Economic Models for Social Interactions

Larry Blume

Cornell University & IHS & The Santa Fe Institute & HCEO

hceconomics.org

SSSI 2017
New Economic School

Introduction

- Role of theory
- Psychology vs sociology
- What are the consequences for received economics of embedding economic behavior in social life? Choice in economic tradition is determined by tastes, beliefs and constraints. If we are to pull this off we must map social interactions into these objects — social influence must effect these.
  - Non-cooperative game theory is a model of interactive beliefs.
  - Challenges for economic theory: How does one do welfare economics with endogenously determined, inderdependent tastes? In some cases this is not insurmountable. In evaluating a Nash equilibrium, we ask, are there belief profiles that would give everyone a higher expected payoff?

Social Life **and** Economics

- "The outstanding discovery of recent historical and anthropological research is that man's economy, as a rule, is submerged in his social relationships. He does not act so as to safeguard his individual interest in the possession of material goods; he acts so as to safeguard his social standing, his social claims, his social assets. He values material goods only in so far as they serve this end." (Polanyi, 1944)

- "Economics is all about how people make choices. Sociology is all about why they don't have any choices to make." (Duesenberry, 1960)

Social Life and Economics

- "The outstanding discovery of recent historical and anthropological research is that man's economy, as a rule, is submerged in his social relationships. He does not act so as to safeguard his individual interest in the possession of material goods; he acts so as to safeguard his social standing, his social claims, his social assets. He values material goods only in so far as they serve this end." (Polanyi, 1944)
- "Economics is all about how people make choices. Sociology is all about why they don't have any choices to make." (Duesenberry, 1960)

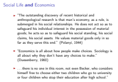
- Duesenberry describes a situation which prevailed in the 50s and 60s — sociologists and economists just didn't talk to each other. That changed in the 1980s, as economists became interested in what sociologists had observed about issues of common interest and sociologists became interested in economists' methods, rational actor models in particular.

- Duesenberry was in fact arguing for more of a sociological perspective. "…there is no one in this room, not even Becker, who considers himself free to choose either two children who go to university or four children who stop their education after high school." p. 233.

- The study of socially embedded economic behavior requires examining the co-dependence of social structures and individual behavior. One vehicle for doing that is "social interactions models", which usually are focused on social networks..

Social Life **and** Economics

- "The outstanding discovery of recent historical and anthropological research is that man's economy, as a rule, is submerged in his social relationships. He does not act so as to safeguard his individual interest in the possession of material goods; he acts so as to safeguard his social standing, his social claims, his social assets. He values material goods only in so far as they serve this end." (Polanyi, 1944)
- "Economics is all about how people make choices. Sociology is all about why they don't have any choices to make." (Duesenberry, 1960)
- ...there is no one in this room, not even Becker, who considers himself free to choose either two children who go to university or four children who stop their education after high school."



- Duesenberry describes a situation which prevailed in the 50s and 60s — sociologists and economists just didn't talk to each other. That changed in the 1980s, as economists became interested in what sociologists had observed about issues of common interest and sociologists became interested in economists' methods, rational actor models in particular.

- Duesenberry was in fact arguing for more of a sociological perspective. "...there is no one in this room, not even Becker, who considers himself free to choose either two children who go to university or four children who stop their education after high school." p. 233.

- The study of socially embedded economic behavior requires examining the co-dependence of social structures and individual behavior. One vehicle for doing that is "social interactions models", which usually are focused on social networks..

What are "Social Interactions"

- Social learning
- Social norms
- Peer effects
- Strategic complementarities that come from "non-economic activities"

Where do Social Interactions Appear?

Phenomena

- Labor markets
  - Career Choices
  - Retirement
- Fertility
- Health
- Education Outcomes
- Violence

Mechanisms

- Peer effects
  - Stigma
- Role models
- Social Norms
- Social Learning
- Social Capital?

Where do Social Interactions Appear?

Phenomena

- Labor markets
  - Career Choices
  - Retirement
- Fertility
- Health
- Education Outcomes
- Violence

Mechanisms

- Peer effects
  - Stigma
- Role models
- Social Norms
- Social Learning
- Social Capital?

- Job search — we'll come back to this. Random matching of firms and consumers gives a random bipartite graph.

- Wayne Baker on the social structure of even seemingly anonymous markets.

- Social capital - Nan Lin, the resources that are embedded in the ties of social networks.

Example I & II. Social Learning and Social Norms

	D	C
D	1	1
C	0	$1+a$

where $a > 0$.

Players from a large population are randomly matched to play this game. They learn from experience. For example, they believe outcomes are iid and have beta-distribution priors.

The play of the population will almost surely settle into one equilibrium. Which one?

Example I & II. Social Learning and Social Norms

	D	C
D	1	1
C	0	$1+a$

where $a > 0$.

Players from a large population are randomly matched to play this game. They learn from experience. For example, they believe outcomes are iid and have beta-distribution priors.

The play of the population will almost surely settle into one equilibrium. Which one?

- Beta distributions are conjugate priors for binomial draws.
- Answer: Both are reachable with positive probability.
- Interpret the equilibria as trust and no-trust.
- Notice the difference between this and reciprocal altruism.

Economics and Social Interactions

Example III. Peer effects

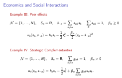
$$\mathcal{N} = \{1, \ldots, N\}, \quad S_n = \mathbf{R}, \quad \bar{s}_{-n} = \sum_{k \neq n} a_{nk} s_k, \quad \sum_k a_{nk} = 1, \quad \beta_n \geq 0$$

$$u_n(s_n, s_{-n}) = h_n s_n - \frac{1}{2} s_n^2 - \frac{\beta_n}{2} \left(s_n - \bar{s}_{-n} \right)^2.$$

Example IV. Strategic Complementarities

$$\mathcal{N} = \{1, \ldots, N\}, \quad S_n = \mathbf{R}, \quad \sum_{k \neq n} g_{nk} = 1, \quad \beta_n > 0$$

$$u_n(s_n, s_{-n}) = h_n s_n - \frac{1}{2} s_n^2 + \beta_n \sum_{k \neq n} g_{nk} s_n s_k.$$


Economics and Social Interactions

Example III. Peer effects

$$\mathcal{N} = \{1, \ldots, N\}, \quad S_n = \mathbf{R}, \quad \bar{s}_{-n} = \sum_{k \neq n} a_{nk} s_k, \quad \sum_k a_{nk} = 1, \quad \beta_n \geq 0$$

$$u_n(s_n, s_{-n}) = h_n s_n - \frac{1}{2} s_n^2 - \frac{\beta_n}{2} (s_n - \bar{s}_{-n})^2.$$

Example IV. Strategic Complementarities

$$\mathcal{N} = \{1, \ldots, N\}, \quad S_n = \mathbf{R}, \quad \sum_{k \neq n} g_{nk} = 1, \quad \beta_n > 0$$

$$u_n(s_n, s_{-n}) = h_n s_n - \frac{1}{2} s_n^2 + \beta_n \sum_{k \neq n} g_{nk} s_n s_k.$$

Expanding the peer effects payoff function,

$$u_n = h_n s_n - \frac{1+\beta}{2} s_n^2 + \beta_n \sum_{k \neq n} a_{nk} s_n s_k - \frac{\beta_n}{2} \bar{s}_{-n}$$

which is payoff-equivalent to

$$u_n' = \frac{h_n}{1+\beta_n} s_n - \frac{1}{2} s_n^2 + \frac{\beta_n}{1+\beta_n} \sum_{k \neq n} a_{nk} s_n s_k$$

In the peer effects model, equilibrium always exists. In the complementarities model equilibrium existence may not exist if $\beta_n > 1$.

Implications of Social Interactions

$$\mathcal{N} = [0,1] \text{ with measure } \mu, \qquad S_n = \mathbf{R}, \qquad \bar{s} = \int s(n)\, d\mu(n),$$

$$u(s, \bar{s}, n) = h(n)s - \frac{1}{2}s^2 - \frac{\beta}{2}(s - \bar{s})^2.$$

First-order conditions, for each n,

$$h(n) - s(n) - \beta\big(s(n) - \bar{s}\big) = 0$$

so

$$\bar{h} - \bar{s} = 0$$

and

$$s(n) = \frac{h(n) + \beta\bar{h}}{1 + \beta}.$$

Implications of Social Interactions

$$\mathcal{N} = [0,1] \text{ with measure } \mu, \qquad S_n = \mathbf{R}, \qquad \bar{s} = \int s(n) d\mu(n),$$

$$u(s, \bar{s}, n) = h(n)s - \frac{1}{2}s^2 - \frac{\beta}{2}(s - \bar{s})^2.$$

First-order conditions, for each n,

$$h(n) - s(n) - \beta(s(n) - \bar{s}) = 0$$

so

$$\bar{h} - \bar{s} = 0$$

and

$$s(n) = \frac{h(n) + \beta \bar{h}}{1 + \beta}.$$

- Mean does not depend on β, but variance s_n equals $\sigma_h^2/(1+\beta)^2$.

Implications of Social Interactions

$$\mathcal{N} = [0,1] \text{ with measure } \mu, \qquad S_n = \mathbf{R}, \qquad \bar{s} = \int s(n) d\mu(n),$$

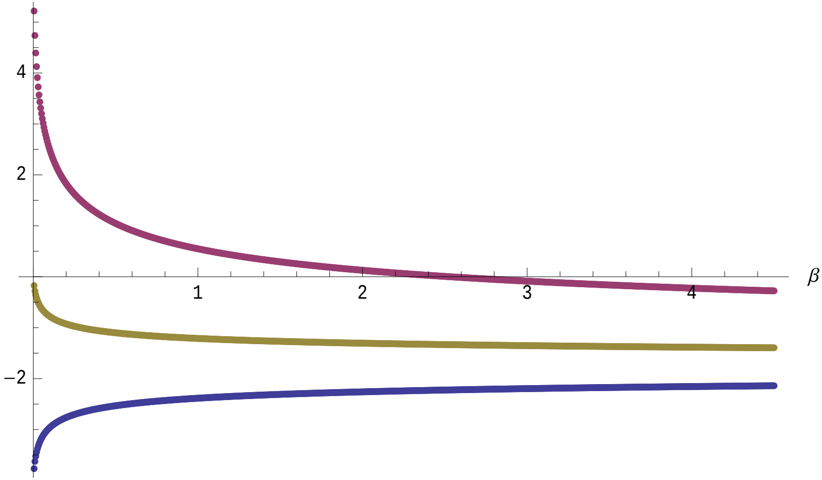
$$u(s, \bar{s}, n) = h(n)s - \frac{1}{2}s^2 - \frac{\beta}{4}(s - \bar{s})^4.$$

First-order conditions, for each n,

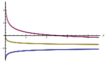
$$h(n) - s(n) - \beta\big(s(n) - \bar{s}\big)^3 = 0$$

$$\text{sign}\{\bar{h} - \bar{s}\} = \text{sign}\{E\{(s(n) - \bar{s})^3\}.$$

Implications of Social Interactions



$$\mu\{-4\} = 0.6, \quad \mu\{6\} = 0.4.$$



$\mu\{-4\} = 0.6, \quad \mu\{6\} = 0.4.$

- $E\{\bar{h} - \bar{s}\} > 0$,

Implications of Social Interactions

$$\mathcal{N} = [0,1] \text{ with measure } \mu, \quad S_h = \mathbf{R}, \quad \bar{s} = \int s(h) d\mu(h),$$

$$u(s, \bar{s}, h) = hs - \frac{1}{2}s^2 - \frac{\beta}{2}(s - \bar{s})^2.$$

Suppose individuals sort themselves into two groups and respond only to their group. For individual h in group g,

$$s(h) = \frac{h + \beta \bar{h}_g}{1 + \beta}.$$

$$u(\text{high}, h) - u(\text{low}, h) =$$

$$\frac{\beta}{2(1+\beta)}(2h - \bar{h}_{\text{high}} - \bar{h}_{\text{low}})(\bar{h}_{\text{high}} - \bar{h}_{\text{low}}).$$

Implications of Social Interactions

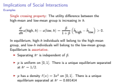
Examples

Single crossing property: The utility difference between the high-mean and low-mean group is increasing in h.

$$\frac{d}{dh}u(\text{high}, h) - u(\text{low}, h) = \frac{\beta}{1+\beta}\left(\bar{h}_{\text{high}} - \bar{h}_{\text{low}}\right) > 0.$$

In equilibrium, high-h individuals will belong to the high-mean group, and low-h individuals will belong to the low-mean group. Equilibrium is assortative.

