Diffusion and Peer Influence
Diffusion & Peer Influence

“you are who you associate with”
1. Diffusion
   A. Compartamental Models
   B. Network Diffusion
      i. Topology
      ii. Timing
      iii. Structural Transmission
          a. Complex contagion

2. Peer Influence
Network Diffusion & Peer Influence

Basics

Coleman, Katz and Menzel, “Diffusion of an innovation among physicians” Sociometry (1957)

Our substantive interest in networks is often in how things move through them, from disease to ideas to behavior.

Fig. 2. Cumulative proportion of doctors introducing gamma/nm: differences in integration on friendship criterion.
Network Diffusion & Peer Influence

Basics

Figure 2. Binge Drinking Predicted Probabilities by Gender and Friends’ Prior Drinking.

Romantic partnerships in high school lead to adoption of partners’ friends’ behaviors.
Depressed students form ties through non-normative network processes.

**Table 4. Proportion of Ties Created or Maintained Over Time by Network Process and Depression Level.**

<table>
<thead>
<tr>
<th>Depression Level</th>
<th>0 to .4</th>
<th>.41 to .8</th>
<th>.81 to 1.2</th>
<th>1.2+</th>
<th>(b^a)</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Students</td>
<td>598</td>
<td>734</td>
<td>337</td>
<td>151</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Number of Ties</td>
<td>4.45</td>
<td>3.67</td>
<td>3.36</td>
<td>2.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network Processes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reciprocity</td>
<td>.37</td>
<td>.35</td>
<td>.33</td>
<td>.37</td>
<td>-.002</td>
<td>.658</td>
</tr>
<tr>
<td>Transitivity</td>
<td>.20</td>
<td>.17</td>
<td>.18</td>
<td>.15</td>
<td>-.014</td>
<td>.094</td>
</tr>
<tr>
<td>Popularity</td>
<td>.41</td>
<td>.44</td>
<td>.42</td>
<td>.38</td>
<td>-.012</td>
<td>.070</td>
</tr>
<tr>
<td>Unknown</td>
<td>.24</td>
<td>.25</td>
<td>.27</td>
<td>.31</td>
<td>.023</td>
<td>.042</td>
</tr>
<tr>
<td>Homophilous Ties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unattributable to Network Processes</td>
<td>.23</td>
<td>.26</td>
<td>.32</td>
<td>.36</td>
<td>.045</td>
<td>.045</td>
</tr>
</tbody>
</table>

*Note: Columns for network processes do not sum to one because multiple processes may contribute to the presence of a tie.*  
*\(^a\)We obtained beta coefficients (\(b\)) by regressing the proportion of ties attributable to each process on depression level (coded 1 to 4).*
Network Diffusion & Peer Influence

*Basics*

Classic (disease) diffusion makes use of compartmental models. Large N and homogenous mixing allows one to express spread as generalized probability models. Works very well for highly infectious bits in large populations…

---

![SI(S) model](https://example.com/sis-model.png)

**SI(S) model** – actors are in only two states, susceptible or infectious.

\[
\Delta I = \beta \left( \frac{S}{P} \right) I - \gamma I - \mu I
\]

\[
\Delta S = -\beta \left( \frac{S}{P} \right) I + \gamma I + \mu I
\]

---

![SIIR(S) model](https://example.com/siir-model.png)

**SIIR(S) model** – adds an “exposed” but not infectious state and recovered.

\[
\Delta S = -\beta \left( \frac{S}{P} \right) I + \alpha R + \mu (P - S)
\]

\[
\Delta E = \beta \left( \frac{S}{P} \right) I - \epsilon E - \mu E
\]

\[
\Delta I = \epsilon E - \gamma I - \mu I
\]

\[
\Delta R = \gamma I - \alpha R - \mu R
\]

Network Diffusion & Peer Influence

Basics – might even help understand the zombie apocalypse

Figure 1: The basic model

Figure 6: Model Equations for the Quarantine model

http://loe.org/images/content/091023/Zombie%20Publication.pdf
Network Diffusion & Peer Influence

Basics

Network Models
Same basic SI(R,Z,etc) setup, but connectivity is not assumed random, rather it is structured by the network contact pattern.

If $p_{ij}$ is small or the network is very clustered, these two can yield very different diffusion patterns.*

*these conditions do matter. Compartmental models work surprisingly well if the network is large, dense or the bit highly infectiousness…because most networks have a bit of randomness in them. We are focusing on the elements that are unique/different for network as opposed to general diffusion.
Network Diffusion & Peer Influence

Basics

If $0 < p_{ij} < 1$
If $0 < p_{ij} < 1$
Network Diffusion & Peer Influence

*Network diffusion features*

Key Question: What features of a network contribute most to diffusion potential?

*In addition to* the dyadic probability that one actor passes something to another \( (p_{ij}) \), two factors affect flow through a network:

**Topology**
- the shape, or form, of the network
- Example: one actor cannot pass information to another unless they are either directly or indirectly connected

**Time**
- the timing of contact matters
- Example: an actor cannot pass information he has not receive yet

Use simulation tools to explore the relative effects of structural connectivity features

*This is a big conditional! – lots of work on how the dyadic transmission rate may differ across populations.*
Network Diffusion & Peer Influence

Network diffusion features

We need:

(1) reachability
(2) distance
(3) local clustering
(4) multiple routes
(5) star spreaders

- A network has to be connected for a bit to pass over it
- If transmission is uncertain, the longer the distance the lower the likelihood of spread.

Distance and diffusion \( p(\text{transfer}) = p_{ij}^{\text{dist}} \)

Here \( p_{ij} \) of 0.6
Network Diffusion & Peer Influence

*Network diffusion features*

We need:

1. reachability
2. distance
3. local clustering
4. multiple routes
5. star spreaders

- Local clustering turns flow “in” on a potential transmission tree
Network Diffusion & Peer Influence

*Network diffusion features*

We need:

1. reachability
2. distance
3. local clustering
4. multiple routes
5. star spreaders

- The more *alternate routes* one has for transmission, the more likely flow should be.
  - *Operationalize alternate routes with structural cohesion*
Network Diffusion & Peer Influence

Network diffusion features

**Probability of transfer**
by distance and number of non-overlapping paths, assume a constant $p_{ij}$ of 0.6

10 paths

5 paths

2 paths

1 path

Cohesion $\Rightarrow$ Redundancy $\Rightarrow$ Diffusion
Network Diffusion & Peer Influence

Network diffusion features

Structural Cohesion:

A network’s structural cohesion is equal to the minimum number of actors who, if removed from the network, would disconnect it.

Node Connectivity
As number of node-independent paths
Network Diffusion & Peer Influence

Network diffusion features

STD Transmission danger: sex or drugs?

Structural core more realistic than nominal core

Data from “Project 90,” of a high-risk population in Colorado Springs
Network Diffusion & Peer Influence

*Network diffusion features*

We need:

1. reachability
2. distance
3. local clustering
4. multiple routes
5. star spreaders

- Much of the work on “core groups” or “at risk” populations focus on high-degree nodes. The assumption is that high-degree nodes are likely to contact lots of people.
Much of the work on “core groups” or “at risk” populations focus on high-degree nodes. The assumption is that high-degree nodes are likely to contact lots of people.
Network Diffusion & Peer Influence

Network diffusion features

*Assortative mixing:*

A more traditional way to think about “star” effects.
Network Diffusion & Peer Influence

*Network diffusion features: simulation test*

- Simulation study: How do these different features compare over a collection of observed nets?
  - For each network trial:
    - Fix dyadic transmission probability
    - Randomly select a node as seed
    - Trace the diffusion path across the network
      - Measure speed & extent of spread
      - Model extent of spread by structural characteristics
  - First run: Add Health:
    - simple diffusion process
    - dyadic probability set to 0.08
    - 500 trials in each network.
  - Then expand to:
    - Different assumptions of dyadic transmission probability
    - Different data set (Facebook)
    - Complex diffusion models
Network Diffusion & Peer Influence

Network diffusion features: simulation test

Node size is proportional to ln(degree+1).

