Linear regression for causal inference on social networks via network embeddings

Alex Hayes

2024-06-17

University of Wisconsin-Madison

Christakis and Fowler 2007 catalyzed methodology for social networks

The Spread of Obesity in a Large Social Network Over 32 Years

Nicholas A. Christakis, M.D., Ph.D., M.P.H., and James H. Fowler, Ph.D.

ABSTRACT

BACKGROUND

The prevalence of obesity has increased substantially over the past 30 years. We performed a quantitative analysis of the nature and extent of the person-to-person spread of obesity as a possible factor contributing to the obesity epidemic.

METHODS

We evaluated a densely interconnected social network of 12,067 people assessed repeatedly from 1971 to 2003 as part of the Framingham Heart Study. The body-mass index was available for all subjects. We used longitudinal statistical models to examine whether weight gain in one person was associated with weight gain in his

Shalizi and Thomas 2011 argued the issue was failure to account for homophily

Sociological Methods & Research
40(2) 211–239

DOI: 10.1177/0049124111404820
http://smr.sagepub.com
SAGE

Homophily and Contagion Are Generically Confounded in Observational Social Network Studies

Cosma Rohilla Shalizi[1] and
Andrew C. Thomas[1]

There is now an explosion of approaches that account for homophily

Estimating Causal Peer Influence in Homophilous Social Networks by Inferring Latent Locations

Edward McFowland III[a] and Cosma Rohilla Shalizi[b]

[a]Department of Information and Decision Sciences, Carlson School of Management, University of Minnesota, Minneapolis, MN; [b]Statistics Department, Carnegie Mellon University, and the Santa Fe Institute, Pittsburgh, PA

ABSTRACT

Social influence cannot be identified from purely observational data on social networks, because such influence is generically confounded with latent homophily, that is, with a node's network partners being informative about the node's attributes and therefore its behavior. If the network grows according to either a latent community (stochastic block) model, or a continuous latent space model, then latent homophilous attributes can be consistently estimated from the global pattern of social ties. We show that, for common versions of those two network models, these estimates are so informative that controlling for estimated attributes allows for asymptotically unbiased and consistent estimation of social-influence effects in linear models. In particular, the bias shrinks at a rate that directly reflects how much information the network provides about the latent attributes. These are the first results on the consistent nonexperimental estimation of social-influence effects in the presence of latent homophily, and we discuss the prospects for generalizing them.

ARTICLE HISTORY
Received February 2018
Accepted June 2021

KEYWORDS
Causal Inference; Homophily; Social Networks; Peer Influence

Homophily is the tendency for behavior to vary with social group

Nodes colored by social group

Nodes	(people)	i	$\in \{1, \dots, n\}$
Edge $i \sim j$	(friends?)	A_{ij}	$\in \{0, 1\}$
Outcome		Y_i	$\in \mathbb{R}$
Treatment		T_i	$\in \mathbb{R}$
Latent position	(social group)	$X_{i\cdot}$	$\in \mathbb{R}^d$

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

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

Social groups X are latent (i.e., unobserved)

There are many extensions to the stochastic blockmodel

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

A natural approach to model homophily is to include X in regression models

$$\underbrace{\mathbb{E}[Y_i \mid T_i, C_{i\cdot}, X_{i\cdot}]}_{\mathbb{R}} = \beta_0 + T_i \beta_{\mathrm{t}} + C_{i\cdot} \beta_{\mathrm{c}} + X_{i\cdot} \beta_{\mathrm{x}} \qquad \text{outcome}$$

$$\underbrace{\mathbb{E}[X_{i\cdot} \mid T_i, C_{i\cdot}]}_{\mathbb{R}^{1 \times d}} = \theta_0 + T_i \theta_{\mathrm{t}} + C_{i\cdot} \Theta_{\mathrm{c}} \qquad \text{latent positions}$$

β_{x} is “the effect of belonging to a particular social group”

Edge $i \sim j$	A_{ij}	$\in \mathbb{R}$
Latent position	$X_{i\cdot}$	$\in \mathbb{R}^{1 \times d}$
Treatment	T_i	$\in \{0, 1\}$
Outcome	Y_i	$\in \mathbb{R}$
Confounders	$C_{i\cdot}$	$\in \mathbb{R}^{1 \times p}$

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

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

Ordinary least squares 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

$$\sqrt{n}\,\widehat{\Sigma}_{\beta}^{-1/2}\begin{pmatrix}\widehat{\beta}_{\mathsf{w}}-\beta_{\mathsf{w}}\\ Q\,\widehat{\beta}_{\mathsf{x}}-\beta_{\mathsf{x}}\end{pmatrix}\to\mathcal{N}(0,I_d),\ \textit{and}$$