- Separating h^* is independent of β.
- μ is uniform on $[0, 1]$. There is a unique equilibrium separated at $h^* = 1/2$.
- μ has a density $f(x) = 3x^2$ on $[0, 1]$. There is a unique equilibrium separated at $h^* = 0.691414$



- equilibrium is independent of β, and at the midpoint between the two means.
- Any $K > 2$-group equilibrium is *assortative*.
- The equilibrium separation condition is that $u(\text{high}) - u(\text{low}) = 0$.
- with cubed example, there will be multiple equilibrium.

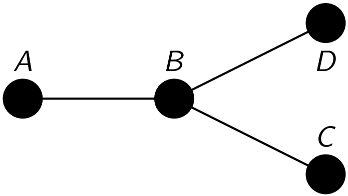
Plan

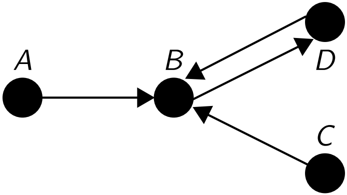
- Network Science
- Consequences of Social Networks
- Properties of Social Networks
- Labor Markets — Weak and Strong Ties
- Peer Effects and Complementarities — Games on Networks
- Matching and Network Formation
- Social Capital
- Social Learning
- Diffusion

Network Science

Graphs

A directed graph $\mathcal{G}$ is a pair (V, E) where V is a set of vertices, or nodes, and E is a set of Edges. In a directed graph, an edge is an ordered pair (v, w) of vertices, meaning that there is a connection **from** v **to** w. In an undirected graph, an edge is an unordered pair of vertices.





$$V = \{A, B, C, D\}$$
$$G = \{(A, B), (B, C), (B, D)\}$$

$$V = \{A, B, C, D\}$$
$$G = \{(A, B), (C, B), (B, D), (D, B)\}$$

Graphs

A directed graph $\mathcal{G}$ is a pair (V, E) where V is a set of vertices, or nodes, and E is a set of Edges. In a directed graph, an edge is an ordered pair (v, w) of vertices, meaning that there is a connection **from** v **to** w. In an undirected graph, an edge is an unordered pair of vertices.

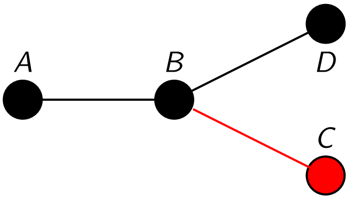
$V = \{A, B, C, D\}$
$G = \{(A, B), (B, C), (B, D)\}$

$V = \{A, B, C, D\}$
$G = \{(A, B), (C, B), (B, D), (D, B)\}$

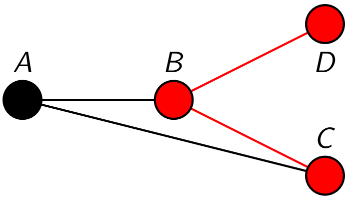
We talk only about undirected graphs from here on out.

The degree of a node in an undirected graph $\mathcal{G}$ is $\#\{w : (v, w) \in E\}$.

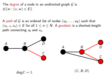
A path of $\mathcal{G}$ is an ordered list of nodes $(v_0, \ldots, v_N)$ such that $(v_{n-1}, v_n) \in E$ for all $1 \leq n \leq N$. A geodesic is a shortest-length path connecting v_0 and v_n.



$\deg C = 1.$



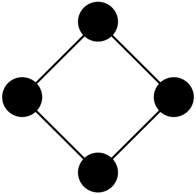
(C, B, D)

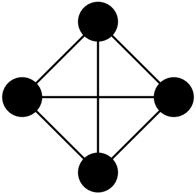


- Chapter 2 each of EK and J.
- Newman (2003)
- For directed graphs we talk about in-degree and out-degree.
- c, a, b, d and c, b, d are two paths from c to d. The latter is a geodesic. The distance from c to d is 2, not 3.

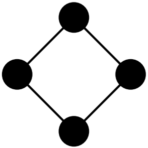
Graphs

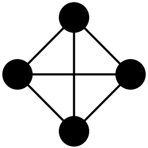
A subset of vertices is connected if there is a path between every two of them. A component of $\mathcal{G}$ is a set of vertices maximal with respect to connectedness. A clique is a component for which all possible edges are in E.





A graph $\mathcal{G}$ has a matrix representation. A adjacency matrix for a graph (V, E) is a $\#V \times \#V$ matrix A such that $a_{vw} = 1$ if $(v, w) \in E$, and 0 otherwise. A weighted adjacency matrix has non-zero numbers corresponding to edges in E.





$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

A graph $\mathcal{G}$ has a matrix representation. A adjacency matrix for a graph (V, E) is a $\#V \times \#V$ matrix A such that $a_{vw} = 1$ if $(v, w) \in E$, and 0 otherwise. A weighted adjacency matrix has non-zero numbers corresponding to edges in E.

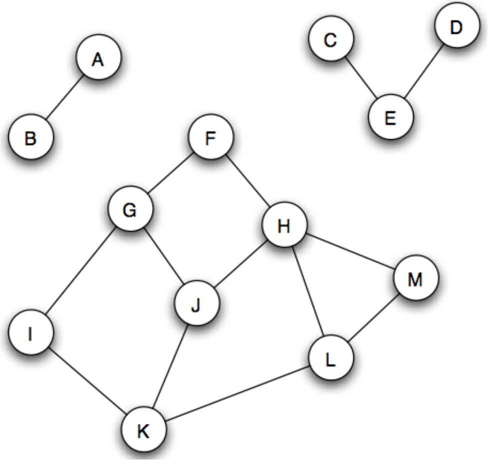
$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

- adjacency matrices aso called sociomatrices.
- Adjacency matrices of undirected graphs are symmetric; not so for directed graphs.

Common Network Measurements

- Graph diameter — maximal geodesic length.
- Mean geodesic length.
- Degree distribution.
- Clustering coefficient — the average (over vertices) of the number of length 2 paths containing i that are part of a triangle. (Measures degree of transitivity.)
- Component size distribution

Graphs



- 3 Components, $\{A, B\}$, $\{C, D, E\}$, $\{F, \ldots, M\}$.
- Min degree $= 1$.
- Max deg $= 4$.
- Diam Comp. $3 = 3$.
- Degree Dist. $1 : 4/13,\ 2 : 4/13,\ 3 : 4/13,\ 4 : 1/13$.
- Clustering coefficient: $1/15$.

Probabilistic Models of Graphs

Going beyond descriptive statistics of individual networks to inference about network properties requires probabilistic models of network structure.

- Having observed data from some networks, what can I infer about the properties of other networks?
- Having observed some data from a network, what can I infer about other properties of this network?

Two kinds of models

- Descriptive statistics: Stochastic block models, exponential random graphs
- Structural models: Models of network formation.
  - Algorithmic
  - Strategic

Probabilistic Models of Graphs

Going beyond descriptive statistics of individual networks to inference about network properties requires probabilistic models of network structure.

- Having observed data from some networks, what can I infer about the properties of other networks?
- Having observed some data from a network, what can I infer about other properties of this network?

Two kinds of models

- Descriptive statistics: Stochastic block models, exponential random graphs
- Structural models: Models of network formation.
  - Algorithmic
  - Strategic

- Bernoulli random graphs
- Preferential attachment

- pairwise and Nash stability

Stochastic Social Network Analysis

- Treat networks as realizations of variables
- Propose a model for the distribution of those variables
- Fit the model to some observed data
- With the learned model
  - Interpret the parameters to gain insight into the properties of the network
  - Use the model to predict properties of the network or of other networks

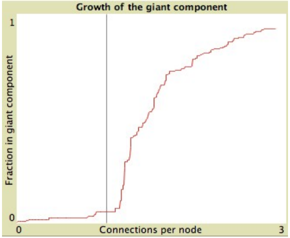
Erdös-Rényi Random Graphs

Undirected graph. Every pair of vertices is chosen as an edge independently with probability p.

Poisson random graphs: A sequence of graphs $\mathcal{G}_n$ with $|V_n| = n$ and p such that $p \cdot (n-1) \to z$.

Large n facts:

- Phase transition at $z = 1$.
- $z < 1$: Exponential component size distribution.
- $z > 1$: A giant connected component of size $O(n)$.
- Clustering coefficient is $C^2 = O(n^{-1})$.
- Poisson degree distribution with mean z.



Simulation of Erdös-Rényi random sets on 300 nodes.

Erdös-Rényi Random Graphs

Undirected graph. Every pair of vertices is chosen as an edge independently with probability p.

Poisson random graphs: A sequence of graphs $\mathcal{G}_n$ with $|V_n| = n$ and p such that $p \cdot (n-1) \to z$.

Large n facts:

- Phase transition at $z = 1$.
- $z < 1$: Exponential component size distribution.
- $z > 1$: A giant connected component of size $O(n)$.
- Clustering coefficient is $C^2 = O(n^{-1})$.
- Poisson degree distribution with mean z.

Simulation of Erdös-Rényi random sets on 300 nodes.

- Compare Erdös-Rényi/Poisson with social networks. Note differences, similarites

- Single giant component is very general, as is emergence as a critical phenomenon

- Preferential attachment, other random models designed to generate random graphs that "look like" random graphs.

- Brief discussion of generating interesting random graphs — exponential random graphs.

- There is a strategic theory of graph formation that we will discuss later.

Preferential Attachment

- A source of power laws.
- Introduced by Eggenberger and Polya (1923).
- Popularized by Zipf (1949) (city size) and Simon (1955) (wealth).

Preferential Attachment

A directed graph.

- A vertex set V of size N.
- For nodes $i > 1$, with probability p i links to a randomly chosen node $j < i$.
- With probability $(1 - p)$ i links to the immediate ancestor of j.

The graph is surely connected.

For large n the fraction of nodes with in-degree k is $1/k^r$ where r depends on p. The fraction P_r of vertices with r edges converges as N gets large, and $P_r = \Theta(r^{-\frac{2-p}{1-p}})$. See Kumar et al. (2000).

Preferential Attachment

A directed graph.

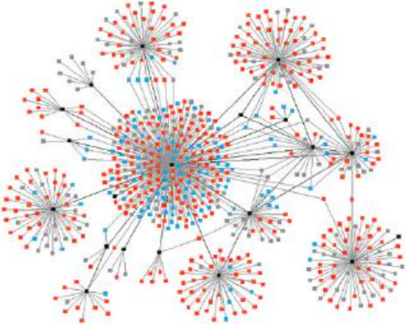
- A vertex set V of size N.
- For nodes $i > 1$, with probability p i links to a randomly chosen node $j < i$.
- With probability $(1-p)$ i links to the immediate ancestor of j.

The graph is surely connected.

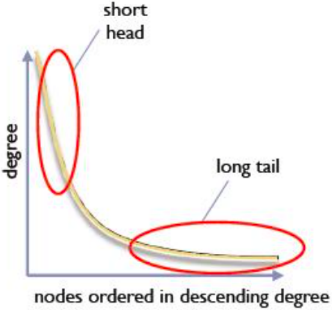
For large n the fraction of nodes with in-degree k is $1/k^r$ where r depends on p. The fraction P_r of vertices with r edges converges as N gets large, and $P_r = \Theta(r^{-\frac{2-p}{1-p}})$. See Kumar et al. (2000).