School 13

School 14

Legend:
- Orange: 7th Grade
- Blue: 8th Grade
- Purple: 9th Grade
- Green: 10th Grade
- Magenta: 11th Grade
- Gray: 12th Grade
Network Diffusion & Peer Influence

Network diffusion features: simulation test

Figure 2. Distribution of Diffusion Curves, School 13
Define as a general measure of the “diffusion susceptibility” of a graph’s structure as the ratio of the area under the observed curve to the area under the curve for a matching random network. As this gets smaller than 1.0, you get effectively slower median transmission.
# Network Diffusion & Peer Influence

*Network diffusion features: simulation test*

## Table 2. OLS Regression of Relative Diffusion Ratio on Network Structure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.62 ***</td>
<td>1.90 ***</td>
<td>1.02 ***</td>
<td>1.81 ***</td>
</tr>
<tr>
<td><strong>Connectivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td>-0.207 ***</td>
<td>-0.179 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent Paths</td>
<td>-0.077 ***</td>
<td>-0.056 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance x Paths</td>
<td>0.023 ***</td>
<td>0.015 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Clustering</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering Coefficient</td>
<td></td>
<td>-0.692 ***</td>
<td>-0.653 ***</td>
<td></td>
</tr>
<tr>
<td>Grade Homophily</td>
<td></td>
<td>-0.026 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peer Group Strength</td>
<td></td>
<td>-0.868 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Degree Distribution</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree Skew</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assortative Mixing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Control Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network Size/100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion Isolated</td>
<td></td>
<td></td>
<td></td>
<td>-1.106 ***</td>
</tr>
<tr>
<td>Non-Complete</td>
<td></td>
<td></td>
<td>-0.052 *</td>
<td></td>
</tr>
<tr>
<td>Adj- R²</td>
<td></td>
<td></td>
<td></td>
<td>0.76</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td>124</td>
</tr>
</tbody>
</table>

Adj- R² slopes are significant at: *** p < 0.001, ** p < 0.01, * p < 0.05.
# Network Diffusion & Peer Influence

*Network diffusion features: simulation test*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.90 ***</td>
<td>1.02 ***</td>
<td>1.02 ***</td>
<td>1.02 ***</td>
</tr>
<tr>
<td><strong>Connectivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent Paths</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance x Paths</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Clustering</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering Coefficient</td>
<td>-0.692 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade Homophily</td>
<td>-0.026 **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peer Group Strength</td>
<td>-0.868 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Degree Distribution</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree Skew</td>
<td></td>
<td></td>
<td>-0.023</td>
<td></td>
</tr>
<tr>
<td>Assortative Mixing</td>
<td></td>
<td></td>
<td>-0.189 *</td>
<td></td>
</tr>
<tr>
<td><strong>Control Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network Size/100</td>
<td>-0.005 ***</td>
<td>-0.005 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion Isolated</td>
<td>-1.106 ***</td>
<td>-0.984 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Complete</td>
<td>-0.052 *</td>
<td>-0.078 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj- R²</td>
<td>0.76</td>
<td>0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>124</td>
<td>124</td>
<td>124</td>
<td>124</td>
</tr>
</tbody>
</table>
## Network Diffusion & Peer Influence

*Network diffusion features: simulation test*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.62***</td>
<td>1.90***</td>
<td>1.02***</td>
<td>1.81***</td>
</tr>
<tr>
<td><strong>Connectivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td>-0.207***</td>
<td></td>
<td>-0.179***</td>
<td></td>
</tr>
<tr>
<td>Independent Paths</td>
<td>-0.077***</td>
<td></td>
<td>-0.056***</td>
<td></td>
</tr>
<tr>
<td>Distance $\times$ Paths</td>
<td>0.023***</td>
<td></td>
<td></td>
<td>0.015***</td>
</tr>
<tr>
<td><strong>Clustering</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering Coefficient</td>
<td></td>
<td>-0.692***</td>
<td></td>
<td>-0.653***</td>
</tr>
<tr>
<td>Grade Homophily</td>
<td></td>
<td>-0.026**</td>
<td></td>
<td>-0.007</td>
</tr>
<tr>
<td>Peer Group Strength</td>
<td></td>
<td>-0.868***</td>
<td></td>
<td>-0.141</td>
</tr>
<tr>
<td><strong>Degree Distribution</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree Skew</td>
<td></td>
<td></td>
<td>-0.023</td>
<td>-0.007</td>
</tr>
<tr>
<td>Assortative Mixing</td>
<td></td>
<td></td>
<td></td>
<td>-0.189* -0.059</td>
</tr>
<tr>
<td><strong>Control Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network Size/100</td>
<td>0.005***</td>
<td>-0.005***</td>
<td>-0.005***</td>
<td>0.004*</td>
</tr>
<tr>
<td>Proportion Isolated</td>
<td>-0.007</td>
<td>-1.106***</td>
<td>-0.984***</td>
<td>-0.300*</td>
</tr>
<tr>
<td>Non-Complete</td>
<td>-0.006</td>
<td>-0.052*</td>
<td>-0.078**</td>
<td>-0.006</td>
</tr>
<tr>
<td>Adj- $R^2$</td>
<td>0.85</td>
<td>0.76</td>
<td>0.60</td>
<td>0.90</td>
</tr>
<tr>
<td>N</td>
<td>124</td>
<td>124</td>
<td>124</td>
<td>124</td>
</tr>
</tbody>
</table>
Figure 4. Relative Diffusion Ratio
By Distance and Number of Independent Paths

Network Diffusion & Peer Influence

Network diffusion features: simulation test
Network Diffusion & Peer Influence
A closer look at emerging connectivity

Traditional “core group” models have a local-vision understanding of risk: those with lots of ties (high degree) are the focus for intervention and actions.

• In the short time-windows necessary for STD transfer, *low-degree networks* are the relevant features for transmission. What sorts of networks emerge when average degree (in the short run) is held to small numbers?

• How does the shape of the degree distribution matter? If activity is homogeneous do we get fundamentally different networks than if it is very heterogeneous?
Network Diffusion & Peer Influence
*A closer look at emerging connectivity*

Emergent Connectivity in low-degree networks

Partner Distribution

Component Size/Shape
Network Diffusion & Peer Influence

* A closer look at emerging connectivity
In both distributions, a giant component & reconnected core emerges as density increases, but at very different speeds and ultimate extent.
Network Diffusion & Peer Influence
*A closer look at emerging connectivity*

*Very small* changes in degree generate a quick cascade to large connected components. While not quite as rapid, STD cores follow a similar pattern, emerging rapidly and rising steadily with small changes in the degree distribution.

This suggests that, even in *the short run* (days or weeks, in some populations) large connected cores can emerge covering the majority of the interacting population, which can sustain disease.

The analogous curve for the long-tailed (“scale free”) distribution is gentler: with components emerging slowly and bicomponents never reaching a large set.
Network Diffusion & Peer Influence
A closer look at emerging connectivity

What distinguishes these two distributions?

Shape
Network Diffusion & Peer Influence
A closer look at emerging connectivity

What distinguishes these two distributions?
Shape:

The scale-free network’s signature is the long-tail
So what effect does changes in the shape have on connectivity?
Network Diffusion & Peer Influence

*A closer look at emerging connectivity*
Search Procedure:
1) Identify all valid degree distributions with the given mean degree and a maximum of 6 w. brute force search.
2) Map them to this space
3) Simulate networks each degree distribution
4) Measure size of components & Bicomponents
Network Diffusion & Peer Influence

A closer look at emerging connectivity
Network Diffusion & Peer Influence
A closer look at emerging connectivity

Figure 4. Degree Distribution and Simulated Giant Components
Network Diffusion & Peer Influence

*A closer look at emerging connectivity*
Network Diffusion & Peer Influence

A closer look at emerging connectivity
Network Diffusion & Peer Influence

A closer look at emerging connectivity
Network Diffusion & Peer Influence

A closer look at emerging connectivity

C:99%, B: 86%
Network Diffusion & Peer Influence

A closer look at emerging connectivity

**Figure 4.** Giant Components as a function of average degree and degree distribution shape.

**Figure 5.** Size of the Largest bicomponent by mean degree and distribution variability.

Largest Component  
(at least 1 path)

Largest Bicomponent  
(at least 2 paths)
Network Diffusion & Peer Influence

A closer look at emerging connectivity
Network Diffusion & Peer Influence

*Relational Dynamics*

Key Question: What features of a network contribute most to diffusion potential?