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

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

Our 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

Folklore: 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

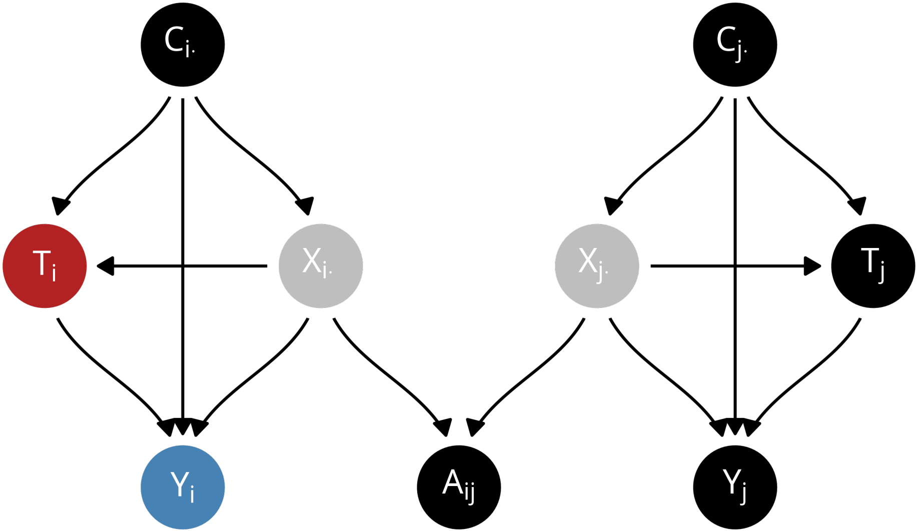
Peer effects

- 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

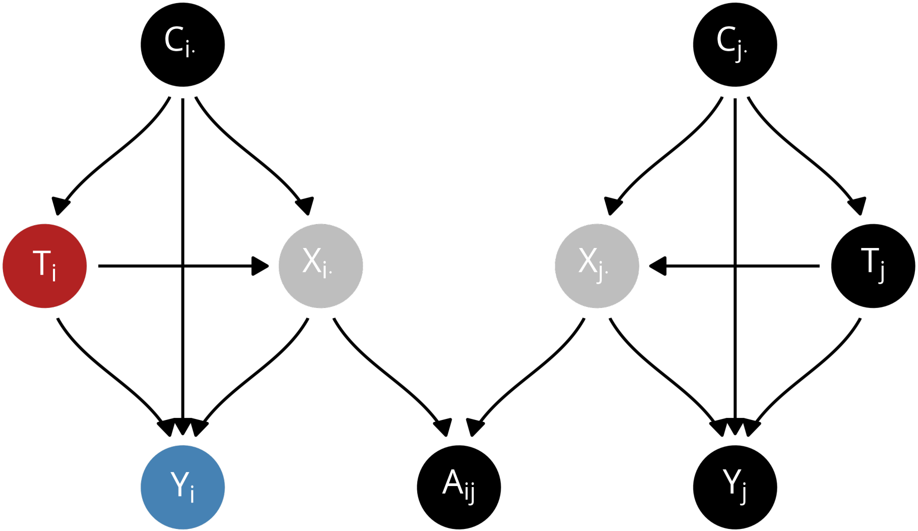
$$\mathbb{E}[Y_i \mid T, C_{i\cdot}, X_{i\cdot}, Y_{-i}, A] = \beta_0 + T_i \beta_{\mathsf{t}} + C_{i\cdot}\beta_{\mathsf{c}} + X_{i\cdot}\beta_{\mathsf{x}} + \underbrace{\beta_{\mathsf{gt}} \sum_{j \neq i} \frac{A_{ij} T_j}{\sum_i A_{ij}}}_{\text{interference}} + \underbrace{\beta_{\mathsf{gy}} \sum_{j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

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

People typically use embeddings to fairly blindly control for latent confounding



What happens if treatment effects latent position rather than vice versa?

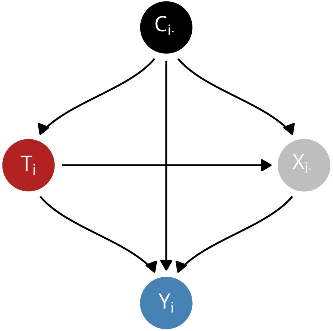


Causal mediation

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

Decompose effect of T_i on Y_i:

1. Effect operating along $T_i \to Y_i$ path (direct)
2. Effect operating along $T_i \to X_{i\cdot} \to Y_i$ path (indirect)



$$\Psi_{\text{ate}} = \Psi_{\text{nde}} + \Psi_{\text{nie}}$$
$$\Psi_{\text{nde}} = \mathbb{E}[Y_i(t, X_{i\cdot}(t^*)) - Y_i(t^*, X_{i\cdot}(t^*))]$$
$$\Psi_{\text{nie}} = \mathbb{E}[Y_i(t, X_{i\cdot}(t)) - Y_i(t, X_{i\cdot}(t^*))]$$

When latent positions are confounders:

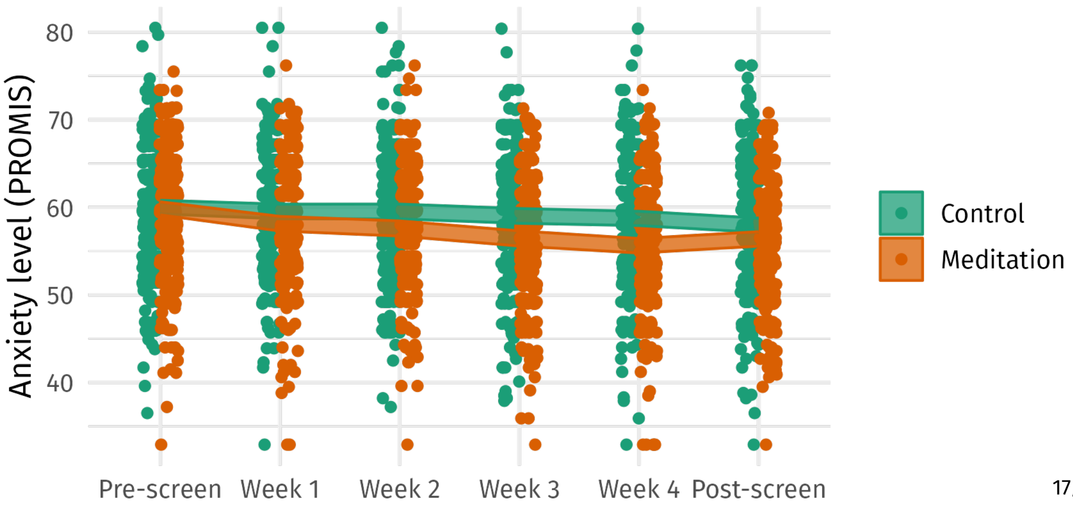
$$\mathbb{E}[Y_i \mid T_i, C_{i\cdot}, X_{i\cdot}] = \beta_0 + T_i \underbrace{\beta_{\mathrm{t}}}_{\substack{\text{average}\\\text{treatment}\\\text{effect}}} + C_{i\cdot}\beta_{\mathrm{c}} + X_{i\cdot}\beta_{\mathrm{x}}$$

When latent positions are mediators:

$$\mathbb{E}[Y_i \mid T_i, C_{i\cdot}, X_{i\cdot}] = \beta_0 + T_i \underbrace{\beta_{\mathrm{t}}}_{\substack{\text{natural}\\\text{direct}\\\text{effect}}} + C_{i\cdot}\beta_{\mathrm{c}} + X_{i\cdot} \underbrace{\beta_{\mathrm{x}}}_{\substack{\text{effect of}\\ X \text{ on } Y\text{ (?)}}}$$

$$\mathbb{E}[X_{i\cdot} \mid T_i, C_{i\cdot}] = \theta_0 + T_i \underbrace{\theta_{\mathrm{t}}}_{\substack{\text{effect of}\\ T \text{ on } X}} + C_{i\cdot}\Theta_{\mathrm{c}}$$

Experiments (n = 662) show that smartphone-guided meditation reduces anxiety



We study the causal effect of four weeks of meditation

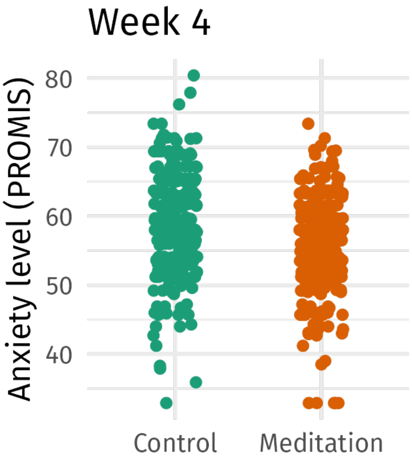
Counterfactual outcomes

anxiety, no meditation	$Y_i(0)$	$\in \mathbb{R}$
anxiety, meditation	$Y_i(1)$	$\in \mathbb{R}$

Observed data

Treatment	meditation	T_i	$\in \{0,1\}$
Outcome	anxiety	Y_i	$\in \mathbb{R}$

$$\Psi_{\text{ate}} = \mathbb{E}[Y_i(1) - Y_i(0)]$$
$$= \mathbb{E}[Y_i \mid T_i = 1] - \mathbb{E}[Y_i \mid T_i = 0]$$
$$\widehat{\Psi}_{\text{ate}} = -2.7 \pm 1.3$$



Psychologists want to know why the Healthy Minds Program reduces distress

The meditation program is designed to alter latent cognitive factors

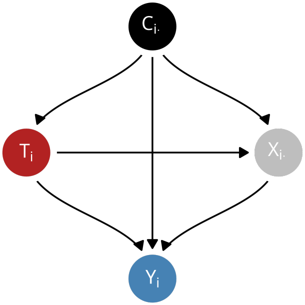
1. Awareness (mindful action)
2. Connection (social connection, reducing loneliness)
3. Insight (cognitive defusion)
4. Purpose (presence of meaning)

The hope: improving these latent cognitive factors reduces anxiety

Multi-stage effects can be formalized as causal mediation

Decompose effect of T_i on Y_i:

1. Effect operating along $T_i \rightarrow Y_i$ path (direct)
   - i.e., guided breathing reduces anxiety
2. Effect operating along $T_i \rightarrow X_{i\cdot} \rightarrow Y_i$ path (indirect)
   - i.e., meditation program decreases loneliness, which in turn decreases anxiety



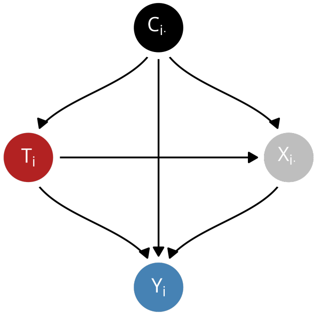
$$\Psi_{\text{ate}} = \Psi_{\text{nde}} + \Psi_{\text{nie}}$$