- A simple version: Every vertex links to only 1 other node.
- Called the copy model, introduced by Kumar et al. (2000).
- A graphical version of the Yule-Simon process: An urn process in which balls are added to a growing number of urns, each ball being allocated to an urn with probability linear in the number the urn already contains.

Preferential Attachment



Example of network with preferential attachment



Sketch of long-tailed degree distribution

Consequences of Social Networks

This section is an interlude - to remind us of economic questions before going on with theory of networks.

Macro Crime Statistics

Glaeser Sacerdote and Scheinkman (1996)

We believe that the most puzzling aspect of crime is not its overall level, or that level's relationship with the overall quantity of deterrence. Rather, . . . , we believe that the most inexplicable aspect of crime is its large variance across time and space.

If agents' decisions are independent, then city crime levels represent averages of large numbers of independent decisions. Elementary statistics tells us that these averages should be free of the effects of random idiosyncratic error terms and they should be close to the expected population mean.

However, even casual empiricism suggests that differences in observable local area characteristics can account for little of the variation in crime rates across cities in the U.S. or across precincts in New York City.

Macro Crime Statistics
Glaeser Sacerdote and Scheinkman (1996)

We believe that the most puzzling aspect of crime is not its overall level, or that level's relationship with the overall quantity of deterrence. Rather, ..., we believe that the most inexplicable aspect of crime is its large variance across time and space.

If agents' decisions are independent, then city crime levels represent averages of large numbers of independent decisions. Elementary statistics tells us that these averages should be free of the effects of random idiosyncratic error terms and they should be close to the expected population mean.

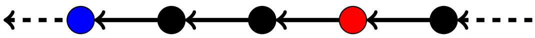
However, even casual empiricism suggests that differences in observable local area characteristics can account for little of the variation in crime rates across cities in the U.S. or across precincts in New York City.

- In the iid world, variation in crime rates across cities is driven by variation in city characteristics.

- Aggregate behavior of social systems is determined by how individuals are linked.

A Model (of sorts)

- $2N + 1$ individuals live on the integer lattice at points $-N, \ldots, N$.
- Type 0s never commit a crime; Type 1's always do; Type 2's imitate the neighbor to the right.
- Type of individual i is p_i.



A Model (of sorts)

- Expected distance between fixed agents determines group size — range of interaction effects.
- Social interactions magnify the effect of fixed agents.

$$E\{a_i\} = \frac{p_1}{p_0 + p_1} \equiv p, \qquad S_n = \sum_{|i| \le n} \frac{a_i - p}{2n + 1}.$$

$$\sqrt{2n+1} S_n \to N(0, \sigma^2), \qquad \sigma^2 = p(1-p)\frac{2-\pi}{\pi}$$

where $\pi = p_0 + p_1$.

A Model (of sorts)

- Expected distance between fixed agents determines group size — range of interaction effects.
- Social interactions magnify the effect of fixed agents.

$$E\{a_i\} = \frac{p_1}{p_0 + p_1} = p, \qquad S_n = \sum_{|i| \le n} \frac{a_i - p}{2n+1}.$$

$$\sqrt{2n+1}S_n \to N(0, \sigma^2), \qquad \sigma^2 = p(1-p)\frac{2-\pi}{\pi}$$

where $\pi = p_0 + p_1$.

- The smaller the fraction of committed types, the more the conformity and the larger the variance.

Adoption of a New Technology
Conley and Udry (2010)

- The adoption of new technology is a central feature of the transformation of farming systems during the process of economic development.
- How do farmers learn about a new technology?
  - Farmer's own experimentation.
  - Extension service, media.
  - Social learning, from neighbors' experiments.

Adoption of a New Technology
Conley and Udry (2010)

- The adoption of new technology is a central feature of the transformation of farming systems during the process of economic development.
- How do farmers learn about a new technology?
  - Farmer's own experimentation.
  - Extension service, media.
  - Social learning, from neighbors' experiments.

Over the past decade, in part of Ghana's Eastern region, an established system of maize and cassava intercropping for sale to urban consumers has begun to be replaced by intensive production of pineapple for export to European markets. An important component of this transformation is the adoption of agricultural chemicals that were not used in the previous farming system. In this paper, we consider how farmers in this area might learn about the appropriate use of fertilizer in this new farming system.

Adoption of a New Technology

A basic model: Besley and Case (1993); Foster and Rosenzweig (1995); Munshi (2004)

- A village is a learning unit.
- Some farmers experiment, others do not.
- Each farmer in the village observes the farming activities of each of the other farmers.
- Each farmer then updates his or her own opinion regarding the technology.
- Each farmer makes decisions regarding cultivation for the next season.

⋮

Adoption of a New Technology

A basic model: Besley and Case (1993); Foster and Rosenzweig (1995); Munshi (2004)

- A village is a learning unit.
- Some farmers experiment, others do not.
- Each farmer in the village observes the farming activities of each of the other farmers.
- Each farmer then updates his or her own opinion regarding the technology.
- Each farmer makes decisions regarding cultivation for the next season.

⋮

Fails against the data

Adoption of a New Technology

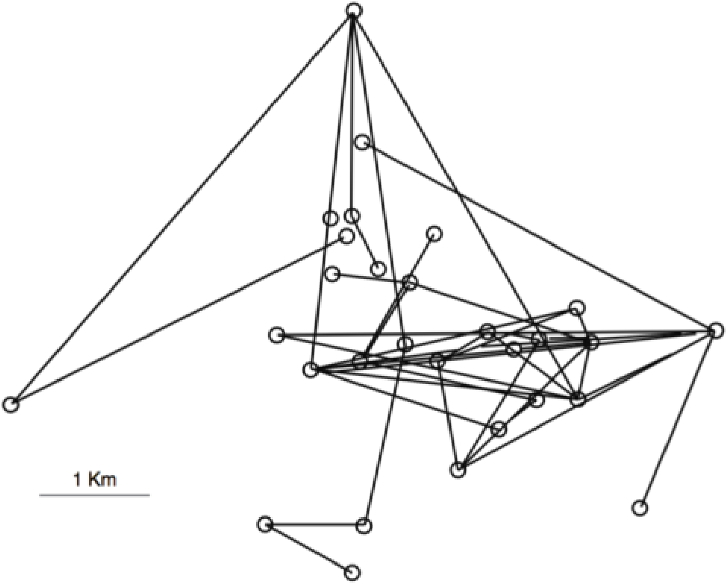
- A survey was conducted of approximately 450 individuals in four clusters of villages in Ghana's Eastern Region over a period of 21 months in 1996-1998. Two aspects of the data are relevant here.
  - Plot level data on inputs and outputs at frequent intervals from the respondents.
  - a variety of data on farmer interactions was collected. For example, data was collected on respondents' knowledge of inputs and outputs on the plots of other respondents and on respondents' conversations about farming (and specifically about fertilizer) with other farmers.

Adoption of a New Technology

- Each respondent was matched randomly with 10 other farmers in his/her village.
- In only 11 percent of these matches had one of the two individuals ever received advice about farming from the other.
- In 30 percent of the matches, the respondent indicates that he **could** approach the other farmer for advice about fertilizer.
- Respondents are able to provide some information on harvests and inputs used on approximately 7 percent of random matches between respondents and pineapple plots cultivated by other farmers in the village.

Information flows through a sparse social network.

Adoption of a New Technology



Adoption of a New Technology

- Fertilizer is used at time t to produce output at time $t + 1$.
- Production is subject to a random shock. Shocks are iid draws from a common and unknown distribution.
- Farmers are Bayesians.
- Consider both limited communication and limited information.
- Contrast pineapple production with a known technology.

Adoption of a New Technology

- Fertilizer is used at time t to produce output at time $t+1$.
- Production is subject to a random shock. Shocks are iid draws from a common and unknown distribution.
- Farmers are Bayesians.
- Consider both limited communication and limited information.
- Contrast pineapple production with a known technology.

Why is learning hard? Think about the updating rule.

Adoption of a New Technology

We find that farmers are more likely to change input levels upon the receipt of bad news about the profitability of their previous level of input use, and less likely to change when they observe bad news about the profitability of alternative levels of inputs. Farmers tend to increase (decrease) input use when an information neighbor achieves higher than expected profits when using more (less) inputs than they previously used. This holds when controlling for correlations in growing conditions, for common credit shocks using a notion of financial neighborhoods, and across several information metrics. Support for the interpretation of our results as indicating learning is provided by the fact that it is novice farmers who are most responsive to news in their information neighborhoods. Additional support is provided by our finding no evidence of learning when our methodology is applied to a known maize-cassava technology.

Properties of Social Networks

Some Social Networks

Network	Type	n	m	z	ℓ	α	$C^{(1)}$	$C^{(2)}$	r
film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.20	0.78	0.208
company directors	undirected	7 673	55 392	14.44	4.60	–	0.59	0.88	0.276
math coauthorship	undirected	253 339	496 489	3.92	7.57	–	0.15	0.34	0.120
physics coauthorship	undirected	52 909	245 300	9.27	6.19	–	0.45	0.56	0.363
biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	–	0.088	0.60	0.127
telephone call graph	undirected	47 000 000	80 000 000	3.16		2.1			
email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0		0.16	
email address books	directed	16 881	57 029	3.38	5.22	–	0.17	0.13	0.092
student relationships	undirected	573	477	1.66	16.01	–	0.005	0.001	−0.029
sexual contacts	undirected	2 810				3.2			

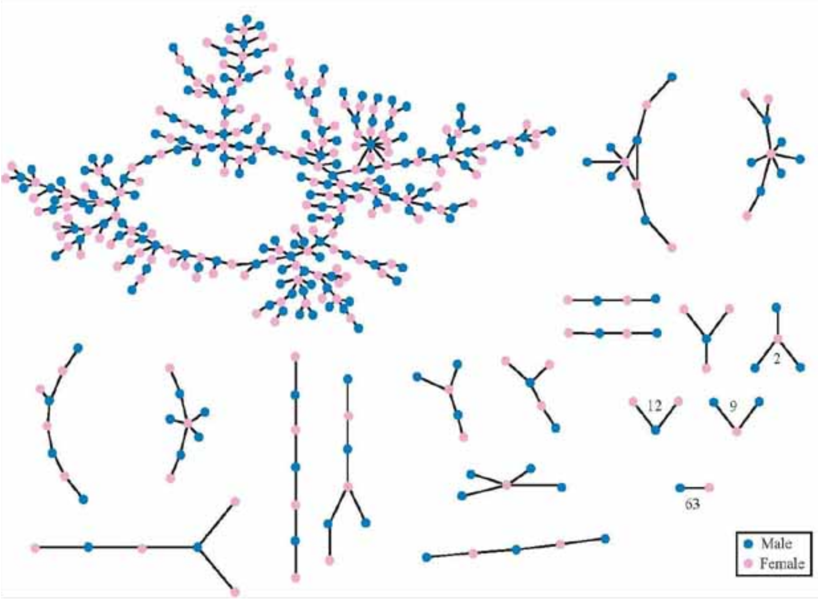
n – # nodes, m – # edges, z – mean degree,
l – mean geodesic length, α – exponent of degree dist.,
$C^{(k)}$ - clustering coeff.s, r degree corr. coeff.

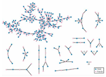
Network	Type	n	m	z	l	α	$C^{(1)}$	$C^{(2)}$	r
film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.20	0.78	0.208
company directors	undirected	7 673	55 392	14.44	4.60	–	0.59	0.88	0.276
math coauthorship	undirected	253 339	496 489	3.92	7.57	–	0.15	0.34	0.120
physics coauthorship	undirected	52 909	245 300	9.27	6.19	–	0.45	0.56	0.363
biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	–	0.088	0.60	0.127
telephone call graph	undirected	47 000 000	80 000 000	3.16		2.1			
email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0		0.16	
email address books	directed	16 881	57 029	3.38	5.22	–	0.17	0.13	0.092
student relationships	undirected	573	477	1.66	16.01	–	0.005	0.001	−0.029
sexual contacts	undirected	2 810				3.2			

n – # nodes, m – # edges, z – mean degree,
l – mean geodesic length, α – exponent of degree dist.,
$C^{(k)}$ - clustering coeff.s, r degree corr. coeff.