*In addition to* the dyadic probability that one actor passes something to another \((p_{ij})\), two factors affect flow through a network:

*Topology*
- the shape, or form, of the network
- Example: one actor cannot pass information to another unless they are either directly or indirectly connected

*Time*
- the timing of contact matters
- Example: an actor cannot pass information he has not received yet

Use simulation tools to explore the relative effects of structural connectivity features

*This is a big conditional! – lots of work on how the dyadic transmission rate may differ across populations.*
Network Diffusion & Peer Influence
Relational Dynamics

Three relevant networks

Discussions of network effects on STD spread often speak loosely of “the network.”

There are three relevant networks that are often conflated:

1) The contact network. The set of pairs of people connected by sexual contact. G(V,E).

2) The exposure network. A subset of the edges in the contact network where timing makes it possible for one person to pass infection to another.

3) The transmission network. The subset of the exposure network where disease is actually passed. In most cases this is a tree layered on (2) and rooted on a source/seed node.
Network Diffusion & Peer Influence

*Relational Dynamics*

**Three relevant networks**

Discussions of network effects on STD spread often speak loosely of “the network.”

There are three relevant networks that are often conflated:

Who can “A” reach?

Contact network: Everyone, it is a connected component
Network Diffusion & Peer Influence

Relational Dynamics

Three relevant networks

Discussions of network effects on STD spread often speak loosely of “the network.”

There are three relevant networks that are often conflated:

Who can “A” reach?

Exposure network: here, node “A” could reach up to 8 others
Network Diffusion & Peer Influence

Relational Dynamics

Three relevant networks

Discussions of network effects on STD spread often speak loosely of “the network.”

There are three relevant networks that are often conflated:

Who can “A” reach?

Transmission network: upper limit is 8 through the exposure links (dark blue). Transmission is path dependent: if no transmission to B, then also none to {K,L,O,J,M}
Concurrency

The mapping between the contact network and the exposure network is based on relational timing. In a dynamic network, edge timing determines if something can flow down a path because things can only be passed forward in time.

Definitions:

Two edges are adjacent if they share a node.

A path is a sequence of adjacent edges \((E_1, E_2, \ldots E_d)\).

A time-ordered path is a sequence of adjacent edges where, for each pair of edges in the sequence, the start time \(S_i\) is less than or equal to \(E_j\):

\[ S(E_1) \leq E(E_2) \]

Adjacent edges are concurrent if they share a node and have start and end dates that overlap. This occurs if:

\[ S(E_2) \leq E(E_1) \]
Concurrenty

The mapping between the contact network and the exposure network is based on relational timing. In a *dynamic* network, edge timing determines if something can flow down a path because *things can only be passed forward in time*. 
Network Diffusion & Peer Influence

Relational Dynamics

The constraints of time-ordered paths change our understanding of the system structure of the network. Paths make a network a system: linking actors together through indirect connections. Relational timing changes how paths cumulate in networks.

*Indirect connectivity is no longer transitive:*

Here A can reach C, and C and reach D. But A cannot reach D (nor D A). Why? Because any infection A passes to C would have happened *after* the relation between C and D ended.
Edge time structures are characterized by sequence, duration and overlap.

Paths between $i$ and $j$, have length and duration, but these need not be symmetric even if the constituent edges are symmetric.
Network Diffusion & Peer Influence
Relational Dynamics

Implied Contact Network of 8 people in a ring
All relations Concurrent

Reachability = 1.0
Network Diffusion & Peer Influence
Relational Dynamics

Reachability = 0.71

Implied Contact Network of 8 people in a ring
Serial Monogamy (1)
Network Diffusion & Peer Influence
Relational Dynamics

Implied Contact Network of 8 people in a ring
Mixed Concurrent

Reachability = 0.57
Network Diffusion & Peer Influence

Relational Dynamics

Implied Contact Network of 8 people in a ring

Serial Monogamy (3)

Reachability = 0.43
Timing alone can change mean reachability from 1.0 when all ties are concurrent to 0.42.

In general, ignoring time order is equivalent to assuming all relations occur simultaneously – assumes perfect concurrency across all relations.
**Network Diffusion & Peer Influence**

*Relational Dynamics*

*Path distances no longer simply add*

While *a* is 2 steps from *d*, and *d* is 1 step from *e*, *a* and *e* are 4 steps apart.

This is because the shorter path from *a* to *e* emerges *after* the path from *d* to *e* ended.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
The **geodesic** from A to D is AE, ED and is two steps long.

But the **fastest** path would be AB, BC, CD, which while 3 steps long could get there by time 5 compared to time 7.

I ignore this feature for the remainder…
Concurrency affects exposure by making paths symmetric, which increases exposure down multiple “branches” of a contact sequence. Consider a simplified example:

All edges concurrent
Concurrency affects exposure by making paths symmetric, which increases exposure down multiple “branches” of a contact sequence. Consider a simplified example:

All edges except gd concurrent

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
<th>j</th>
<th>k</th>
<th>l</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>b</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>c</td>
<td>1</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>d</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>e</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>f</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>g</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>h</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>i</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>j</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>k</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>l</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>.</td>
<td>1</td>
</tr>
<tr>
<td>m</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>.</td>
</tr>
</tbody>
</table>
Concurrency affects exposure by making paths symmetric, which increases exposure down multiple “branches” of a contact sequence. Consider a simplified example:

All edges gd after all others:

The concurrency status of \{dg\} determines which “side” of the graph is exposed. Note this effect happens at the system level – the correlation between exposure and node-level timing is essentially zero (d/g excepted)
Resulting infection trace from a simulation (Morris et al, AJPH 2010).
Resulting infection trace from a simulation (Morris et al, AJPH 2010).
Timing constrains potential diffusion paths in networks, since bits can flow through edges that have ended.

This means that:

- Structural paths are not equivalent to the diffusion-relevant path set.
- Network distances don’t build on each other.
- Weakly connected components overlap without diffusion reaching across sets.
- Small changes in edge timing can have dramatic effects on overall diffusion
- Diffusion potential is maximized when edges are concurrent and minimized when they are “inter-woven” to limit reachability.

Combined, this means that many of our standard path-based network measures will be incorrect on dynamic graphs.
Network Diffusion & Peer Influence

Relational Dynamics – other implication of dynamic nets (briefly)

**Structural measurement implications**

The distribution of paths is important for many of the measures we typically construct on networks, and these will be change if timing is taken into consideration:

Centrality:
- Closeness centrality
- Path Centrality
- Information Centrality
- Betweenness centrality

Network Topography
- Clustering
- Path Distance

Groups & Roles:
- Correspondence between degree-based position and reach-based position
- Structural Cohesion & Embeddedness
- Opportunities for Time-based block-models (similar reachability profiles)

In general, any measures that take the *systems* nature of the graph into account will differ.
Network Diffusion & Peer Influence
Relational Dynamics – other implication of dynamic nets (briefly)

**Structural measurement implications**

*New versions of classic reachability measures:*

1) **Temporal reach**: The $ij$ cell = 1 if $i$ can reach $j$ through time.

2) **Temporal geodesic**: The $ij$ cell equals the number of steps in the shortest path linking $i$ to $j$ over time.

3) **Temporal paths**: The $ij$ cell equals the number of time-ordered paths linking $i$ to $j$.

These will only equal the standard versions when all ties are concurrent.

*Duration explicit measures*

4) **Quickest path**: The $ij$ cell equals the shortest time within which $i$ could reach $j$.

5) **Earliest path**: The $ij$ cell equals the real-clock time when $i$ could first reach $j$.

6) **Latest path**: The $ij$ cell equals the real-clock time when $i$ could last reach $j$.

7) **Exposure duration**: The $ij$ cell equals the longest (shortest) interval of time over which $i$ could transfer a good to $j$.

Each of these also imply different types of “betweenness” roles for nodes or edges, such as a “limiting time” edge, which would be the edge whose comparatively short duration places the greatest limits on other paths.
Network Diffusion & Peer Influence

*Structural Moderators of Timing Effects*

Topology & Time interact: How relational sequencing affects diffusion is conditioned by the structural patterns of relations.