Natural direct and indirect effects are defined counterfactually

Treatment	meditation	T_i	$\in \{0, 1\}$
Outcome	anxiety	Y_i	$\in \mathbb{R}$
Mediators	cognitive state	$X_{i\cdot}$	$\in \mathbb{R}^{1\times d}$
Confounders	age & sex	$C_{i\cdot}$	$\in \mathbb{R}^{1\times p}$

$$\Psi_{\text{nde}} = \mathbb{E}[Y_i(t, X_{i\cdot}(t^*)) - Y_i(t^*, X_{i\cdot}(t^*))]$$
$$\Psi_{\text{nie}} = \mathbb{E}[Y_i(t, X_{i\cdot}(t)) - Y_i(t, X_{i\cdot}(t^*))]$$

If we knew cognitive state, we could use standard tools. But cognitive state is latent!



Psychologists measure latent cognitive state using surveys

For example, the NIH Toolbox loneliness survey

1. I feel alone and apart from others
2. I feel left out
3. I feel that I am no longer close to anyone
4. I feel alone
5. I feel lonely

Never	Rarely	Sometimes	Usually	Always
1 ○	2 ○	3 ○	4 ○	5 ○

The experimenters ran weekly surveys of the study participants

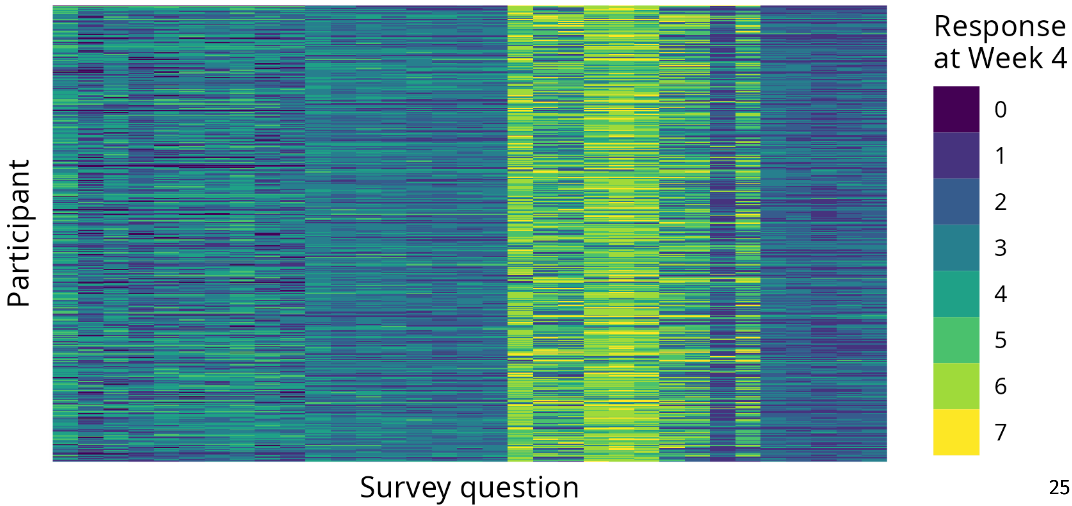
- NIH Toolbox Loneliness (5 questions)
- Five Facet Mindfulness Questionnaire Acting with Awareness subscale (8 questions)
- Drexel Defusion Scale (10 questions)
- Meaning in Life Questionnaire (10 questions)

Another survey: the Meaning in Life Questionnaire

1. I understand my life's meaning.
2. I am looking for something that makes my life feel meaningful.
3. I am always looking to find my life's purpose.
4. My life has a clear sense of purpose.
5. I have a good sense of what makes my life meaningful.
6. I have discovered a satisfying life purpose.
7. I am always searching for something that makes my life feel significant.
8. I am seeking a purpose or mission for my life
9. My life has no clear purpose.
10. I am searching for meaning in my life.

Absolutely untrue	Mostly untrue	Somewhat untrue	Can't say true or false	Somewhat true	Mostly true	Absolutely true
1 ○	2 ○	3 ○	4 ○	5 ○	6 ○	7 ○

The survey responses form a bipartite network



We model survey responses using a factor model

Suppose $A \in \mathbb{R}^{n \times m}$ is the matrix of survey responses. Then

$$\mathbb{E}[A \mid X, Z] = XZ^T$$

$$\begin{array}{lll} \text{Participant embeddings} & X & \in \mathbb{R}^{n \times d} \\ \text{Question embeddings} & Z & \in \mathbb{R}^{m \times d} \end{array}$$

We estimate latent cognitive state by embedding the network

Definition (ASE)

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

$$\widehat{X} = \widehat{U}\widehat{S}^{1/2} \text{ and } \widehat{Z} = \widehat{V}\widehat{S}^{1/2}$$