- Source: Newman (2003).
- film: Amaral et al. (2000) and Watts and Strogatz (1998). corporate directors: Davis, Yoo, and Baker (2003). sexual relations: Liljeros, Edling, and Amaral (2003).
- Notice how small l is relative to n or m.

Some Social Networks I

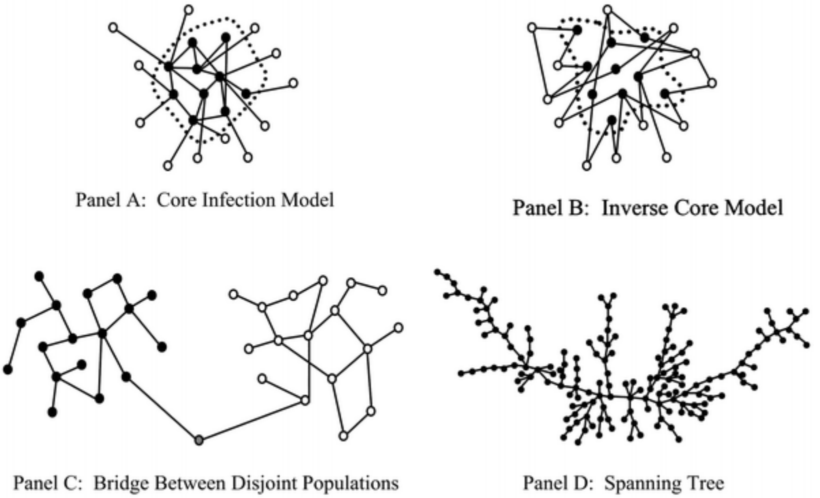


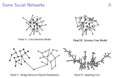


- Bearman, Moody, and Stovel (2004)

- “While this large component involves the vast majority of individuals with multiple partners, it has numerous short branches. Further, it is very broad: the two most distant individuals are 37 steps apart. Most surprising, it is characterized by the almost complete absence of short cycles. Thus the network closely approximates a chainlike spanning tree.”

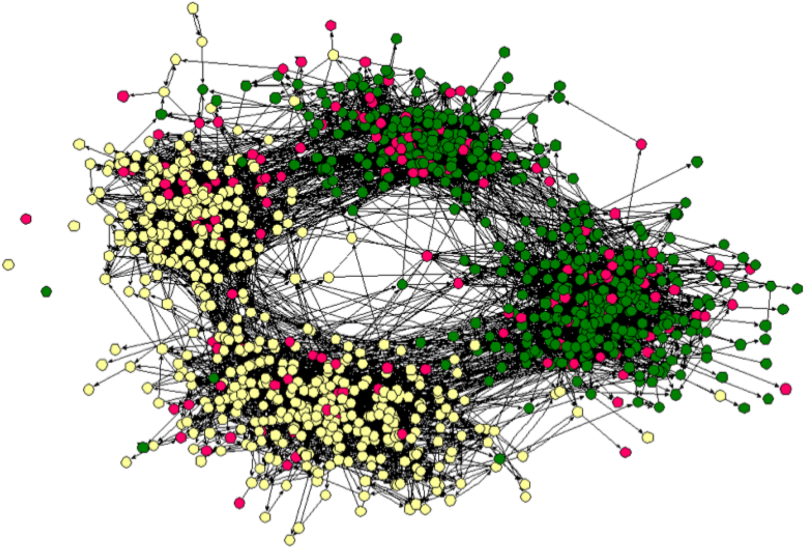
Some Social Networks II

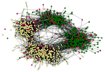




- No one dates their ex's partner's ex. No 4-cycles.
- They simulated very similar-looking graphs with a simple model where the probability of a tie is higher for two individuals with the similar amounts of sexual experience, plus a no-four-cycle rule.
- Under core and inverse core structures, it matters enormously which actors are reached, while under a spanning tree structure the key is not so much which actors are reached, just that some are. This is because given the dynamic tendency for unconnected dyads and triads to attach to the main component, the structure is equally sensitive to a break (failure to transmit disease) at any site in the graph. In this way, relatively low levels of behavior change even by low-risk actors, who are perhaps the easiest to influence can easily break a spanning tree network into small disconnected components, thereby fragmenting the epidemic and radically limiting its scope.

Some Social Networks III





- Moody (2001)

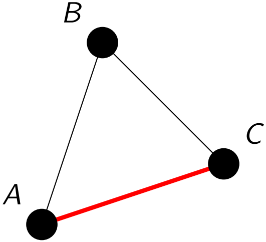
- Moody's "Countryside" High School, friendship network, illustrates homophily. Stratification by age (vertical) and race (ethnicity).

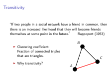
- Some Properties of Social Networks

Transitivity

"If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in the future." Rappoport (1953)

- Clustering coefficient: Fraction of connected triples that are triangles.
- Why transitivity?





Why transitivity?

- **Constraints:** Opportunity to meet
- **Tastes:**
  - homophily — our friends' friends are like our friends, and thus like us.
  - incentives — reducing latent stress; an old idea.

Centrality

Types of Centrality Measures:

Degree Centrality How many vertices can a vertex reach directly?

Betweeness Centrality How likely is this vertex to be on the geodesic between two randomly chosen vertices?

Closeness Centrality How fast can this vertex reach all vertices in the network.

Eigenvector Centrality How much does this vertex influence other important vertices?

Which nodes are important?

Let A be a weighted adjacency matrix for a directed graph. $A_{ij} > 0$ if j influences i. e is the vector of 1's. $\lambda > 0$ is the Perron eigenvalue of A.

- Degree Centrality: How many nodes can a node directly influence?

'

$$c_j^d = \sum_i A_{ij} \qquad c^d = eA$$

Katz (1953) Centrality: How many nodes can a node reach?

$$c_j^k(\alpha) = \sum_i \left(\sum_{k>0} \alpha^k A^k \right)_{ij} \qquad c^k(\alpha) = e(I - \alpha A)^{-1} - e.$$

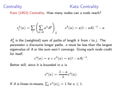
A_{ij}^k is the (weighted) sum of paths of length k from i to j. The parameter α discounts longer paths. α must be less than the largest eigenvalue of A or the sum won't converge. Giving each node credit for itself,

$$c^a(\alpha) = e + c^k(\alpha) = e(I - \alpha A)^{-1}.$$

Better still, since it is bounded in α is

$$c^\alpha(\alpha) = \frac{\lambda - \alpha}{\lambda} c^a(\alpha).$$

If A is linear-in-means, $\sum_j c^\alpha(\alpha)_j = 1$ for $\alpha \leq 1$.



- Why divide by λ? Because as $\alpha \uparrow \lambda$, c^α converges to $\liminf_T \frac{1}{T+1} \sum_0^T A^k$.

Eigenvector Centrality: The centrality of j is proportional to the sum of the centralities of the nodes she influences.

$$c_j^e = \mu \sum_i c_i a_{ij} \qquad c^e = \mu c^e A$$

where $\mu > 0$ and $c^e \geq 0$.. If the network is strongly connected, then (Perron Frobenius Theorem) there is a unique scalar μ and a one- dimensional set of vectors $c \gg 0$ that solve this. μ is the inverse of the Perron eigenvalue, and c is in the corresponding left eigenspace. (Bonacich, 1987; Bonacich and Lloyd, 2001).

It is not necessary, but useful, to choose from the positive half-eigenspace the vector whose components sum to 1, that is, of l_1-norm 1.

Suppose A is indecomposable. Assume the Perron eigenvalue of A is 1 (e.g. lim). Let $C^a(\alpha) = diag\ c^a(\alpha)$. Let $C^e = diag\ c^e$

$$(1-\alpha)C^a(\alpha) = (1-\alpha)\sum_{k\geq 0}\alpha^k A^k,$$

$$(1-\alpha)C^a(\alpha) - C^e = (1-\alpha)\sum_{k\geq 0}\alpha^k(A^k - C^e) \to_{\alpha\uparrow 1} 0,$$

so

$$\lim_{\alpha\to 1}(1-\alpha)c^a(\alpha) = \lim_{\alpha\to 1}(1-\alpha)eC^a(\alpha) = eC^e = c^e.$$

More generally, with Perron eigenvalue $\lambda > 0$,

$$\lim_{\alpha\to\lambda}(\frac{\lambda-\alpha}{\lambda})c^a(\alpha) = \lim_{\alpha\to\lambda}(\frac{\lambda-\alpha}{\lambda})eC^a(\alpha) = eC^e = c^e.$$

Finally, $c^d = \lim_{\alpha\to 0} c^k(\alpha)$.

Two sources of centrality:

- Who you are connected to.
- What you 'bring to the table'.

$$\begin{aligned} c^{\alpha}(\alpha, d) &= \alpha c^{\alpha}(\alpha, d)A + d \\ &= d(I - \alpha A)^{-1} \\ &= d(I + \alpha A + \alpha^2 A^2 + \cdots) \end{aligned}$$

α-centrality takes $d = e$:

$$c^{\alpha}(\alpha) = c^{\alpha}(\alpha, e).$$

A quadratic game in which each player is influenced by the average play of his neighbors.

$$u_i(x_i, x_{-i}) = h_i x_i - \frac{x_i^2}{2} - \frac{\beta}{2}(x_i - \bar{x}_i)^2, \qquad \bar{x}_i = \sum_j a_{ij} x_j.$$

The equilibrium is unique:

$$x = (1-\phi)(I - \phi A)^{-1} h, \qquad \phi = \beta/1 + \beta.$$

Average play in the population is

$$\begin{aligned} \frac{1}{n} e \cdot x &= \frac{1}{n}(1-\phi) e \left(I - \phi A\right)^{-1} h \\ &= \frac{1}{n}(1-\phi) c^a(\phi) h. \end{aligned}$$

Individual i's influence on the average choice of the population is proportional to $c^a(\phi)$.

A quadratic game in which each player is influenced by the average play of his neighbors.

$$u_i(x_i, x_{-i}) = h_i x_i - \frac{x_i^2}{2} - \frac{\beta}{2}(x_i - \bar{x}_i)^2, \qquad \bar{x}_i = \sum_j a_{ij} x_j.$$

The equilibrium is unique:

$$x = (1-\phi)(I - \phi A)^{-1} h, \qquad \phi = \beta/1+\beta.$$

Average play in the population is

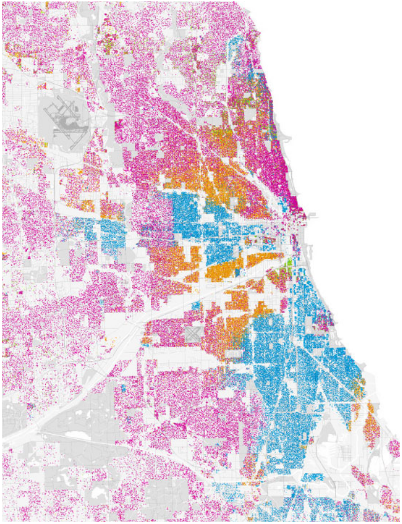
$$\frac{1}{n} e \cdot x = \frac{1}{n}(1-\phi) e\,(I - \phi A)^{-1}\, h = \frac{1}{n}(1-\phi) c^{\alpha}(\phi) h.$$

Individual i's influence on the average choice of the population is proportional to $c^{\alpha}(\phi)$.

- Weights sum to 1.

- When β is small, social influence is weak, ϕ is small and so only short paths matter. When social influence is strong, ϕ is near 1 and long paths matter most.

Homophily



"Similarity begets friendships."