Examples:
- Time limitations mean star nodes can’t interact with everyone at each time tick; effects of high degree are thus limited by schedule/availability.
- If within-cluster ties are also more frequent than between cluster ties, then the effects of communities will be magnified.
- Multiple connectivity should provide routes around breaks built by temporal sequence.
Network Diffusion & Peer Influence

*Structural Moderators of Timing Effects*

**Measures**

Dependent variable: Reachability in the exposure graph. This is the proportion of pairs in the network that are reachable in time.
Network Diffusion & Peer Influence

*Structural Moderators of Timing Effects*

Measures: Independent variables

Features of the topology. Of key interest is the level of structural cohesion.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>--</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>--</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>--</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>--</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>--</td>
</tr>
</tbody>
</table>

Cell Value = highest k-connected component pair belongs to.

Average = 1.25
Network Diffusion & Peer Influence

Structural Moderators of Timing Effects

Measures: Independent variables

Features of the topology. Of key interest is the level of structural cohesion.

Cell Value = highest k-connected component pair belongs to.

Average = 1.6
Network Diffusion & Peer Influence

*Structural Moderators of Timing Effects*

**Measures: Independent variables**

Features of the topology. Of key interest is the level of structural cohesion.

![Average connectivity](image)

**Average connectivity**

- Mean 1.196506
- std 0.191079
## Network Diffusion & Peer Influence

*Structural Moderators of Timing Effects*

<table>
<thead>
<tr>
<th></th>
<th>Volume</th>
<th>Distance</th>
<th>Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>“High Cohesive”</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nodes: 148</td>
<td>Mean: 3.59</td>
<td>Largest BC: 0.51</td>
<td></td>
</tr>
<tr>
<td>Mean Deg: 6.16</td>
<td>Diameter: 5</td>
<td>Pairwise K: 1.57</td>
<td></td>
</tr>
<tr>
<td>Density: 0.042</td>
<td>Centralization: 0.312</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centralization: 0.187</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nodes: 80</td>
<td>Mean: 3.02</td>
<td>Largest BC: 0.33</td>
<td></td>
</tr>
<tr>
<td>Mean Deg: 5.27</td>
<td>Diameter: 5</td>
<td>Pairwise K: 1.34</td>
<td></td>
</tr>
<tr>
<td>Density: 0.067</td>
<td>Centralization: 0.413</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centralization: 0.373</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nodes: 154</td>
<td>Mean: 4.99</td>
<td>Largest BC: 0.08</td>
<td></td>
</tr>
<tr>
<td>Mean Deg: 3.71</td>
<td>Diameter: 8</td>
<td>Pairwise K: 1.07</td>
<td></td>
</tr>
<tr>
<td>Density: 0.025</td>
<td>Centralization: 0.259</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centralization: 0.147</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nodes: 128</td>
<td>Mean: 4.55</td>
<td>Largest BC:</td>
<td></td>
</tr>
<tr>
<td>Mean Deg: 3.39</td>
<td>Diameter: 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Density: 0.027</td>
<td>Centralization: 0.301</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centralization: 0.205</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Exemplar independent variables**
Network Diffusion & Peer Influence
Structural Moderators of Timing Effects

Relation between Concurrency & Exposure in 4 contact networks

Proportion of pairs reachable in the exposure network vs. Proportion of relations that overlap in time
Network Diffusion & Peer Influence

Structural Moderators of Timing Effects

Relation between Concurrency & Exposure in 4 contact networks

“Low Cohesive”
Network Diffusion & Peer Influence

Structural Moderators of Timing Effects
Network Diffusion & Peer Influence

*Structural Moderators of Timing Effects*

Colors=different nets
Panels=cohesion level
Figure 5, Exposure Reachability by Concurrency
Predicted values for models 1 and 4 based on Table 2, shaded areas represent 95% CIs
Network Diffusion & Peer Influence

*Structural Moderators of Timing Effects*

1. Concurrency has a necessarily positive effect on potential diffusion exposure
   1. This implies that we should see greater transmission given greater concurrency
   2. This works by creating “multiple routes” in the exposure path structure

2. Structural cohesion captures multiple routes in the contact graph
   1. Higher levels of cohesion increase exposure by directly increasing the underlying transmission substrate

3. There is a negative interaction between cohesion and concurrency: as cohesion increases, the relative returns to concurrency decrease.
   1. But this comes at the cost of a higher base-level of exposure.
Network Diffusion & Peer Influence

*Structural Transmission Dynamics: beyond disease diffusion*

**Complex Contagion**

Thus far we have focused on a “simple” dyadic diffusion parameter, $p_{ij}$, where the probability of passing/receiving the bit is purely dependent on discordant status of the dyad, sometimes called the “independent cascade model” (), which suggests a monotonic relation between the number of times you are exposed through peers.

High exposure could be due to repeated interaction with one person or weak interaction with many, effectively equating:

Alternative models exist. Under “complex contagion” for example, the likelihood that I accept the bit that flows through the network depends on the proportion of my peers that have the bit.
Complex Contagion

Assume adoption requires $k$ neighbors having adopted, then transmission can only occur within dense clusters:
Network Diffusion & Peer Influence

*Structural Transmission Dynamics: beyond disease diffusion*

Complex Contagion

Assume adoption requires $k$ neighbors having adopted, then transmission can only occur within dense clusters:

Assume $p_{ij}=1$, $k=2$, starting nodes in yellow
Network Diffusion & Peer Influence

*Structural Transmission Dynamics: beyond disease diffusion*

Complex Contagion

Assume adoption requires $k$ neighbors having adopted, then transmission can only occur within dense clusters:

For this network under weak complex diffusion ($k=2$), the maximum risk size is 8.
Network Diffusion & Peer Influence

*Structural Transmission Dynamics: beyond disease diffusion*

Complex Contagion

Assume adoption requires $k$ neighbors having adopted, then transmission can only occur within dense clusters:

For this network under weak complex diffusion ($k=2$), the maximum risk size is reaches 98%.
Complex Contagion

Can lead to widely varying sizes of potential diffusion cascades. Here’s the distribution across all PROPSPER schools:

Distribution is largely bimodal (even with a connected pair start)
Network Diffusion & Peer Influence

*Structural Transmission Dynamics: beyond disease diffusion*

Complex Contagion

Can lead to widely varying sizes of potential diffusion cascades. Here’s the distribution across all PROPSPER schools:

The governing factors are (a) curved effect of local redundancy and (b) structural cohesion
Network Diffusion & Peer Influence

Structural Transmission Dynamics: beyond disease diffusion

Complex Contagion

Does get used for real health work:

Here, authors assume a CC process, seeded with observed depressive cases, turn that into a Markov model and ask what parameters would maximize fit from simulated to observed.

Spreading of healthy mood in adolescent social networks

E. M. Hill¹, F. E. Griffiths² and T. House¹,³

¹Centre for Complexity Science and Warwick Infectious Disease Epidemiology Research Centre, and ²Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK
³School of Mathematics, University of Manchester, Manchester M13 9PL, UK

Depression is a major public health concern worldwide. There is evidence that social support and befriending influence mental health, and an improved understanding of the social processes that drive depression has the potential to bring significant public health benefits. We investigate transmission of mood on a social network of adolescents, allowing flexibility in our model by making no prior assumption as to whether it is low mood or healthy mood that spreads. Here, we show that while depression does not spread, healthy mood among friends is associated with significantly reduced risk of developing and increased chance of recovering from depression. We found that this spreading of healthy mood can be captured using a non-linear complex contagion model. Having sufficient friends with healthy mood can halve the probability of developing, or double the probability of recovering from, depression over a 6–12-month period on an adolescent social network. Our results suggest that promotion of friendship between adolescents can reduce both incidence and prevalence of depression.
Complex diffusion is just the most well studied of the options that combine transmission with some pairwise positional feature. This is a wide-open area for future research.

The basic idea is that transmission is increased/decreased if there is some third structural property that the susceptible & infected pair share.

This leads us into the general problem of peer influence models…when do peers change each other’s behaviors?
Network Diffusion & Peer Influence

*Peer Influence Dynamics*

**Background:**

- Long standing research interest in how our relations shape our attitudes and behaviors.