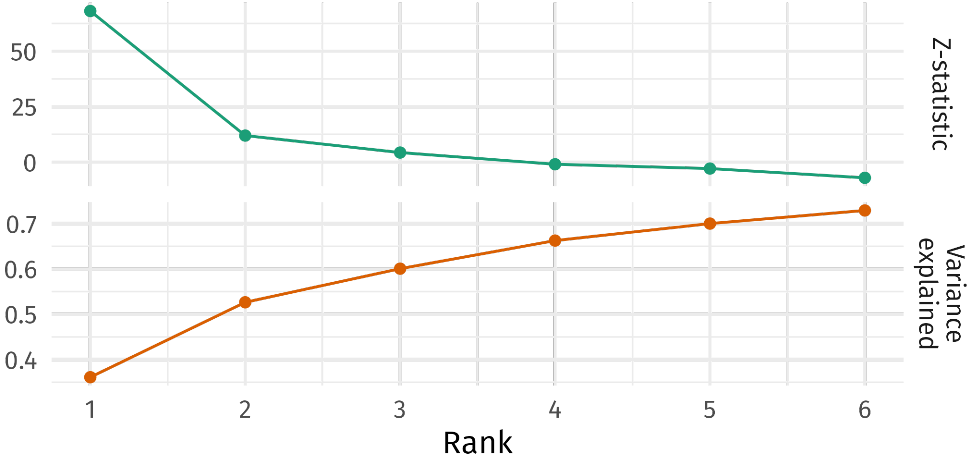
where $\widehat{U}\widehat{S}\widehat{V}^T$ is the rank-d truncated singular value decomposition of A.

Lemma

Under a suitable low-rank model, there are $d \times d$ orthogonal matrices Q_1, Q_2 such that

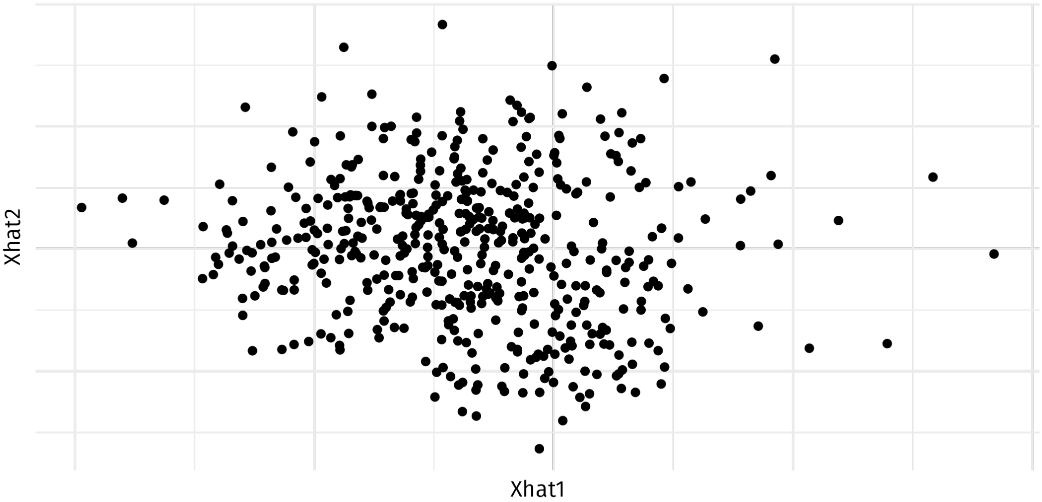
$$\max_{i \in [n]} \left\| \widehat{X}_{i\cdot} - X_{i\cdot} Q_1 \right\| = o_p(1) \text{ and } \max_{i \in [n]} \left\| \widehat{V}_{i\cdot} - V_{i\cdot} Q_2 \right\| = o_p(1).$$