Plato

"All things akin and like are for the most part pleasant to each other, as man to man, horse to horse, youth to youth. This is the origin of the proverbs: The old have charms for the old, the young for the young, like to like, beast knows beast, ever jackdaw to jackdaw, and all similar sayings." Aristotle, *Nicomachean Ethics*

Sources of Homophily

- Status Homophily: We feel more comfortable when we interact with others who share a similar cultural background.
- Value Homophily: We often feel justified in our opinions when we are surrounded by others who share the same opinions.
- Opportunity Homophily: We mostly meet people like us.

Sources of Homophily

- **Status Homophily**: We feel more comfortable when we interact with others who share a similar cultural background.
- **Value Homophily**: We often feel justified in our opinions when we are surrounded by others who share the same opinions.
- **Opportunity Homophily**: We mostly meet people like us.

- Sociologist use descriptive language to talk about their concepts. To make use of them we must map them into our economic models.

- Lazarsfeld and Merton (1954). Value and status homophily have to do with tastes. Patterns of interaction are driven by preferences for similarity. Opportunity homophily is about constraints.

- white = pink   black = blue

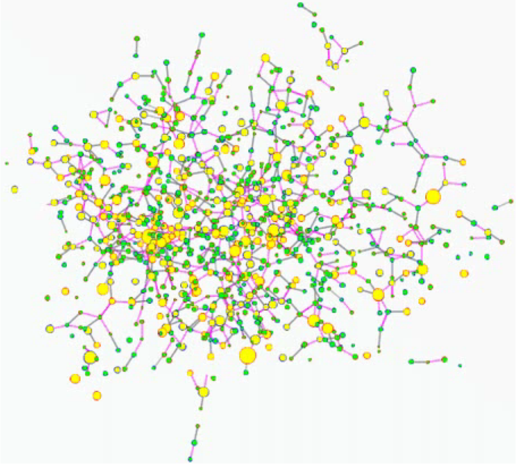
- asian = green   hispanic = orange

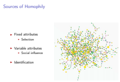
- other = grey

- Measuring homophily — if nodes can be typed, compare edges to what random mixing would do.

Sources of Homophily

- Fixed attributes
  - Selection
- Variable attributes
  - Social influence
- Identification





- Christakis and Fowler (2007)
- purple lines — genetic link. Size proportional to bmi, kg/m^2. Yellow — obese. Red rim — female.
- note clustering. Dynamics are whole network grew, clustering increased, whole network got heavier.
- Why clustering?
  - selection — people choose their friends to match their weight.
  - homophily on an unobserved attribute which is highly correlated with obesity.
  - change in friends' weight affects your own view of weight.
  - Much is wrong with this work. Cohen-Cole and Fletcher (2008)

Labor Markets

Job Search

VOL. 81 NO. 5 MONTGOMERY: SOCIAL NETWORKS AND LABOR-MARKET OUTCOMES 1409

TABLE 1—JOB-FINDING METHODS USED BY WORKERS

Source/data	Percentage of jobs found using each method					
	Friends/relatives	Gate application	Employment agency	Ads	Other	Sample size
Myers and Shultz (1951)/sample of displaced textile workers:						
First job	62	23	6	2	7	144
Mill job	56	37	3	2	2	144
Present job	36	14	4	0	46[a]	144
Rees and Shultz (1970)/Chicago labor-market study, 12 occupations:[b]						
Typist	37.3	5.5	34.7	16.4	6.1	343
Keypunch operator	35.3	10.7	13.2	21.4	19.4	280
Accountant	23.5	6.4	25.9	26.4	17.8	170
Tab operator	37.9	3.2	22.2	22.2	14.5	126
Material handler	73.8	6.9	8.1	3.8	7.4	286
Janitor	65.5	13.1	7.3	4.8	9.3	246
Janitress	63.6	7.5	5.2	11.2	12.5	80
Fork-lift operator	66.7	7.9	4.7	7.5	13.2	175
Punch-press operator	65.4	5.9	7.7	15.0	6.0	133
Truck driver	56.8	14.9	1.5	1.5	25.3	67
Maintenance electrician	57.4	17.1	3.2	11.7	10.6	129
Tool and die maker	53.6	18.2	1.5	17.3	9.4	127
Granovetter (1974)/sample of residents of Newton, MA:						
Professional	56.1	18.2	15.9[c]	—[c]	9.8	132
Technical	43.5	24.6	30.4	—	1.4	69
Managerial	65.4	14.8	13.6	—	6.2	81
Corcoran et al. (1980)/Panel Study of Income Dynamics, 11th wave:						
White males	52.0	—[d]	5.8	9.4	33.8[d]	1,499
White females	47.1	—	5.8	14.2	33.1	988
Black males	58.5	—	7.0	6.9	37.6	667
Black females	43.0	—	15.2	11.0	30.8	605

[a]Most of these workers were rehired at a previous mill job or hired at a new mill established in one of the vacated mills.
[b]In computing the percentages, workers rehired by previous employers and those not reporting the job-finding source are excluded from the denominator and subtracted from the sample size.
[c]Agencies and ads are combined under the heading "formal means."
[d]Gate applications are included under "other."

The Strength of Weak Ties

"...[T]he strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie. Each of these is somewhat independent of the other, though the set is obviously highly intracorrelated. Discussion of operational measures of and weights attaching to each of the four elements is postponed to future empirical studies. It is sufficient for the present purpose if most of us can agree, on a rough intuitive basis, whether a given tie is strong, weak, or absent."

Granovetter (1973, p. 1361)

"...[T]he strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie. Each of these is somewhat independent of the other, though the set is obviously highly intracorrelated. Discussion of operational measures of and weights attaching to each of the four elements is postponed to future empirical studies. It is sufficient for the present purpose if most of us can agree, on a rough intuitive basis, whether a given tie is strong, weak, or absent."

Granovetter (1973, p. 1361)

- Sample of professional/managerial job changers in Newton MA. Asked those who found the new job through a contact how often they saw the contact around that time. Also asked where the informant got the job info.
  - Often 16.7%, occasionally 55.6%, 27.8% rarely.
  - Path length 39.1%, 45.3% one link, 12.5% two links, 3.1% more than 2.

- Valery Yakubovich (2005) uses a large scale survey of hires made in 1998 in a major Russian metropolitan area and finds that a worker is more likely to find a job through weak ties than through strong ones.

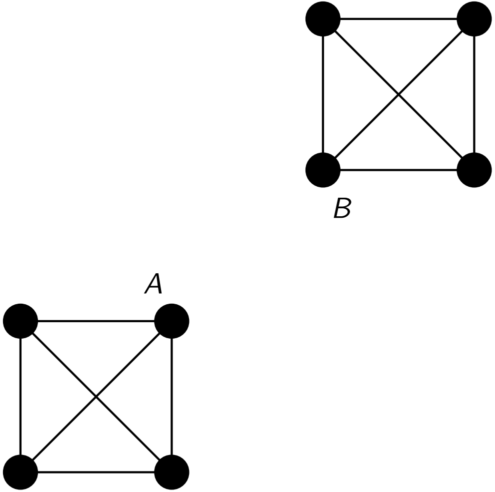
"...[T]he strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie. Each of these is somewhat independent of the other, though the set is obviously highly intracorrelated. Discussion of operational measures of and weights attaching to each of the four elements is postponed to future empirical studies. It is sufficient for the present purpose if most of us can agree, on a rough intuitive basis, whether a given tie is strong, weak, or absent."

Granovetter (1973, p. 1361)

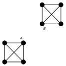
- Rapoport and Horvath (1961) ask each student in a Michigan middle school $N = 851$ to list their eight best friends in order. They take random samples of 9 students, and trace out who can be reached through first and second choices; same for second and third; ...; seventh and eighth. The smallest number of strudents are reached by 1 and 2; the most by 7 and 8.

Why do Weak Ties Matter? I

Two cliques.



Two cliques.



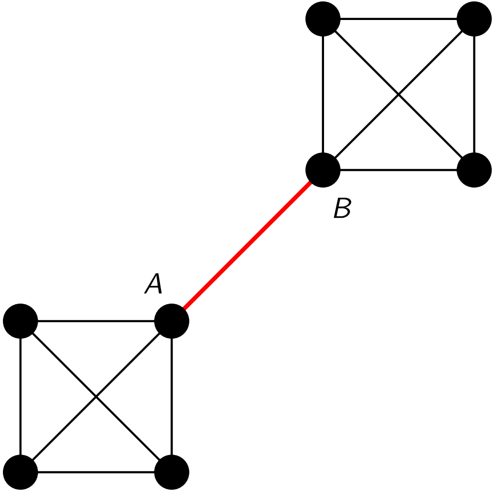
- The hypothesis which enables us to relate dyadic ties to larger structures is: the stronger the tie between A and B, the larger the propor- tion of individuals in S to whom they will both be tied, that is, connected by a weak or strong tie. This overlap in their friendship circles is predicted to be least when their tie is absent, most when it is strong, and intermediate when it is weak.

- it is sufficient for my purpose in this paper to say that the triad which is most unlikely to occur, under the hypothesis stated above, is that in which A and B are strongly linked, A has a strong tie to some friend C, but the tie between C and B is absent. (Granovetter, 1973, p. 1363).

Why do Weak Ties Matter? I

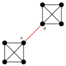
Two cliques.

A–*B* is a **bridge**.



Why do Weak Ties Matter? I

Two cliques.
A–*B* is a **bridge**.



- The hypothesis which enables us to relate dyadic ties to larger structures is: the stronger the tie between A and B, the larger the propor- tion of individuals in S to whom they will both be tied, that is, connected by a weak or strong tie. This overlap in their friendship circles is predicted to be least when their tie is absent, most when it is strong, and intermediate when it is weak.

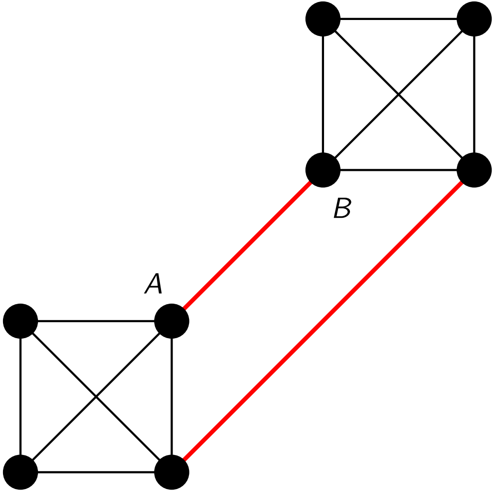
- it is sufficient for my purpose in this paper to say that the triad which is most unlikely to occur, under the hypothesis stated above, is that in which A and B are strongly linked, A has a strong tie to some friend C, but the tie between C and B is absent. (Granovetter, 1973, p. 1363).

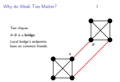
Why do Weak Ties Matter? I

Two cliques.

A–*B* is a **bridge**.

Local bridge's endpoints have no common friends.





- The hypothesis which enables us to relate dyadic ties to larger structures is: the stronger the tie between A and B, the larger the propor- tion of individuals in S to whom they will both be tied, that is, connected by a weak or strong tie. This overlap in their friendship circles is predicted to be least when their tie is absent, most when it is strong, and intermediate when it is weak.

- it is sufficient for my purpose in this paper to say that the triad which is most unlikely to occur, under the hypothesis stated above, is that in which A and B are strongly linked, A has a strong tie to some friend C, but the tie between C and B is absent. (Granovetter, 1973, p. 1363).

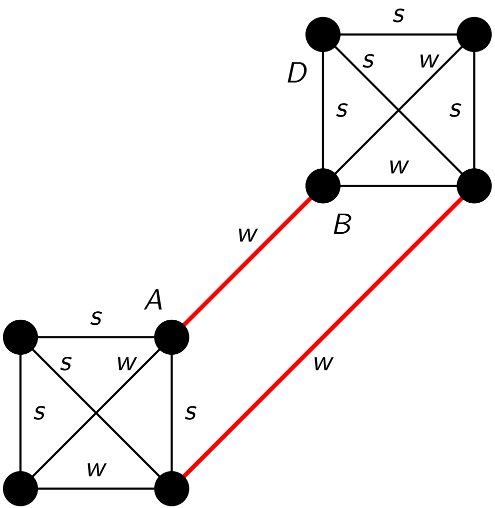
Why do Weak Ties Matter? I

Two cliques.