- Most often assumed mechanism is that people (through conversation or similar) change each others beliefs/opinions, which changes behavior. This implies that position in a communication network should be related to attitudes.

- **Alternatives:**
  - Modeling behavior: ego copies behavior of alter to gain respect, esteem, etc.
  - Distinction: Ego tries to be different from (some) alter to gain respect, esteem, etc.
  - Access: Ego wants to do Y, but can only do so because alter provides access (say, being old enough to buy cigarettes).
Network Diffusion & Peer Influence

Peer Influence Dynamics

Background:

- Early work was ego-centric – people informed on their peers
  - Seems to have inflated PI effects by ~50% or so…either through projection of ego behavior onto peers or selective interaction (what alters do with ego may be different than what alter does all the time).

- Then to cross sectional associations based on alter self-reports
  - Better, but still likely conflates selection with influence

- Next to dynamic models:
  - Ego Behavior(t) \sim f(ego behavior(t-1) + alter behavior (t-1) + controls
  - Much better; still debate on (a) correct estimation functions, (b) unobserved selection features that confound causal inference.
  - Development of Actor-oriented models (SIENA)
Network Diffusion & Peer Influence

*Peer Influence Dynamics*

Background:

- Finally: Experimental manipulation of peer exposure
  - “Gold standard” for isolation of peer effects
    - Likely strongly underestimates effects (as measure intent to treat, not take-up of treatment, since people may not care about relations that can be manipulated).

$\beta(\text{Peer}(y))$: Ego Inform $<$ Alter Inform $<$ Cross Sectional $<$ Dynamic $<$ Experimental. Still often find peer effects, but my sense is that we’ve (strongly) over-corrected at this point.
Network Diffusion & Peer Influence

Peer Influence Dynamics

Freidkin’s *Structural Theory of Social Influence* :

Two-part model:

Beliefs are a function of two sources:

a) Individual characteristics
   • Gender, Age, Race, Education, Etc. Standard sociology

b) Interpersonal influences
   • Actors negotiate with others
Network Diffusion & Peer Influence

Peer Influence Dynamics

\[ Y^{(1)} = XB \] \hspace{1cm} (1)

\[ Y^{(t)} = \alpha W Y^{(T-1)} + (1 - \alpha) Y^{(1)} \] \hspace{1cm} (2)

\( Y^{(1)} \) = an \( N \times M \) matrix of initial opinions on \( M \) issues for \( N \) actors

\( X \) = an \( N \times K \) matrix of \( K \) exogenous variable that affect \( Y \)

\( B \) = a \( K \times M \) matrix of coefficients relating \( X \) to \( Y \)

\( \alpha \) = a weight of the strength of endogenous interpersonal influences

\( W \) = an \( N \times N \) matrix of interpersonal influences
This is the standard sociology model for explaining anything: the General Linear Model.

It says that a dependent variable (Y) is some function (B) of a set of independent variables (X). At the individual level, the model says that:

$$Y_i = \sum_k X_{ik} B_k$$

Usually, one of the X variables is $\varepsilon$, the model error term.
This part of the model taps social influence. It says that each person’s final opinion is a weighted average of their own initial opinions and the opinions of those they communicate with (which can include their own current opinions).
The key to the peer influence part of the model is $W$, a matrix of interpersonal weights. $W$ is a function of the communication structure of the network, and is usually a transformation of the adjacency matrix. In general:

\[ 0 \leq w_{ij} \leq 1 \]

\[ \sum_{j} w_{ij} = 1 \]

Various specifications of the model change the value of $w_{ii}$, the extent to which one weighs their own current opinion and the relative weight of alters.
### Network Diffusion & Peer Influence

**Peer Influence Dynamics**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

### Self weight:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.33</td>
<td>.33</td>
<td>.33</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>.33</td>
<td>.33</td>
<td>.33</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>.25</td>
<td>.25</td>
<td>.25</td>
<td>.25</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>.50</td>
<td>.50</td>
</tr>
</tbody>
</table>

### Even

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.50</td>
<td>.25</td>
<td>.25</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>.25</td>
<td>.50</td>
<td>.25</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>.20</td>
<td>.20</td>
<td>.40</td>
<td>.20</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>.33</td>
<td>.67</td>
</tr>
</tbody>
</table>

### 2*Self

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.50</td>
<td>.25</td>
<td>.25</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>.25</td>
<td>.50</td>
<td>.25</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>.17</td>
<td>.17</td>
<td>.50</td>
<td>.17</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>.50</td>
<td>.50</td>
</tr>
</tbody>
</table>

### Degree

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.50</td>
<td>.25</td>
<td>.25</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>.25</td>
<td>.50</td>
<td>.25</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>.17</td>
<td>.17</td>
<td>.50</td>
<td>.17</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>.50</td>
<td>.50</td>
</tr>
</tbody>
</table>
Network Diffusion & Peer Influence

Peer Influence Dynamics

Formal Properties of the model

\[ Y^{(t)} = \alpha W Y^{(T-1)} + (1 - \alpha) Y^{(1)} \] (2)

When interpersonal influence is complete, model reduces to:

\[ Y^{(t)} = 1 W Y^{(T-1)} + 0 Y^{(1)} = W Y^{(T-1)} \]

When interpersonal influence is absent, model reduces to:

\[ Y^{(t)} = 0 W Y^{(T-1)} + Y^{(1)} = Y^{(1)} \]
Formal Properties of the model

If we allow the model to run over $t$ and $W$ remains constant:

$$Y^{(\infty)} = \alpha W Y^{(\infty)} + (1 - \alpha) XB$$

The model is directly related to spatial econometric models:

$$Y^{(\infty)} = \alpha W Y^{(\infty)} + \tilde{X} \beta + \varepsilon$$

Where the two coefficients ($\alpha$ and $\beta$) are estimated directly (See Doreian, 1982, SMR).

This is the linear network auto correlation model, best bet with cross-sectional data (and randomization trick to estimate se)
Network Diffusion & Peer Influence

*Peer Influence Dynamics*

Simple example

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.33</td>
<td>.33</td>
<td>.33</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>.33</td>
<td>.33</td>
<td>.33</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>.25</td>
<td>.25</td>
<td>.25</td>
<td>.25</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>.50</td>
<td>.50</td>
<td>7</td>
</tr>
</tbody>
</table>

\[ \alpha = .8 \]

<table>
<thead>
<tr>
<th>T</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>2.60</td>
<td>2.81</td>
<td>2.93</td>
<td>2.98</td>
<td>3.00</td>
<td>3.01</td>
<td>3.01</td>
</tr>
<tr>
<td>3</td>
<td>3.00</td>
<td>3.00</td>
<td>3.21</td>
<td>3.33</td>
<td>3.38</td>
<td>3.40</td>
<td>3.41</td>
<td>3.41</td>
</tr>
<tr>
<td>5</td>
<td>5.00</td>
<td>4.20</td>
<td>4.20</td>
<td>4.16</td>
<td>4.14</td>
<td>4.14</td>
<td>4.13</td>
<td>4.13</td>
</tr>
<tr>
<td>7</td>
<td>7.00</td>
<td>6.20</td>
<td>5.56</td>
<td>5.30</td>
<td>5.18</td>
<td>5.13</td>
<td>5.11</td>
<td>5.10</td>
</tr>
</tbody>
</table>

By t=7, still variability in Y
Network Diffusion & Peer Influence
Peer Influence Dynamics

Simple example

\[
\begin{array}{cccccc}
1 & 2 & 3 & 4 & \text{Y} \\
1 & .33 & .33 & .33 & 0 & 1 \\
2 & .33 & .33 & .33 & 0 & 3 \\
3 & .25 & .25 & .25 & .25 & 5 \\
4 & 0 & 0 & .50 & .50 & 7 \\
\end{array}
\]

\[\alpha = 1.0\]

\[
\begin{array}{cccccccc}
\text{T:} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
1.00 & 3.00 & 3.33 & 3.56 & 3.68 & 3.74 & 3.78 & 3.81 \\
3.00 & 3.00 & 3.33 & 3.56 & 3.68 & 3.74 & 3.78 & 3.81 \\
5.00 & 4.00 & 4.00 & 3.92 & 3.88 & 3.86 & 3.85 & 3.84 \\
7.00 & 6.00 & 5.00 & 4.50 & 4.21 & 4.05 & 3.95 & 3.90 \\
\end{array}
\]