We must estimate the number of latent factors *d*

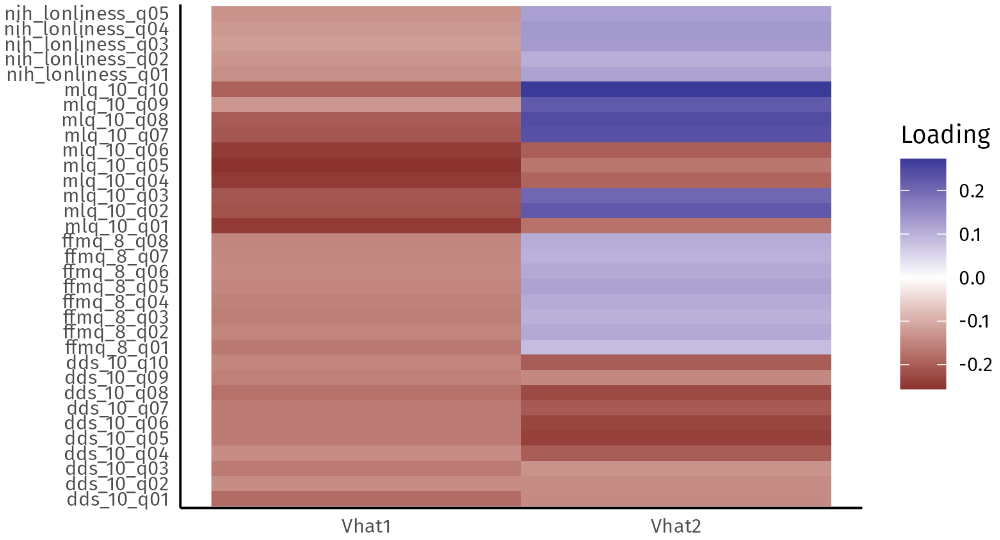


Cross-validated eigenvalue method selects d = 2

Participants embed into an inocuous latent space



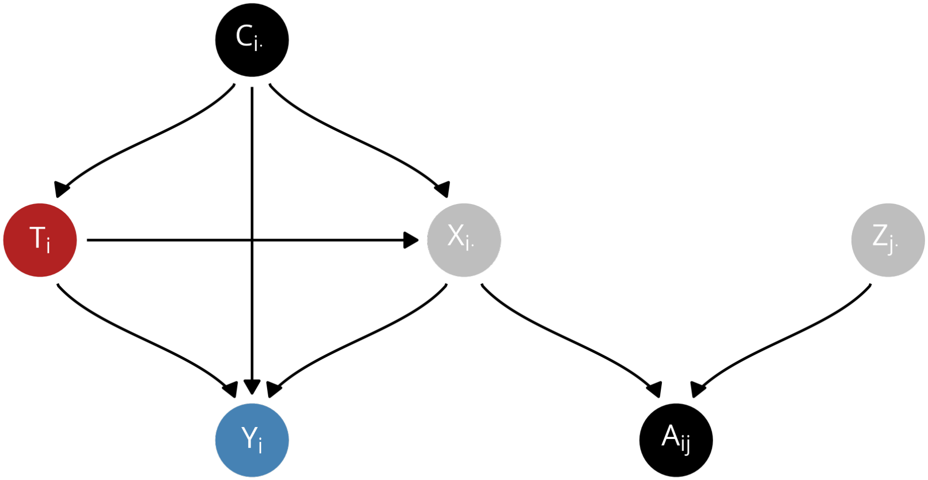
The survey questions are largely redundant



Recap so far

- Psychologists hoped to find four latent factors corresponding to loneliness, meaning in life, etc, etc
- We only found two ”do you feel good”, and ”do you feel bad”
- People basically answer all the questions exactly the same
- Still curious how much the latent well-being factors account for decrease in anxiety

We use survey responses to estimate latent participant embeddings



The identifying assumptions can be expressed counterfactually

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:

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

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

Under semi-parametric assumptions, causal effects are regression coefficients

If the non-parametric identification conditions hold and also

$$\underbrace{\mathbb{E}[Y_i \mid T_i, C_{i\cdot}, X_{i\cdot}]}_{\mathbb{R}} = \underbrace{\beta_0}_{\mathbb{R}} + \underbrace{T_i}_{\{0,1\}} \underbrace{\beta_\mathsf{t}}_{\mathbb{R}} + \underbrace{C_{i\cdot}}_{\mathbb{R}^{1\times p}} \underbrace{\beta_\mathsf{c}}_{\mathbb{R}^p} + \underbrace{X_{i\cdot}}_{\mathbb{R}^{1\times d}} \underbrace{\beta_\mathsf{x}}_{\mathbb{R}^d},$$

$$\underbrace{\mathbb{E}[X_{i\cdot} \mid T_i, C_{i\cdot}]}_{\mathbb{R}^{1\times d}} = \underbrace{\theta_0}_{\mathbb{R}^{1\times d}} + \underbrace{T_i}_{\{0,1\}} \underbrace{\theta_\mathsf{t}}_{\mathbb{R}^{1\times d}} + \underbrace{C_{i\cdot}}_{\mathbb{R}^{1\times p}} \underbrace{\Theta_\mathsf{c}}_{\mathbb{R}^{p\times d}}$$

Then

$$\Psi_{\mathrm{nde}}(t, t^*) = (t - t^*)\, \beta_\mathsf{t}$$

$$\Psi_{\mathrm{nie}}(t, t^*) = (t - t^*)\, \theta_\mathsf{t}\, \beta_\mathsf{x}$$

We using estimated $\widehat{X}$ in place of unobserved X in regression estimators

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

$$\begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_t \\ \widehat{\beta}_c \\ \widehat{\beta}_x \end{bmatrix} = \left(\widehat{D}^T \widehat{D}\right)^{-1} \widehat{D}^T Y \quad \text{and} \quad \widehat{\Theta} = \left(L^T L\right)^{-1} L^T \widehat{X}.$$

$$\widehat{\Psi}_{\mathrm{nde}} = (t - t^*)\, \widehat{\beta}_t$$
$$\widehat{\Psi}_{\mathrm{nie}} = (t - t^*)\, \widehat{\theta}_t\, \widehat{\beta}_x$$

Regression coefficients are asymptotically normal

Theorem

Under a suitably well-behaved network model and some moment 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 heteroscedasticity robust covariance estimators, with* $\widehat{X}$ *plugged in for X.*