A–*B* is a **bridge**.

Local bridge's endpoints have no common friends.

Triadic closure: A length-2 path containining only strong edges is a closed triad.

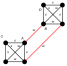


Two cliques.

A–B is a **bridge**.

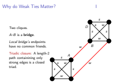
Local bridge's endpoints have no common friends.

Triadic closure: A length-2 path containining only strong edges is a closed triad.

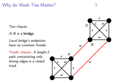


- The hypothesis which enables us to relate dyadic ties to larger structures is: the stronger the tie between A and B, the larger the propor- tion of individuals in S to whom they will both be tied, that is, connected by a weak or strong tie. This overlap in their friendship circles is predicted to be least when their tie is absent, most when it is strong, and intermediate when it is weak.

- it is sufficient for my purpose in this paper to say that the triad which is most unlikely to occur, under the hypothesis stated above, is that in which A and B are strongly linked, A has a strong tie to some friend C, but the tie between C and B is absent. (Granovetter, 1973, p. 1363).



- Two tightly-connected groups. Notice that the two components are cliques. Why? Triadic closure.
- The A–B link is a **bridge**. A bridge, if cut, creates two components.
- More generally, there may be a few paths that connect two components. A–B is a *local bridge* if its endpoints have no friends in common.
- Label edges with types.
- Triadic closure property: If a length-2 path contains only strong edges, then the triad is closed with some kind of edge.
- Triadic closure implies that any local bridge to a node having another tie which is strong, is a weak tie. If A–B were strong, then triadic closure implies A–D exists.



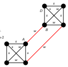
If homophilic ties tend to be strong, then local bridges tend to be weak.

Two cliques.

A–B is a **bridge**.

Local bridge's endpoints have no common friends.

Triadic closure: A length-2 path containining only strong edges is a closed triad.



- Granovetter: Weakly tied nodes are less similar, supply different information.

- Weak ties are less costly to maintain. More of them, so a greater information flow.

Ties and Inequality

Montgomery (1991)

- Workers live for two periods, $\#W$ identical in both periods.
- Half of the workers are high-ability, produce 1.
- Half of the workers are low-ability, produce 0.
- Workers are observationally indistinguishable.

- Each firm employs 1 worker.
- $\pi =$ employee productivity $-$ wage.
- Free entry, risk-neutral entrepreneurs.

- Equilibrium condition: Firms expected profit is 0. Wage offers are expected productivity.

Ties and Inequality I

Montgomery (1991)

- Workers live for two periods, $\#W$ identical in both periods.
- Half of the workers are high-ability, produce 1.
- Half of the workers are low-ability, produce 0.
- Workers are observationally indistinguishable.

- Each firm employs 1 worker.
- π = employee productivity − wage.
- Free entry, risk-neutral entrepreneurs.

- Equilibrium condition: Firms expected profit is 0. Wage offers are expected productivity.

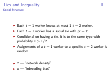
- Montgomery (1991)
- Point of the model is to demonstrate the aggregate effect — emergent effect — of social structure.
- Output price fixed at 1.
- With no structure, wage offers are $1/2$.

Ties and Inequality II

Social Structure

- Each $t = 1$ worker knows at most 1 $t = 2$ worker.
- Each $t = 1$ worker has a *social tie* with pr $= \tau$.
- Conditional on having a tie, it is to the same type with probability $\alpha > 1/2$.
- Assignments of a $t = 1$ worker to a specific $t = 2$ worker is random.

- τ — "network density"
- α — "inbreeding bias"



- Social structure
- Some $t = 2$ workers may have several ties, others none.
- Some have only one tie, and this is a source of market power for the firms. This is the subtlest point.

Ties and Inequality III

Timing

- Firms hire period 1 workers through the anonymous market, clears at wage w_{m1}.
- Production occures. Each firm learns its worker's productivity.
- Firm f sets a referral offer, w_{rf}, for a second period worker.
- Social ties are assigned.
- $t = 1$ workers with ties relay w_{ri}.
- $t = 2$ workers decide either to accept an offer or enter the market.
- Period 2 market clears at wage w_{m2}.
- Production occurs

- Only firms with 1-workers will make referral offers.
- Referral wages offers are distributed on an interval $[w_{m2}, w_R]$.
- $0 < w_{m2} < 1/2$.
- $\pi_2 > 0$.
- $w_{m1} = E\{\text{ production value} + \text{referral value }\} > 1/2$.

Ties and Inequality IV

Equilibrium

- Only firms with 1-workers will make referral offers.
- Referral wages offers are distributed on an interval $[w_{m2}, w_R]$.
- $0 < w_{m2} < 1/2$.
- $\pi_2 > 0$.
- $w_{m1} = E\{\text{ production value} + \text{referral value }\} > 1/2$.

- Why is there not a deterministic equilibrium? Because the probability that a $t = 2$ worker receives exactly one offer is strictly between 0 and 1.
- w_R has to be solved for.
- $w_{m2} < 1/2$ because 1-workers are stripped off by referrals at a higher rate than 0-workers. It is not 0 because some high types don't get offers.
- Because of imperfect competition in the referral market, expected $\pi_2 > 0$.
- $w_{m1} = E\{\text{production value} + \text{referral value}\}$ and $\pi_2 > 0$ for high types. Workers are valued for who they may know.

Ties and Inequality

V

Comparative Statics

$$\alpha, \tau \uparrow \quad \Longrightarrow \quad \begin{cases} w_{m2} \downarrow \\ w_R \uparrow \\ \pi_2 \uparrow \\ w_{m1} \uparrow \end{cases}$$

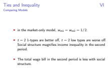
$$a, \tau \uparrow \quad \longrightarrow \quad \begin{cases} w_{m2} \downarrow \\ w_R \uparrow \\ \pi_2 \uparrow \\ w_{m1} \uparrow \end{cases}$$

- Suggestion of increased wage dispersion. See below.
- $w_{m2} \downarrow$ because market pool is worse.
- $w_R \uparrow$ because of increased competition for referred students.
- $\pi_2 \uparrow$ because either more referrals or higher quality referrals.
- $w_{m1} \uparrow$ because $\pi_2 \uparrow$.

Ties and Inequality VI

Comparing Models

- in the market-only model, $w_{m1} = w_{m2} = 1/2$.
- $t = 2$ 1-types are better off, $t = 2$ low types are worse off. Social structure magnifies income inequality in the second period.
- The total wage bill in the second period is less with social structure.



- Compare social structure, no social structure.
- A better analysis would do a steady state, or look at dynamics of an overlapping generations model.
- Degree of inequality should be positive, but less. Wage bill still goes down.

Peer Effects and Complementarities

Behaviors on Networks

Three Types of Network Effects

- Information and social learning.
- Network externalities.
- Social norms.

- Student learn from his classmates, so those who exert extra effort create a positive externality for their friends — complementarity, a network effect.

- A disruptive student takes up teacher's time and slows learning for everyone else — a complementarity

- Student wants to keep up with his peers who do well. This is a social norm effect.

- Student hangs out with those who drink, shoplift, . . . Another social norm effect.

A Common Regression

$$\omega_i = \pi_0 + x_i \pi_1 + \bar{x}_g \pi_2 + y_g \pi_3 + \varepsilon_i$$

Where

- ω_i is a choice variable for an individual,
- x_i is a vector of individual correlates,
- $\bar{x}_g$ is a vector of group averages of individual correlates,
- y_g is a vector of other group effects, and
- ε_i is an unobserved individual effect.

A Common Regression

$$\omega_i = \pi_0 + x_i\pi_1 + \bar{x}_g\pi_2 + y_g\pi_3 + \varepsilon_i$$

Where

- ω_i is a choice variable for an individual,
- x_i is a vector of individual correlates,
- $\bar{x}_g$ is a vector of group averages of individual correlates,
- y_g is a vector of other group effects, and
- ε_i is an unobserved individual effect.

- Datcher (1982)
- The LIM is the most common class of models in the empirical social interactions literature. We want to reverse-engineer it. Not the endpoint.
- Is $\pi_2 \neq 0$ evidence of an endogenous or peer effect? Can't address this without a **structural model**.

For all $g \in G$ and all $i \in g$,

$$\omega_i = \alpha + \beta x_i + \delta x_g + \gamma \mu_i + \varepsilon_i \quad \text{(Behavior)}$$

$$x_g = \frac{1}{N_g} x_i \quad \text{(Behavior)}$$

$$\mu_i = \frac{1}{N_g} \sum_{j \in g} \mathrm{E}\{\omega_j\} \quad \text{(Equilibrium)}$$

The reduced form is

$$\omega_i = \frac{\alpha}{1-\gamma} + \beta x_i + \frac{\gamma\beta + \delta}{1-\gamma} x_g + \varepsilon_i$$

For all $g \in G$ and all $i \in g$,

$$\omega_i = \alpha + \beta x_i + \delta x_g + \gamma \mu_i + \varepsilon_i \qquad \text{(Behavior)}$$

$$x_g = \frac{1}{N_g} x_i \qquad \text{(Behavior)}$$

$$\mu_i = \frac{1}{N_g} \sum_{j \in g} \mathrm{E}\{\omega_j\} \qquad \text{(Equilibrium)}$$

The reduced form is

$$\omega_i = \frac{\alpha}{1-\gamma} + \beta x_i + \frac{\gamma\beta + \delta}{1-\gamma} x_g + \varepsilon_i$$

- The identification problem here is clear. Can't determine δ, γ

General Linear Network Model

$$\omega_i = \beta' x_i + \delta' \sum_j c_{ij} x_j + \gamma' \sum_j a_{ij} \,\mathrm{E}\left\{\omega_j | x\right\} + \eta_i$$

This is the general linear model

$$\Gamma \omega + \Delta x = \eta.$$

Question:

- How do we interpret the parameters?
- What kind of restrictions on the coefficients are reasonable, and do they lead to identification.

These questions require a theoretical foundation.

Incomplete-Information Game

- I individuals; each i described by a type vector $(x_i, z_i) \in \mathbf{R}^2$.
  x_i is publicly observable, z_i is private.
- There is a Harsanyi prior ρ on the space of types $\mathbf{R}^{2I}$.
- Actions are $\omega_i \in \mathbf{R}$.
- Payoff functions:

$$U_i(\omega_i, \omega_{-i}; x, z_i) = \theta_i \omega_i - \frac{1}{2}\omega_i^2 - \frac{\phi}{2}\left(\omega_i - \sum_j a_{ij}\,\omega_j\right)^2$$

- a_{ij} — peer effect of j on i.

Incomplete-Information Game

- I individuals; each i described by a type vector $(x_i, z_i) \in \mathbf{R}^2$. x_i is publicly observable, z_i is private.
- There is a Harsanyi prior ρ on the space of types $\mathbf{R}^{2I}$.
- Actions are $\omega_i \in \mathbf{R}$.
- Payoff functions:

$$U_i(\omega_i, \omega_{-i}; x, z_i) = \theta_i \omega_i - \frac{1}{2}\omega_i^2 - \frac{\phi}{2}\left(\omega_i - \sum_j a_{ij}\omega_j\right)^2$$

- a_{ij} — peer effect of j on i.

Think about this as students choosing effort in school.

- The extremely difficult task is to disentangle neighborhood effects from peer effects and there is no consensus on the importance of peer effects on own achievement in this literature (see the recent literature surveys by Durlauf (2004) and Ioannides and Topa (2010).
- There is an important literature on peer effects in education (for a survey, see Sacerdote (2011)).
- The models in this section come from Blume et al. (2013a).

Private Component

To complete the model, specify how individual characteristics matter.

$$\theta_i = \gamma x_i + \delta \sum_j c_{ij} x_j + z$$

Direct Effect

Contextual Effect

c_{ij} — contextual/direct effect of j on i.