By \(t=7\), almost no variability in \(Y\)
Network Diffusion & Peer Influence

Peer Influence Dynamics

Extended example: building intuition

Consider a network with three cohesive groups, and an initially random distribution of opinions:
Simulated Peer Influence:
75 actors, 2 initially random opinions, Alpha = .8, 7 iterations
Simulated Peer Influence:
75 actors, 2 initially random opinions, Alpha = .8, 7 iterations
Simulated Peer Influence:
75 actors, 2 initially random opinions, Alpha = .8, 7 iterations
Simulated Peer Influence:
75 actors, 2 initially random opinions, Alpha = .8, 7 iterations
Simulated Peer Influence:
75 actors, 2 initially random opinions, Alpha = .8, 7 iterations
Simulated Peer Influence:
75 actors, 2 initially random opinions, Alpha = .8, 7 iterations
Simulated Peer Influence:
75 actors, 2 initially random opinions, Alpha = .8, 7 iterations
Simulated Peer Influence:
75 actors, 2 initially random opinions, Alpha = .8, 7 iterations
Extended example: building intuition

Consider a network with three cohesive groups, and an initially random distribution of opinions:

Now weight in-group ties higher than between group ties
Simulated Peer Influence:
75 actors, 2 initially random opinions, Alpha = 0.8, 7 iterations, in-group tie: 2
Consider the implications for populations of different structures. For example, we might have two groups, a large orthodox population and a small heterodox population. We can imagine the groups mixing in various levels:

Heterodox: 10 people
Orthodox: 100 People

<table>
<thead>
<tr>
<th></th>
<th>Little Mixing</th>
<th>Moderate Mixing</th>
<th>Heavy Mixing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterodox</td>
<td>.95 .001</td>
<td>.95 .008</td>
<td>.95 .05</td>
</tr>
<tr>
<td>Orthodox</td>
<td>.001 .02</td>
<td>.008 .02</td>
<td>.05 .02</td>
</tr>
</tbody>
</table>
Light

Moderate

Heavy
Light mixing
Light mixing
Light mixing
Light mixing
Light mixing
Light mixing
Moderate mixing
Moderate mixing
Moderate mixing
Moderate mixing
Moderate mixing
Moderate mixing
High mixing
High mixing
High mixing
High mixing
High mixing
In an unbalanced situation (small group vs large group) the extent of contact can easily overwhelm the small group. Applications of this idea are evident in:

• Missionary work (Must be certain to send missionaries out into the world with strong in-group contacts)

• Overcoming deviant culture (I.e. youth gangs vs. adults)

• This is also the mechanism behind why most youth peer influence is a *good* thing – most youth are well behavior and civic minded...so are exerting positive influences on their peers.
Friedkin (1998) generalizes the model so that alpha varies across people. 
(1) simply changing $\alpha$ to a vector ($A$), which then changes each person’s opinion directly

(2) by linking the self weight ($w_{ii}$) to alpha.

$$Y^{(t)} = AWY^{(T-1)} + (I - A)Y^{(1)}$$

Were A is a diagonal matrix of endogenous weights, with $0 \leq a_{ii} \leq 1$. A further restriction on the model sets $w_{ii} = 1-a_{ii}$

This leads to a great deal more flexibility in the theory, and some interesting insights. Consider the case of group opinion leaders with unchanging opinions (I.e. many people have high $a_{ii}$, while a few have low):
Peer Opinion Leaders

[Graph showing data points for Group 1, Group 2, and Group 3 leaders.]
Peer Opinion Leaders
Peer Opinion Leaders
Peer Opinion Leaders
Peer Opinion Leaders
Peer Opinion Leaders
Further extensions of the model might:

- Time dependent $\alpha$: people likely value other’s opinions more early than later in a decision context

- Interact $\alpha$ with $XB$: people’s self weights are a function of their behaviors & attributes

- Make $W$ dependent on structure of the network (weight transitive ties greater than intransitive ties, for example)

- Time dependent $W$: The network of contacts does not remain constant, but is dynamic, meaning that influence likely moves unevenly through the network

- And others likely abound….
Network Diffusion & Peer Influence

*Peer Influence Dynamics*

There are two common ways to test for peer associations through networks.

The first estimates the parameters ($\alpha$ and $\beta$) of the network autocorrelation model directly, the second transforms the network into a *dyadic* model, predicting similarity among actors.

Peer influence model:

$$Y^{(\infty)} = \alpha W Y^{(\infty)} + \tilde{X} \beta + \varepsilon$$

This is the linear network autocorrelation model, and as specified, the model makes strong assumptions about equilibrium opinion and static relations.

$\rightarrow$ Some variants on this also expand $e$ to include alternative autocorrelation in the error structure.
There are two common ways to test for peer associations through networks.

The first estimates the parameters ($\alpha$ and $\beta$) of the network autocorrelation model directly, the second transforms the network into a *dyadic* model, predicting similarity among actors.

Peer influence model:

\[
Y^{(\infty)} = \alpha WY^{(\infty)} + \tilde{X}\beta + \varepsilon
\]

Note that since $WY^{\infty}$ is a simple vector -- weighted mean of friends $Y$ -- which can be constructed and added to your GLM model. That is, multiple $Y$ by a $W$ matrix, and run the regression with $WY$ as a new variable, and the regression coefficient is an estimate of $\alpha$. This is what Doriean calls the QAD estimate of peer influence.

It’s wrong, $\alpha$ will be biased, but it’s often not terribly wrong if most obvious selection factors are built into $X$. 
An obvious problem with this specification is that cases are, by definition, not independent, hence “network autocorrelation” terminology.

In practice, the QAD approach (perhaps combined with a GLS estimator) results in empirical estimates that are “virtually indistinguishable” from MLE (Doreian et al, 1984)

The proper way to estimate the peer equation is to use maximum likelihood estimates, and Doreian gives the formulas for this in his paper, and Carter Butts has implemented in in R with the LNAM procedure.

An alternative is to use non-parametric approaches, such as the Quadratic Assignment Procedure, to estimate the effects.
Peer influence through Dyad Models

Another way to get at peer influence is not through the level of Y, but by assessing the similarity of connected peers. Recall the simulated example: peer influence is reflected in how close points are to each other.

Network Diffusion & Peer Influence

Peer Influence Dynamics
Network Diffusion & Peer Influence

Peer Influence Dynamics

Peer influence through Dyad Models

The model is now expressed at the dyad level as:

\[ Y_{ij} = b_0 + b_1 A_{ij} + \sum_k b_k X_k + e_{ij} \]

Where \( Y \) is a matrix of similarities, \( A \) is an adjacency matrix, and \( X_k \) is a matrix of similarities on attributes.

**Advantages** include ease of specifying relation-specific similarity functions. You can add different features of a relation by adjusting/adding “\( A_{ij} \)” variables.

**Disadvantage** is that now in addition to network autocorrelation, you have repeated cases (on both sides).

But these can be dealt with using non-parametric modeling & testing techniques (QAP, for example). (which we will go over this afternoon)
Network Diffusion & Peer Influence

Peer Influence & Health: Current Lit & Controversies

The NEW ENGLAND JOURNAL of MEDICINE

SPECIAL ARTICLE

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.
Network Diffusion & Peer Influence

Peer Influence & Health: Current Lit & Controversies

Used the friend/relative tracking data from a larger heart-health study to identify network contacts, including friends.
Network Diffusion & Peer Influence
Peer Influence & Health: Current Lit & Controversies

Used the friend/relative tracking data from a larger heart-health study to identify network contacts, including friends.
Network Diffusion & Peer Influence

*Peer Influence & Health: Current Lit & Controversies*

Used the friend/relative tracking data from a larger heart-health study to identify network contacts, including friends.
Network Diffusion & Peer Influence

*Peer Influence & Health: Current Lit & Controversies*

Used the friend/relative tracking data from a larger heart-health study to identify network contacts, including friends.
Network Diffusion & Peer Influence
*Peer Influence & Health: Current Lit & Controversies*

Used the friend/relative tracking data from a larger heart-health study to identify network contacts, including friends.
Network Diffusion & Peer Influence

*Peer Influence & Health: Current Lit & Controversies*

Used the friend/relative tracking data from a larger heart-health study to identify network contacts, including friends.
Used the friend/relative tracking data from a larger heart-health study to identify network contacts, including friends.
Network Diffusion & Peer Influence

Peer Influence & Health: Current Lit & Controversies

Used the friend/relative tracking data from a larger heart-health study to identify network contacts, including friends.