Causal estimators are asymptotically normal

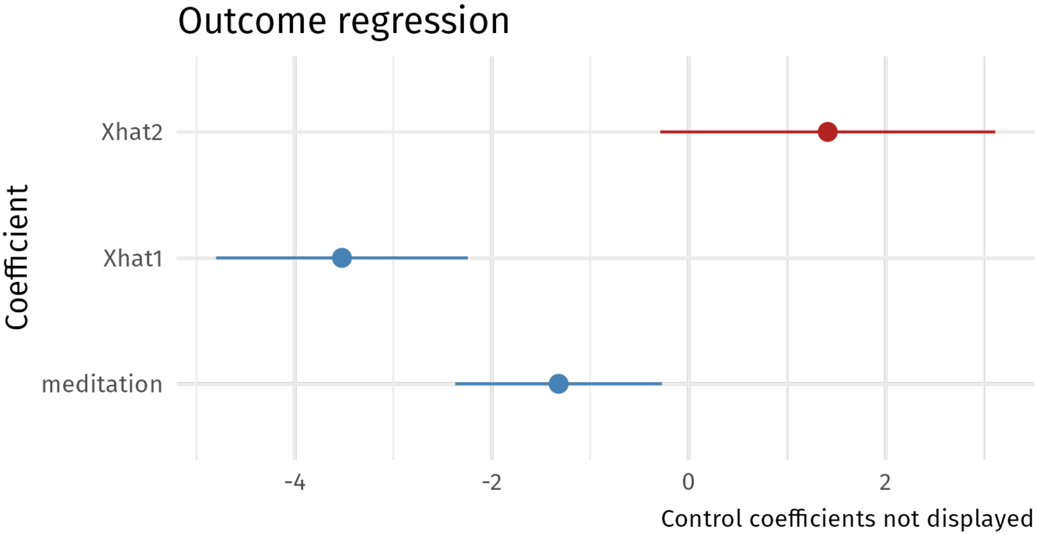
Theorem

Under the same statistical assumptions as before, plus mediating homophily,

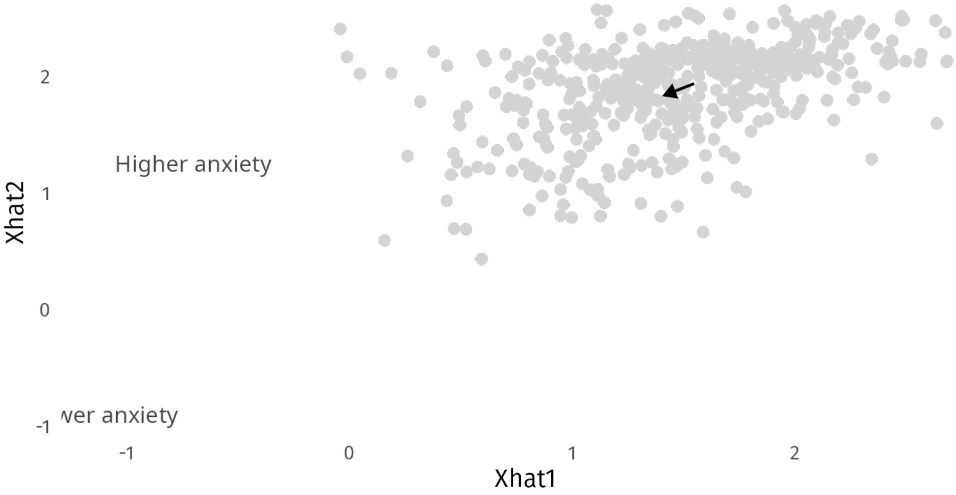
$$\sqrt{n\,\widehat{\sigma}^2_{\mathrm{nde}}}\left(\widehat{\Psi}_{\mathrm{nde}} - \Psi_{\mathrm{nde}}\right) \to \mathcal{N}(0,1), \text{ and}$$
$$\sqrt{n\,\widehat{\sigma}^2_{\mathrm{nie}}}\left(\widehat{\Psi}_{\mathrm{nie}} - \Psi_{\mathrm{nie}}\right) \to \mathcal{N}(0,1).$$

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

Latent factors increase anxiety, meditation decreases anxiety



Meditation causes a small but significant shift in latent space

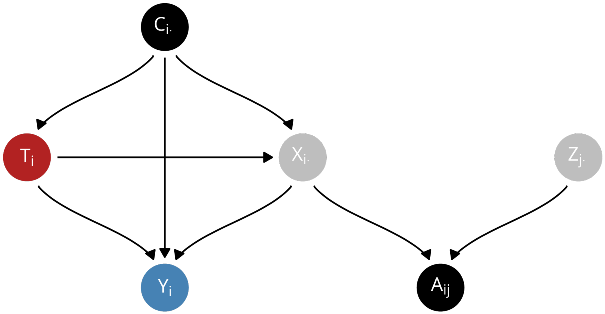


We estimate that most of the effect is along the direct pathway

$$\widehat{\Psi}_{\text{ate}} = -2.7 \pm 1.3$$

$$\widehat{\Psi}_{\text{nde}} = -1.8 \pm 1.1$$

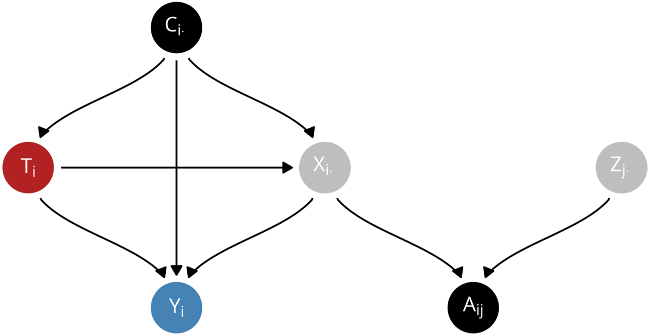
$$\widehat{\Psi}_{\text{nie}} = -1 \pm 0.7$$



Takeaways

- We developed a method to decompose causal effects into effects operating along direct and indirect pathways in a low-rank latent space
- ”Latent causal effects” sounds scary, but can often just look at your data to interpret them

$$\Psi_{\text{ate}} = \Psi_{\text{nde}} + \Psi_{\text{nie}}$$
$$\widehat{\Psi}_{\text{ate}} = -2.7 \pm 1.3$$
$$\widehat{\Psi}_{\text{nde}} = -1.8 \pm 1.1$$
$$\widehat{\Psi}_{\text{nie}} = -1 \pm 0.7$$



Thank you! Questions?