Private Component

To complete the model, specify how individual characteristics matter.

$$\theta_i = \gamma x_i + \delta \sum_j c_{ij} x_j + z$$

Direct Effect Contextual Effect

c_{ij} — contextual/direct effect of j on i.

Explain contextual effects.

Equilibrium

$$(1+\phi)\left(I - \frac{\phi}{1+\phi}A\right)\omega - (\gamma I + \delta C)x = \eta$$

$$\Gamma\omega + \Delta x = \eta.$$

Constraints imposed by the theory:

$$a_{ii} = c_{ii} = 0, \quad \sum_j a_{ij} = \sum_j c_{ij} = 1.$$

$$\Gamma_{ii} = 1+\phi, \quad \sum_{j\neq i} \Gamma_{ij} = -\phi, \quad \Delta_{ii} = -(\gamma+\delta), \quad \sum_{j\neq i} \Delta_{ij} = \delta.$$

Even more constraints if you insist on $A = C$.

When is the first equation identified?

- Order condition: $\#\{j \not\sim_C 1\} + \#\{j \not\sim_A 1\} \geq N - 1$.
- For each (γ, δ) pair there is a generic set of C-matrices such that the rank condition is satisfied.
- If two individuals' exclusions satisfy the order condition, there is a generic set of C-matrices such that the rank condition is satisfied for all γ and δ.

Non-Linear Aggregators

Bad apple The worst student does enormous harm.

Shining light A single student with sterling outcomes can inspire all others to raise their achievement.

Invidious comparison Outcomes are harmed by the presence of better achieving peers.

Boutique A student will have higher achievement whenever she is surrounded by peer with similar characteristics.

Non-Linear Aggregators

Bad apple The worst student does enormous harm.

Shining light A single student with sterling outcomes can inspire all others to raise their achievement.

Invidious comparison Outcomes are harmed by the presence of better achieving peers.

Boutique A student will have higher achievement whenever she is surrounded by peer with similar characteristics.

- Hoxby and Weingarth (2005)
- bad apple: the most relevant peer effects are those provided by the least academically able or least disciplined student in the class- room. This student provides large negative externalities in several possible ways: The bad apple peer may cause so much commotion in the classroom as to distract the teacher and students from productive tasks. Or he may encourage additional raucous or disruptive behavior among other students. Or the bad apple may not be a discipline problem but he may simply have low ability and require extra teaching attention, thereby detracting from the experience of the other students.
- This is essentially a model in which students do best when the environment is made to cater to their type. For instance, in schools, the Boutique model might mean that teachers organize lessons and materials around the learning style of a student if there is a critical mass of his type.

Matching and Network Formation

- Market Design
- Matching problems are models of network formation
  - Bipartite matching with transferable utility
  - Bipartite matching without exchange
  - Generalization to networks

Stable Matches

Given are two sets of objects X and Y. e.g. workers and firms. Both sides have preferences over whom they are matched with, but with no externalities, that is, given that a is matched with x, he does not care if b is matched with y and z. The literature divides over the information parties have when they choose partners, and whether compensating transfers can be made. The organizing principle is that of a stable match.

Assume w.l.o.g. $|X| \leq |Y|$.

Definition: A match is one-to-one map from X to Y. A match is stable if there are no pairs $x \leftrightarrow y$ and $x' \leftrightarrow y'$ such that $y' \succ_x y$ and $x \succ_{y'} x'$.

Find the optimal match by maximizing total surplus:

$$\begin{aligned} v(L \cup F) = \max_x \quad & \sum_{l,f} v_{lf} x_{lf} \\ \text{s.t.} \quad & \sum_f x_{lf} \leq 1 \quad \text{for all } l, \\ & \sum_l x_{lf} \leq 1 \quad \text{for all } f, \\ & x \geq 0 \end{aligned}$$

The vertices for this problem are integer solutions, that is, non-fractional matches. A solution to the primal is an optimal matching.

Find the optimal match by maximizing total surplus:

$$
\begin{aligned}
v(L \cup F) = \max_x \quad & \sum_{l,f} v_{lf} x_{lf} \\
\text{s.t.} \quad & \sum_f x_{lf} \leq 1 \quad \text{for all } l, \\
& \sum_l x_{lf} \leq 1 \quad \text{for all } f, \\
& x \geq 0
\end{aligned}
$$

The vertices for this problem are integer solutions, that is, non-fractional matches. A solution to the primal is an optimal matching.

- What would a social planner do? Maximize output?

- What would a market equilibrium look like?
  - The wages and profit distribution must be feasible.
  - No worker wants to leave his job for another manager who will prefer hire him.
  - No manager wants to hire a worker who would prefer to work for her.

Set of laborers L and firms F. v_{lf} is the value or surplus generated by matching worker l and firm f.

The surplus of a match is split between the firm and worker. Suppose $i \leftrightarrow f$ and $j \leftrightarrow g$. Payments to each are w_i and w_j, and π_i and π_j.

Since this is a division of the surplus,

$$w_i + \pi_f = v_{if} \quad \text{and} \quad w_j + \pi_g = v_{jg}.$$

If $w_i + \pi_g < v_{ig}$, then there is a split of the surplus v_{ig} such that i and g would both prefer to match with each other than with their current partners. The match is not stable. Stability requires

$$w_i + \pi_g \geq v_{ig} \quad \text{and} \quad w_j + \pi_f \geq v_{jf}.$$

Matching with Transferable Utility

The dual has variables for each individual and firm.

$$\begin{aligned} \min_{w,\pi} \quad & \sum_{l,f} w_l + \pi_f \\ \text{s.t.} \quad & \pi_f + w_l \geq v_{lf} \quad \text{for all pairs } l, f, \\ & \pi \geq 0,\ w \geq 0. \end{aligned}$$

Solutions to the dual satisfy the stability condition.

Complementary slackness says that matched laborer-firm pairs split the surplus, $\pi_f + w_l = v_{lf}$.

Characterizing Matches

Theorem: A matching is stable if and only if it is optimal.

Lemma: Each laborer with a positive payoff in any stable outcome is matched in every stable matching.

Proof: Complementary slackness.

Lemma: If laborer l is matched to firm f at stable matching x, and there is another stable matching x' which l likes more, then f likes it less.

Proof: Formalize this as follows: If x is a stable matching and $\langle w', \pi' \rangle$ is another stable payoff, then $w' > w$ implies $\pi > \pi'$. This follows from complementary slackness, since $w_l + \pi_f = v_{lf} = w'_l + \pi'_f$.

Suppose X and Y are each partially-ordered sets, and $v : X \times Y \to \mathbf{R}$ is a function.

Definition: $v : X \times Y \to \mathbf{R}$ has increasing differences iff $x' > x$ and $y' > y$ implies that

$$v(x', y') + v(x, y) \geq v(x', y) + v(x, y').$$

An important special case is where X and Y are intervals of $\mathbf{R}$, each with the usual order, and v is C^2.

$$v(x', y') - v(x, y') \geq v(x', y) - v(x, y).$$

Then

$$D_x v(x, y') \geq D_x v(x, y)$$

From this it follows that $D_{xy} v(x, y) \geq 0$.

Generalizations

- Matching without exchange. Gale and Shapley (1962).
- The roommate problem.
- Generalization of non-transferable matching to networks. Jackson and Wolinsky (1996).

The Rooommate Problem

Let $\delta(v)$ denote the edge set of vertex v. $E(S)$ is the set of edges with endpoints in S.

$$
\begin{aligned}
v = \max_x \quad & \sum_{e \in E} v_e x_e \\
\text{s.t.} \quad & \sum_{e \in \delta(v)} x_e \leq 1 \quad \text{for all } v, \\
& x \in \{0, 1\}.
\end{aligned}
$$

The Rooommate Problem

Let $\delta(v)$ denote the edge set of vertex v. $E(S)$ is the set of edges with endpoints in S.

$$
\begin{aligned}
v = \max_x \quad & \sum_{e \in E} v_e x_e \\
\text{s.t.} \quad & \sum_{e \in \delta(v)} x_e \leq 1 \quad \text{for all } v, \\
& x \in \{0, 1\}.
\end{aligned}
$$

$$
\begin{aligned}
\text{s.t.} \quad & \sum_{e \in \delta(v)} x_e \leq 1 \quad \text{for all } v, \\
& \sum_{e \in \mathrm{E}\{(\}S)} x_e \leq k \quad \text{for all } S \text{ s.t. } |S| = 2k + 1.
\end{aligned}
$$

Network Formation with Contagious Risk

Blume et al. (2013b)

A set V of N agents form no more than Δ bilateral relationships with each other, thereby constructing a graph $G = (V, E)$. Each agent receives payoff $a > 0$ from each of her links.

Then, cascades occur. Each node fails independently with probably q. Each failed node transmits failure to her neighbors with independent probability p, and so on. The edges that transmit, and the nodes they connect are the live-edge subgraph.

A falied agent loses all benefits and pays a cost b.

$$\pi_i = ad_i(1 - \phi_i) - b\phi_i$$

where d_i is the degree of agent i and ϕ_i is the probability i fails.

Network Formation with Contagious Risk

Rawlsian welfare — minimum welfare among all agents.

Definition: A graph is stable if:

- no node can strictly increase its payoff by deleting all its incident links (hence removing itself from the network), and
- there is no pair of unconnected nodes (i, j) such that adding an (i, j) edge to G would make them both better off.

Rawlsian welfare — minimum welfare among all agents.

Definition: A graph is stable if:

- no node can strictly increase its payoff by deleting all its incident links (hence removing itself from the network), and
- there is no pair of unconnected nodes (i, j) such that adding an (i, j) edge to G would make them both better off.

Stability is a superset of pairwise Nash stability which allows agents to drop any subset of links.

Assumptions

- $a > pqb$.
- $a < pb$.
- $a < qb$.

We want the bounds to hold very loosely. "Separation parameter" δ:

Assumption $\mathcal{P}(\delta)$: There is a small constant δ such that

$$\delta^{-1} pqb < a < \delta \min\{pb, qb\}.$$

Assumptions

- $a > pqb$.
- $a < pb$.
- $a < qb$.

We want the bounds to hold very loosely. "Separation parameter" δ:

Assumption $\mathcal{P}(\delta)$: There is a small constant δ such that

$$\delta^{-1} pqb < a < \delta \min\{pb, qb\}.$$

- We imagine p and q very small.
- N is very large.
- Imagine 2 nodes i and j, and i is deciding whether to link to j.
  - The probability that j fails and infects i is pq, so if $a < pqb$ no links will form.
  - If i knows that j will fail, he is protected only by transmission failure. We don't want this link to form, otherwise everyone will link as much as they can. So $a < pb$.
  - Analogously, we want no links to form if transmission is certain.
- Spreading parameter — interesting idea.

- Results provide asymptotically tight characterizations of the welfare obtained by both socially optimal and stable graphs.
- If each node forms more than $1/p$ links, the live-edge subgraph has a giant connected component.
- "..., we find roughly that social optimality occurs just beyond the edge of a phase transition that controls how failures propagate, while stable graphs lie slightly further still past this phase transition, at a point where most of the welfare has already been wiped out."

Social Capital

Networks and Social Capital

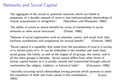
"the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance or recognition." (Bourdieux and Wacquant, 1992)

"the ability of actors to secure benefits by virtue of membership in social networks or other social structures." (Portes, 1998)

"features of social organization such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit." (Putnam, 1995)

"Social capital is a capability that arises from the prevalence of trust in a society or in certain parts of it. It can be embodied in the smallest and most basic social group, the family, as well as the largest of all groups, the nation, and in all the other groups in between. Social capital differs from other forms of human capital insofar as it is usually created and transmitted through cultural mechanisms like religion, tradition, or historical habit." (Fukuyama, 1996)

"naturally occurring social relationships among persons which promote or assist the acquisition of skills and traits valued in the marketplace..." (Loury, 1992)



Bourdieux and Wacquant (1992) An Invitation to Reflexive Sociology

Portes (1998) “Social capital: its origins and applications in modern sociology.”