The network shows significant evidence of weight-homophily.
Network Diffusion & Peer Influence

Peer Influence & Health: Current Lit & Controversies

Used the friend/relative tracking data from a larger heart-health study to identify network contacts, including friends.

Effects of peer obesity on ego, by peer type

Edge-wise regressions of the form:

\[ E_{\text{Current}} = \beta_1(Alt_{\text{Current}}) + \beta_2(Alt_{\text{previous}}) + \beta_3(Ego_{\text{previous}}) + \text{Controls} \]

Ego is repeated for all alters; models include random effects on ego id.
This modeling strategy pools observations on edges and estimates a global effect net of change in ego/alter as a control. Here color is a single ego, number is wave (only 2 egos and 3 waves represented).

$$Ego_{Current} = \beta_1(Alt_{Current}) + \beta_2(Alt_{previous}) + \beta_3(Ego_{previous}) + Controls$$
Alterative specifications include using change-change models and allowing for a random effect of peers. This allows for greater variability in peer effects, and the potential to model differences.

\[ E_{\text{Current}} = \beta_e (A_{\text{Current}}) + \beta_2 (A_{\text{previous}}) + \beta_3 (E_{\text{previous}}) + \text{Controls} \]

\[ \beta_e = \beta + \nu_e \]

Or difference models:

\[ (E_{\text{previous}} - E_{\text{Current}}) = \beta_e (A_{\text{previous}} - A_{\text{current}}) + \text{Controls} \]

\[ \beta_e = \beta + \nu_e \]
Network Diffusion & Peer Influence

Peer Influence & Health: Current Lit & Controversies

Critiques of C&F

The C&F studies – of obesity, but also other work on the FHS data – turn on the validity of the causal association.

All turn on some issue of model miss-specification, typically:

• Can’t truly distinguish a network effect from other sources of common influence
  • “Selection” (“homophily”) or “Common influence” (“Shared environment”)
  • The most strident work in this area (Salizi

• Statistical errors
  • Misinterpretation of confidence intervals
  • Poorly specified/estimated models

C&H do a nice job of laying out their responses here:
http://jhfowler.ucsd.edu/examining_dynamic_social_networks.pdf and here:
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2597062/
Critiques of C&F


Use the same models as C&F on Add Health to show that things which are theoretically unlikely to be contagious appear to be in this form of model.

Note these coefficients are substantially smaller than C&F and only significant at the 0.1 level; and not robust to any sensitivity analysis.
1) C&F claim that differences in directional effects support a PI story:

- C&F: While mutual friends and ego→alter friends are > 0, alter→ego is not, means ego is emulating alter.
- Lyons notes these CIs overlap too much to make any claim about distinguishing them from each other.

*Figure 1*. Coefficient estimates and 2 SE (95%) confidence intervals for directional effects. For each study, the order from top to bottom is (1) mutual friendship, (2) FP named LP, then (3) LP named FP. The CIs overlap so much that one cannot infer that the differences are statistically significant. Sources: Christakis and Fowler (2007, suppl. p. 3); Christakis and Fowler (2008, suppl. p. 18); Fowler and Christakis (2008a, suppl. p. 9); Cacioppo et al. (2009, pp. 983–984).
Network Diffusion & Peer Influence

Peer Influence & Health: Current Lit & Controversies

Critiques of C&F  
Lyons, 2011.

2) Insufficient controls for Homophily
   • C& F: Use of alter’s lagged Y to control for homophily. Logic is that any feature that selected us to be friends at t-1 would have had it’s effect then.
   • Lyons notes that current and lagged have opposite signs, which seems suspect, and anyway is an insufficient control. He’s likely right here…

3) Directionality cannot distinguish the source of association
   • C& F: the ordering: mutual, ego→alter, alter→ego suggests an “esteem” model, where ego copies the behavior of alter.
   • Lyons argues that we would expect the same logic from a simple “foci” of similarity. *I don’t find this argument convincing.*

4) Random permutation tests cannot establish 3-degree rule
   • C& F: Association between alters at 1, 2, 3 degrees of separation are higher than we’d expect by chance, based on a permutation test.
   • Lyons invalid if the data are incomplete, which they certainly are. *I don’t find this argument convincing…data are always incomplete…*
5) The models are statistically inconsistent (if not incoherent)

- C&F: Use separate models for each type of tie, with random effects on ego.
- Lyons notes that these really should be treated as simultaneous equations, with shared error structures and so forth. Doing so (a) leads to unidentified models that must force the estimation of the peer effect to 0. That observed \(^0\) indicates something amiss.

- *Strikes me as a bit down in the weeds and I’m not convinced here that he’s critiquing them for what they are really doing (argues there are more equations than data, which is patently not true).*
Critiques of C&F  
Lyons, 2011.

My sense is that the strategy C&F took was not fundamentally misguided, but the model specification is probably thin; certainly in the obesity paper – less so in some of the later papers appearing after these debates.

Ideally you’d have a much better *direct* model for selection – perhaps even a separate two-stage model (see the Siena module), but here there are very limited observational controls, which would have been easy to add. In later specifications, they do add fixed effects for ego and still find similar results.

Commenting on the debate on SocNET – and a related conclusion that only experiments could provide valid inference – Tom Snijders says:

“The logical consequence of this is that we are stuck with imperfect methods. Lyons argues as though only perfect methods are acceptable, and while applauding such lofty ideals I still believe that we should accept imperfection, in life as in science. Progress is made by discussion and improvement of imperfections, not by their eradication.”

For a full general discussion, see: https://www.lists.ufl.edu/cgi-bin/wa?A2=ind1106&L=SOCNET&P=R11428
Network Diffusion & Peer Influence

Peer Influence & Health: Current Lit & Controversies

Shalizi & Thomas: PI is *generally* confounded

So long as there is an unobserved X that causes both ties and behavior, the effect of peers is unidentified.
Shalizi & Thomas: PI is *generally* confounded

![Diagram of causal graphs]

Figure 3: Modifications of the causal graph shown in Figure 1, in which observable covariates (Z) conveys enough information about X that contagion effects are unconfounded with latent homophily. In a (left), Z carries all of the causal effect from X to the observable outcome Y; in b (right), Z carries all of the effect from X to the social network tie A.

Only route out is to make X fully informed (or informing) by an observable Z….but realistically there are few things that (a) cause behavior exclusively without any selection pressure (a) or cause ties exclusively without any influence pressure (b) (though note b is what experimental assignments do)
Shalizi & Thomas: PI is *generally* confounded

Should be noted that this is true for *any* effect – there’s always the potential that an unobserved latent variable is creating a spurious effect;

This sort of work argues that the only solution is to use experimental (or, sometimes, propensity score style models)…but that’s simply not always feasible practically.

*We need to beware of making the best the enemy of the good enough...lest we make no progress at all...*
How to correct this problem?

• Essentially, this is an omitted variable problem, and my “solution” has been to identify as many potentially relevant alternative variables as I can find.

• The strongest possible correction is to use fixed-effects* models that control for all non-varying individual covariates. These have their own problems…

• Dual model for influence & selection.
  • Two-stage model “Heckman” and IV sorts of models
  • Dynamic SAOM models
  • Experiments

*“Adding fixed effects to dynamic panel models with many subjects and few repeat observations creates severe bias towards zero coefficients. This has been demonstrated both analytically (Nickell 1981) and through simulations (Nerlove 1971) for OLS and other regression models and has been well-known by social scientists, including economists, for a very long time. In fact, CCF even note that they do not add fixed effects to their logit regression model for this reason, but they strangely assert that fixed effects are necessary in the OLS model.” Estimating Peer Effects on Health in Social Networks : A Response to Cohen-Cole and Fletcher; Trogdon, Nonnemaker, Pais J.H. Fowler, PhD and N.A. Christakis, MD, PhD
Causal status of such similarity is hard to know,
- Identification strategies are *stringent*
- *My* sense is we’re over-correcting on this front; let’s figure out what’s there first.

