fact: semi-parametric identification of natural mediated effects

Under previous assumption, when natural direct and indirect effect are non-parametrically identified, we have

$$\Psi_{\text{nde}} = \beta_{\mathsf{t}}$$

$$\Psi_{\text{nie}} = \theta_{\mathsf{t}}\, \beta_{\mathsf{x}}$$

Problem: We don't see the latent social groups X, can't fit regressions

Latent positions X can be estimated via principle components analysis

Definition (ASE)

Given a network A, the d-dimensional adjacency spectral embedding of A is

$$\widehat{X} = \widehat{U}\widehat{S}^{1/2}$$

where $\widehat{U}\widehat{S}\widehat{U}^T$ is the rank-d truncated singular value decomposition of A.

Lemma

Under a suitable network model, there is a $d \times d$ orthogonal matrix Q such that

$$\max_{i \in [n]} \left\| \widehat{X}_{i\cdot} - X_{i\cdot} Q \right\| = o_p(1).$$

Estimated latent positions can plug directly into ordinary least squares

Let $\widehat{D} = \begin{bmatrix} 1 & T & C & \widehat{X} \end{bmatrix} \in \mathbb{R}^{n \times (2+p+d)}$ and $L = \begin{bmatrix} 1 & T & C \end{bmatrix} \in \mathbb{R}^{n \times (p+2)}$.

$$\begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_\mathsf{t} \\ \widehat{\beta}_\mathsf{c} \\ \widehat{\beta}_\mathsf{x} \end{bmatrix} = \left(\widehat{D}^T \widehat{D}\right)^{-1} \widehat{D}^T Y \quad \text{and} \quad \widehat{\Theta} = \left(L^T L\right)^{-1} L^T \widehat{X}.$$

$$\widehat{\Psi}_{\text{nde}} = \widehat{\beta}_\mathsf{t} \quad \text{and} \quad \widehat{\Psi}_{\text{nie}} = \widehat{\theta}_\mathsf{t} \, \widehat{\beta}_\mathsf{x}$$

Ordinary least squares regression estimates are asymptotically normal

Theorem

Under a suitably well-behaved network model and some moments conditions on regression errors, there is an unknown orthogonal matrix Q such that

$$\sqrt{n}\,\widehat{\Sigma}_{\beta}^{-1/2} \begin{pmatrix} \widehat{\beta}_w - \beta_w \\ Q\,\widehat{\beta}_x - \beta_x \end{pmatrix} \to \mathcal{N}(0, I_d), \textit{and}$$

$$\sqrt{n}\,\widehat{\Sigma}_{\mathrm{vec}(\Theta)}^{-1/2} \left(\mathrm{vec}\left(\widehat{\Theta}\,Q^T\right) - \mathrm{vec}(\Theta)\right) \to \mathcal{N}\left(0, I_{pd}\right).$$

where $\widehat{\Sigma}_{\mathrm{vec}(\Theta)}^{-1/2}$ *and* $\widehat{\Sigma}_{\beta}^{-1/2}$ *are the typical heteroscedastic robust covariance estimators, with* $\widehat{X}$ *plugged in for* X*.*

Note: must correctly specify dimension of latent positions d

Aside: our regression results are substantially more general than similar work

- Network can be weighted, rather than binary
- No parametric assumptions on edge noise
- No parametric assumptions on regression errors
- Regression errors can be heteroscedastic
- Can model latent positions as outcomes

Principle components + ordinary least squares is a very general, distributionally agnostic tool for network regression

<u>Folklore</u>: similar results hold for asymmetric A, rectangular A, bipartite A, graph Laplacian embeddings rather than adjacency matrix embeddings, general regression M-estimators other than ordinary least squares

Causal estimators are asymptotically normal

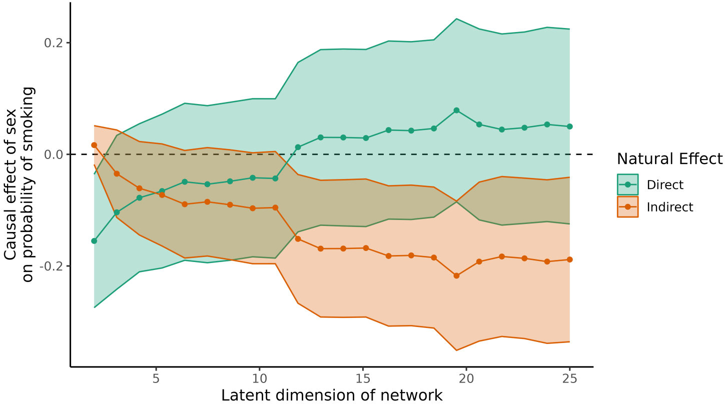
Theorem

Under the same statistical assumptions as before, plus counterfactual assumptions required for causal identification,

$$\sqrt{n\,\widehat{\sigma}^2_{\mathrm{nde}}}\left(\widehat{\Psi}_{\mathrm{nde}} - \Psi_{\mathrm{nde}}\right) \to \mathcal{N}(0,1), \text{ and}$$
$$\sqrt{n\,\widehat{\sigma}^2_{\mathrm{nie}}}\left(\widehat{\Psi}_{\mathrm{nie}} - \Psi_{\mathrm{nie}}\right) \to \mathcal{N}(0,1).$$

where $\widehat{\sigma}^2_{\mathrm{nde}}$ and $\widehat{\sigma}^2_{\mathrm{nie}}$ are rather unfriendly variance estimators derived via the delta method and the previous theorem.