Read the manuscript at *https://arxiv.org/abs/2212.12041*

R package latentnetmediate

Stay in touch

- @alexpghayes
- alex.hayes@wisc.edu
- *https://www.alexpghayes.com*
- *https://github.com/alexpghayes*

I'm looking for a post-doc starting Fall 2024, say hi if this work interests you!

Semi-parametric network model

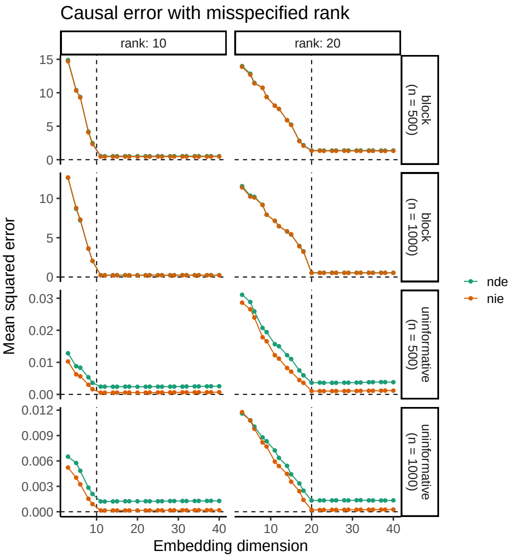
Definition

Let $A \in \mathbb{R}^{n \times n}$ be a random symmetric matrix, such as the adjacency matrix of an undirected graph. Let $P = \mathbb{E}[A \mid X] = XX^T$ be the expectation of A conditional on $X \in \mathbb{R}^{n \times d}$, which has independent and identically distributed rows $X_{1\cdot}, \ldots, X_{n\cdot}$. That is, P has $\mathrm{rank}(P) = d$ and is positive semi-definite with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d > 0 = \lambda_{d+1} = \cdots = \lambda_n$. Conditional on X, the upper-triangular elements of $A - P$ are independent (ν_n, b_n)-sub-gamma random variables.

Remark

$P = XX^T = (XQ)(XQ)^T$ for any $d \times d$ orthogonal matrix Q, the latent positions X are only identifiable up to an orthogonal transformation.

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



mindful action subscale

1. When I do things, my mind wanders off and I'm easily distracted.
2. I don't pay attention to what I'm doing because I'm daydreaming, worrying, or otherwise distracted.
3. I am easily distracted.
4. I find it difficult to stay focused on what's happening in the present.
5. It seems I am 'running on automatic' without much awareness of what I'm doing.
6. I rush through activities without being really attentive to them.
7. I do jobs or tasks automatically without being aware of what I'm doing.
8. I find myself doing things without paying attention.

Never or very rarely true	Rarely true	Sometimes true	Often true	Very often or always true
1 ○	2 ○	3 ○	4 ○	5 ○

Drexel Defusion Scale

1. Feelings of anger. You become angry when someone takes your place in a long line. To what extent would you normally be able to defuse from feelings of anger?
2. Cravings for food. You see your favorite food and have the urge to eat it. To what extent would you normally be able to defuse from cravings for food?
3. Physical pain. Imagine that you bang your knee on a table leg. To what extent would you normally be able to defuse from physical pain?
4. Anxious thoughts. Things have not been going well at school or your job, and work just keeps piling up. To what extent would you normally be able to defuse from anxious thoughts like ”I’ll never get this done.”?
5. Thoughts of self. Imagine you are having a thought such as ”no one likes me.” To what extent would you normally be able to defuse from negative thoughts about yourself?
6. Thoughts of hopelessness. You are feeling sad and stuck in a difficult

Network model

Definition

Let $A \in \mathbb{R}^{n \times n}$ be a random symmetric matrix, such as the adjacency matrix of an undirected graph. Let $P = \mathbb{E}[A \mid X] = XX^T$ be the expectation of A conditional on $X \in \mathbb{R}^{n \times d}$, which has independent and identically distributed rows $X_{1\cdot}, \dots, X_{n\cdot}$. That is, P has $\operatorname{rank}(P) = d$ and is positive semi-definite with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d > 0 = \lambda_{d+1} = \cdots = \lambda_n$. Conditional on X, the upper-triangular elements of $A - P$ are independent (ν_n, b_n)-sub-gamma random variables.

Remark

$P = XX^T = (XQ)(XQ)^T$ for any $d \times d$ orthogonal matrix Q, the latent positions X are only identifiable up to an orthogonal transformation.

References

Christakis, N. A. and J. H. Fowler (2007, July). The Spread of Obesity in a Large Social Network over 32 Years. New England Journal of Medicine 357(4), 370–379.

Shalizi, C. R. and A. C. Thomas (2011, May). Homophily and Contagion Are Generically Confounded in Observational Social Network Studies. Sociological Methods & Research 40(2), 211–239.