Weak instruments bias us toward null effects
Possible solutions:

• Theory: Given what we know about how friendships form, is it reasonable to assume a bi-directional cause? That is, work through the meeting, socializing, etc. process and ask whether it makes sense that $Y$ is a cause of $W$. *This will not convince a skeptical reader, but you should do it anyway.*

• Models:
  - Time Order. Necessary but not sufficient. We are on somewhat firmer ground if $W$ precedes $Y$ in time, but the Shalizi & Thomas problem of an as-yet-earlier joint confounder is still there.
  
  - Simultaneous Models. Model *both* the friendship pattern and the outcome of interest simultaneously. Best bet for direct estimation

• Sensitivity Analysis:
  I think the most reasonable solution…take error potential seriously, attempt to evaluate how big a problem it really is.
### Table 4. Selected SIENA Parameter Estimates: Parental Knowledge, Parental Discipline, and Drinking

| Model 2 | Selection parameters | Alter effects: Who is more often named as a friend? | Ego effects: Who names more friends? | Similarity effects: Choosing friends similar to oneself | Behavioral parameters: Influence on Drinking |
|--------|----------------------|---------------------------------------------------|------------------------------------|------------------------------------------------|-------------------------------------------------
|        | Alter effects: Who is more often named as a friend? | Ego effects: Who names more friends? | Similarity effects: Choosing friends similar to oneself | Behavioral parameters: Influence on Drinking |
|        | Parameter | Model 2 | t | p | Parameter | Model 2 | t | p | Parameter | Model 2 | t | p | Parameter | Model 2 | t | p | Parameter | Model 2 | t | p |
|        |           | b | SE |        |           | b | SE |        |           | b | SE |        |           | b | SE |        |           | b | SE |        |           | b | SE |        |           | b | SE |        |           |
| Alter effects: Who is more often named as a friend? | Parental knowledge | -0.002 | 0.004 | -0.47 | 0.003 | Parental discipline | -0.004 | 0.002 | -1.55 | 0.001 | Drinking | 0.083 | 0.010 | 8.69 | ** | 0.007 |
| Ego effects: Who names more friends? | Parental knowledge | 0.044 | 0.007 | 5.85 | *** | 0.039 | Parental discipline | -0.002 | 0.005 | -0.36 | 0.023 | Drinking | -0.011 | 0.021 | -0.53 | 0.089 |
| Similarity effects: Choosing friends similar to oneself | Parental knowledge | 0.169 | 0.025 | 6.70 | *** | 0.101 | Parental discipline | 0.151 | 0.017 | 8.86 | *** | 0.035 | Drinking | 0.276 | 0.021 | 13.45 | *** | 0.006 |
| Behavioral parameters: Influence on Drinking | Friends' attributes | Mean Parental knowledge | -0.230 | 0.065 | -3.56 | ** | 0.014 | Mean Parental discipline | -0.051 | 0.043 | -1.18 | 0.011 | Drinking mean similarity | 1.162 | 0.110 | 10.56 | *** | 0.023 |
| Control variables (individual level) | Parental knowledge | -0.122 | 0.014 | -8.93 | *** | 0.004 | Parental discipline | -0.043 | 0.009 | -4.54 | *** | 0.002 |

***p < .001. **p < .01. *p < .05. †p < .10.

*a Models also include rate and shape parameters, structural parameters, and the full set of alter, ego, similarity, and individual-level control parameters.
Possible solutions:

- Sensitivity Analysis:
  
  I think the most reasonable solution...take error potential seriously, attempt to evaluate how big a problem it really is.

---

**Sensitivity Analysis for Contagion Effects in Social Networks**

Tyler J. VanderWeele

**Abstract**

Analyses of social network data have suggested that obesity, smoking, happiness, and loneliness all travel through social networks. Individuals exert “contagion effects” on one another through social ties and association. These analyses have come under critique because of the possibility that homophily from unmeasured factors may explain these statistical associations and because similar findings can be obtained when the same methodology is applied to height, acne, and headaches, for which the conclusion of contagion effects seems somewhat less plausible. The author uses sensitivity analysis techniques to assess the extent to which supposed contagion effects for obesity, smoking, happiness, and loneliness might be explained away by homophily or confounding and the extent to which the critique using analysis of data on height, acne, and headaches is relevant. Sensitivity analyses suggest that contagion effects for obesity and smoking cessation are reasonably robust to possible latent homophily or environmental confounding; those for happiness and loneliness are somewhat less so. Supposed effects for height, acne, and headaches are all easily explained away by latent homophily and confounding. The methodology that has been used in past studies for contagion effects in social networks, when used in conjunction with sensitivity analysis, may prove useful in establishing social influence for various behaviors and states. The sensitivity analysis approach can be used to
Network Diffusion & Peer Influence

Peer Influence & Health: Current Lit & Controversies

Possible solutions:

• Sensitivity Analysis:
  I think the most reasonable solution...take error potential seriously, attempt to evaluate how big a problem it really is.

See details of your sensitivity analysis on other tabs

Sociological Methods & Research 2000
Possible solutions:

• Sensitivity Analysis:
  I think the most reasonable solution...take error potential seriously, attempt to evaluate how big a problem it really is.

https://j michaelrosenberg.shinyapps.io/konfound-it/
Network Diffusion & Peer Influence

Peer Influence & Health: Current Lit & Controversies

Possible solutions:

Experimental designs may alleviate this problem

Exogenously decide peers via natural experiment…

PEER EFFECTS WITH RANDOM ASSIGNMENT: RESULTS FOR DARTMOUTH ROOMMATES*

BRUCE SACERDOTE

This paper uses a unique data set to measure peer effects among college roommates. Freshman year roommates and dormmates are randomly assigned at Dartmouth College. I find that peers have an impact on grade point average and on decisions to join social groups such as fraternities. Residential peer effects are markedly absent in other major life decisions such as choice of college major. Peer effects in GPA occur at the individual room level, whereas peer effects in fraternity membership occur both at the room level and the entire dorm level. Overall, the data provide strong evidence for the existence of peer effects in student outcomes.

I. Introduction

People have long believed that peer quality and behavior are among the most important determinants of student outcomes. This idea is expressed in the Coleman Report [1966], in Supreme Court decisions such as Brown versus Topeka Board of Education (1954), and in the findings of numerous researchers. Betts and Morell [1999] find that high school peer group characteristics affect undergraduate grade point average (GPA). Case and Katz [1991] find large peer effects on youth criminal behavior and drug use. In a summary of the developmental psychology literature, Harris [1998] claims that parental behavior has no direct effect on child outcomes and that peer effects are the only important environmental factors affecting outcomes. A rich literature on neighborhood effects including Jencks and Mayer [1990], Rosenbaum [1992], and Katz, Kling, and Liebman [2001] shows that neighborhood peers can have profound effects on both adults and children.
Network Diffusion & Peer Influence

Peer Influence & Health: Current Lit & Controversies

Possible solutions:

… or by developing an online world where you can define the ties

---

**The Spread of Behavior in an Online Social Network Experiment**

Damon Centola

How do social networks affect the spread of behavior? A popular hypothesis states that networks with many clustered ties and a high degree of separation will be less effective for behavioral diffusion than networks in which locally redundant ties are rewired to provide shortcuts across the social space. A competing hypothesis argues that when behaviors require social reinforcement, a network with more clustering may be more advantageous, even if the network as a whole has a larger diameter. I investigated the effects of network structure on diffusion by studying the spread of health behavior through artificially structured online communities. Individual adoption was much more likely when participants received social reinforcement from multiple neighbors in the social network. The behavior spread farther and faster across clustered-lattice networks than across corresponding random networks.

---

**Fig. 1.** Randomization of participants to clustered-lattice and random-network conditions in a single trial of this study ($N = 128$, $Z = 6$). In each condition, the black node shows the focal node of a neighborhood to which an individual is being assigned, and the red nodes correspond to that individual’s neighbors in the network. In the clustered-lattice network, the red nodes share neighbors with each other, whereas in the random network they do not. White nodes indicate individuals who are not connected to the focal node.