Application to Glasgow data



Estimated direct and indirect effects of sex on tobacco usage in the Glasgow social network, adjusted for age and church attendance. Positive values indicate a greater propensity for adolescent boys to smoke, negative effects a greater propensity for adolescent girls to smoke.

Pre-print

Hayes, Alex, Mark M. Fredrickson, and Keith Levin. "Estimating Network-Mediated Causal Effects via Spectral Embeddings." arXiv, April 14, 2023. https://arxiv.org/abs/2212.12041.

Takeaways:

- Latent network structure as a mediator
- Network regression with causal applications
- Consistency and asymptotic normality of ordinary least squares

Future work:

- Varimax rotation for interpretability
- Extension to include peer effects

Contact info, slides & link to manuscript: https://www.alexpghayes.com

Peer effects

- Identifying peer effects in these models is very nuanced
  - Not one but two forthcoming manuscripts about this
- When peer effects are identified, ordinary least squares sometimes works
- When peer effects are unidentified, typically due to aliasing with β_{x} or β_0
  - Aliasing might occur only in the asymptotic limit

$$\mathbb{E}[Y_i \mid T, C_{i\cdot}, X_{i\cdot}, Y_{-i}, A] = \beta_0 + T_i \beta_{\mathrm{t}} + C_{i\cdot}\beta_{\mathrm{c}} + X_{i\cdot}\beta_{\mathrm{x}} + \underbrace{\beta_{\mathrm{gt}} \sum_{j \neq i} \frac{A_{ij} T_j}{\sum_i A_{ij}}}_{\text{interference}} + \underbrace{\beta_{\mathrm{gy}} \sum_{j \neq i} \frac{A_{ij} Y_j}{\sum_i A_{ij}}}_{\text{“contagion”}}$$

Lots of research attention focused on using this kind of model for causal inference

Identifying assumptions

The random variables $(Y_i, Y_i(t,x), X_{i\cdot}, X_{i\cdot}(t), C_{i\cdot}, T_i)$ are independent over $i \in [n]$ and obey the following three properties.

1. Consistency:

   if $T_i = t$, then $X_{i\cdot}(t) = X_{i\cdot}$ with probability 1, and

   if $T_i = t$ and $X_{i\cdot} = x$, then $Y_i(t,x) = Y_i$ with probability 1

2. Sequential ignorability:

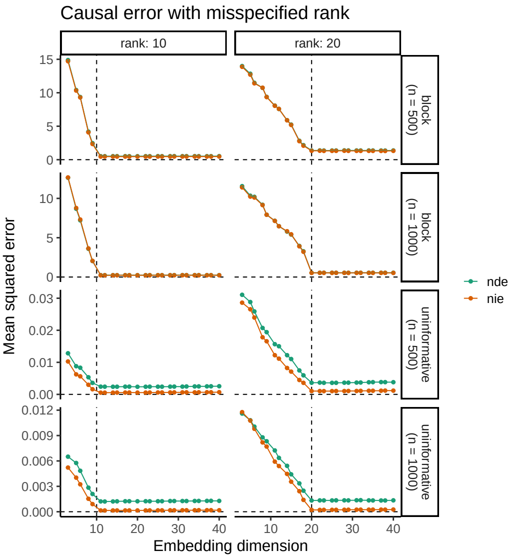
$$\{Y_i(t^*,x), X_{i\cdot}(t)\} \perp\!\!\!\perp T_i \mid C_{i\cdot} \quad \text{and} \quad \{Y_i(t^*,x)\} \perp\!\!\!\perp X_{i\cdot} \mid T_i = t, C_{i\cdot}$$

3. Positivity:

$$\mathbb{P}(x \mid T_i, C_{i\cdot}) > 0 \text{ for each } x \in \mathrm{supp}(X_{i\cdot})$$

$$\mathbb{P}(t \mid C_{i\cdot}) > 0 \text{ for each } t \in \mathrm{supp}(T_i)$$

Choosing $\widehat{d}$: overestimating the embedding dimension is fine



Interventions allowed

Provided that controls $C_{i\cdot}$ are sufficiently informative about group membership $X_{i\cdot}$, treatment T_i is allowed to cause:

- Changes in popularity within a group
- Movement to a new friend group
- Becoming a member of a new friend group while remaining in current friend group
- Friendships becoming more or less likely between distinct friend groups
- Combinations of the above

See Appendix of manuscript for details.

References

Michell, L. and P. West (1996). Peer pressure to smoke: The meaning depends on the method. Health Education Research 11(1), 39–49.