Putnam (1995) “Bowling alone: America’s declining social capital.”

Loury (1992)“The economics of discrimination: Getting to the core of the problem.” Harvard Journal for African American Public Policy 1: 91 – 110.

Networks and Social Capital

"the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance or recognition." (Bourdieux and Wacquant, 1992)

"the ability of actors to secure benefits by virtue of membership in social networks or other social structures." (Portes, 1998)

"features of social organization such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit." (Putnam, 1995)

"Social capital is a capability that arises from the prevalence of trust in a society or in certain parts of it. It can be embodied in the smallest and most basic social group, the family, as well as the largest of all groups, the nation, and in all the other groups in between. Social capital differs from other forms of human capital insofar as it is usually created and transmitted through cultural mechanisms like religion, tradition, or historical habit." (Fukuyama, 1996)

"naturally occurring social relationships among persons which promote or assist the acquisition of skills and traits valued in the marketplace..." (Loury, 1992)

- Putnam is bad because it presumes an effect. Social capital can't always be good. ex. Medieval Church and Crusades.

- "Social capital can be defined simply as the existence of a certain set of informal rules or norms shared among members of a group that permits cooperation among them." Fukuyama goes on: "The sharing of values and norms does not in itself produce social capital, because the values may be the wrong ones... The norms that produce social capital... must substantively include virtues like truth-telling, the meeting of obligations, and reciprocity." What are the criteria for the right norms?

Networks and Social Capital

"...social capital may be defined operationally as *resources embedded in social networks and accessed and used by actors for actions.* Thus, the concept has two important components: (1) it represents resources embedded in social relations rather than individuals, and (2) access and use of such resources reside with actors."

(Lin, 2001)

"...social capital may be defined operationally as *resources embedded in social networks and accessed and used by actors for actions*. Thus, the concept has two important components: (1) it represents resources embedded in social relations rather than individuals, and (2) access and use of such resources reside with actors."

(Lin, 2001)

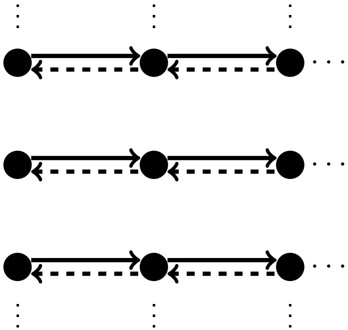
- social capital generates externalities for members of a group;
- these externalities are achieved through shared trust, norms, and values and their consequent effects on expectations and behavior;
- shared trust, norms, and values arise from informal forms of organizations based on social networks and associations. The study of social capital is that of network-based processes that generate beneficial outcomes through norms and trust.

Information

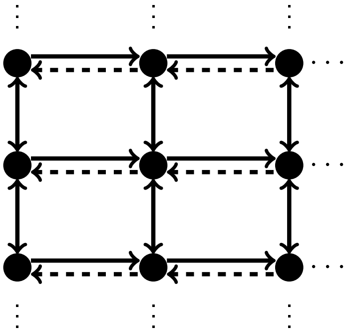
- Search is a classic example according to Lin's (2001) definition.
- Search has nothing to do with values and social norms beyond the willingness to pass on a piece of information.

Intergenerational Transfers

Loury (1981)



Only Intergenerational Transfers



Intergenerational Transfers with Redistribution

Intergenerational Transfers
Loury (1981)


Only Intergenerational Transfers


Intergenerational Transfers with Redistribution

- Two-period overlapping generations model; parents invest in children (education).
- Solid lines are resource flows; dashed lines are utility flows.

x	output
α	ability, realized in adults.
e	investment
c	consumption
y	income
$h(\alpha, e)$	production function
$U(c, V)$	parent's utility

$$c + e = y \quad \text{parental budget constraint}$$

Assumptions:

A.1. U is strictly monotone, strictly concave, C^2, Inada condition at the origin. $\gamma < U_v < 1 - \gamma$ for some $0 < \gamma < 1$.

A.2 h is strictly increasing, strictly concave in e, C^1, $h(0,0) = 0$ and $h(0,e) < e$. $h_\alpha \geq \beta > 0$. For some $\hat{e} > 0$, $h_e \leq \rho < 1$ for all $e > \hat{e}$ and α.

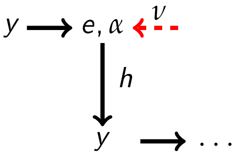
A.3. $0 \leq \alpha \leq 1$, distributed i.i.d. μ. μ has a continuous and strictly positive density on $[0,1]$.

Parent's utility of income y is described by a Bellman equation:

$$V^*(y) = \max_{0 \leq c \leq y} \mathrm{E}\left\{ U\left(c, V^*\left(h(\tilde{\alpha}, y - c)\right)\right)\right\}.$$

- The Bellman equation has a unique solution, and there is a $\bar{y}$ such that $y \leq \bar{y}$ for all α.

The solution defines a Markov process of income.



- If education is a normal good, then the Markov process is ergodic, and the invariant distribution μ has support on $[0, \hat{y}]$, where $\hat{y}$ solves $h\big(1, e^*(y)\big) = y$.

- The Bellman equation has a unique solution, and there is a $\bar{y}$ such that $y \leq \bar{y}$ for all a.

The solution defines a Markov process of income.

- If education is a normal good, then the Markov process is ergodic, and the invariant distribution μ has support on $[0, \hat{y}]$, where $\hat{y}$ solves $h(1, e^*(y)) = y$.

If utility were additively separable in consumption and children's utility, this would be the case.

An *education-specific tax policy* taxes each individual as a function of their education and their income. It is *redistributive* if the aggregate tax collection is 0 for every education level e.

Tax policy τ_1 is more egalitarian than tax policy τ_2 iff the distribution of income under τ_2 is riskier than that of τ_1 conditional on the education level e.

- If τ_1 and τ_2 are redistributive educational tax policies, and τ_1 is more egalitarian than τ_2, then for all income levels y, $V^*_{\tau_1}(y) > V^*_{\tau_2}(y)$.
- A result about universal public education.
- A result on the relationhip between ability and earnings.

Trust

Three Stories about Trust:

Reciprocity: Reputation games, folk theorems, . . .

Social Learning: Generalized trust.

Behavioral Theories: Evolutionary Psychology, prosocial preferences, . . .

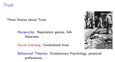


Image: Money as a form of generalized trust.

Three models of trust:

1. trust emerges from self interest — reputation
2. trust is learned beliefs about others. Trust in the context of intentional choice has ultimately to do about beliefs concerning the actions of others which have bearing on the decision. The TRUST variable in Knack and Keefer is essentially a belief variable.
   - 2.1 On individual's characteristics — this leads to Bayesian games.
   - 2.2 Directly on actors' behaviors — this leads to learning models in games.
3. the desire to trust and to be trusted is embodied in human behavior. Self image — the desire to be seen as trustworthy. See Andreoni-Laibson for same in a fairness context.

Trust theorists talk about generalized and particularized trust. Putnam refers to these as bridging and bonding social capital. They may have very different impacts on economic growth.

Inequality and Trust

- Evidence for a correlation between trust and income inequality
  - Rothstein and Uslaner (2005), Uslaner and Brown (2005).
- Trust is correlated with optimism about one's own life chances
  - Uslaner (2002)

Inequality and Trust

- Evidence for a correlation between trust and income inequality
  - Rothstein and Uslaner (2005), Uslaner and Brown (2005).
- Trust is correlated with optimism about one's own life chances
  - Uslaner (2002)

- I say "correlates" here because these studies are not causal. In fact trust is an equilibrium outcome — think of a Nash equilibrium.

- Uslaner and Brown (2005) argue that when inequality is high, people at the top and the bottom of the income distribution will not perceive each other as facing a shared fate. Therefore, they will have less reason to trust people of different backgrounds.

Networks, Trust, and Development

- Informal social organization substitutes for markets and formal social institutions in underdeveloped economies.
- In the US, periods of high growth have also been periods of decline in social capital (Putnam, 2000)
- Possibly: Social capital is needed for economic development only up to some intermediate stage, where generalized trust in institutions takes the place of informal trust arrangements.

Networks, Trust, and Development

- Informal social organization substitutes for markets and formal social institutions in underdeveloped economies.
- In the US, periods of high growth have also been periods of decline in social capital (Putnam, 2000)
- Possibly: Social capital is needed for economic development only up to some intermediate stage, where generalized trust in institutions takes the place of informal trust arrangements.

- Durlauf
- Generalized trust makes possible exchange on a larger scale, and so should be efficiency-enhancing.

Does Social Capital Have an Economic Payoff?

Knaak and Keefer (1997). "Does social capital have a payoff".

$$g_i = \mathbf{X}_i\gamma + \mathbf{Z}_i\pi + \text{CIVIC}_i\alpha + \text{TRUST}_i\beta + \epsilon_i$$

- g_i real per-capita growth rate.
- $\mathbf{X}_i$ control variables — Solow.
- $\mathbf{Z}_i$ control variables — "endogenous" growth models.
- CIVIC_i index of the level of civic cooperation.
- TRUST_i the percentage of survey respondents (after omitting those responding 'don't know') who, when queried about the trustworthiness of others, replied that 'most people can be trusted'.

Does Social Capital Have an Economic Payoff?

Knaak and Keefer (1997). "Does social capital have a payoff".

$$g_i = \mathbf{X}_i\gamma + \mathbf{Z}_i\pi + \text{CIVIC}_i\alpha + \text{TRUST}_i\beta + \epsilon_i$$

- g_i real per-capita growth rate.
- $\mathbf{X}_i$ control variables — Solow.
- $\mathbf{Z}_i$ control variables — "endogenous" growth models.
- CIVIC_i index of the level of civic cooperation.
- TRUST_i the percentage of survey respondents (after omitting those responding 'don't know') who, when queried about the trustworthiness of others, replied that 'most people can be trusted'.

- Why is this a silly regression?

- I pick on this paper because it is highly cited and representative of its type. It is not particularly bad relative to the rest of the literature.

- Some problems with this regression. Our focus: Trust is endogenous. They report on one regression in which they instrument for trust, which is also problematic.

Does Social Capital Have an Economic Payoff?

Knaak and Keefer (1997). "Does social capital have a payoff".

$$g_i = \mathbf{X}_i\gamma + \mathbf{Z}_i\pi + \text{CIVIC}_i\alpha + \text{TRUST}_i\beta + \epsilon_i$$

- g_i real per-capita growth rate.
- $\mathbf{X}_i$ control variables — Solow.
- $\mathbf{Z}_i$ control variables — "endogenous" growth models.
- CIVIC_i index of the level of civic cooperation.
- TRUST_i the percentage of survey respondents (after omitting those responding 'don't know') who, when queried about the trustworthiness of others, replied that 'most people can be trusted'.

How do K&K interpret this regression? Causal statements:

- Low trust can also discourage innovation.
- Government officials in societies with more trust may be perceived as more trustworthy, and their policy pronouncements as thus being more credible. To the extent that this is true, trust also triggers greater investment and other economic activity.
- To the extent that civic norms effectively constrain opportunism, the costs of monitoring and enforcing contracts are likely to be lower, raising the payoffs to many investments and other economic transactions.

BUT this is an equilibrium relationship. Let's build a causal model and see what moves the co-determined variables g and TRUST.

A Model of Trust

- A population of N completely anonymous individuals.
- Individuals have no distinguishing features, and so no one can be identified by any other.
- Individuals are randomly paired at each discrete date t, with the opportunity to pursue a joint venture. Simultaneously with her partner, each individual has to choose whether to participate (P) in the joint venture, or to pursue an independent venture (I). The entirety of her wealth must be invested in one or the other option. The individual with wealth w receives a gross return $w\pi$ from her choice, where π is realized from the following payoff matrix:

A Model of Trust

- $E\tilde{R} > E\tilde{e} > E\tilde{r}$.
- Individuals reinvest a constant fraction β of their wealth.
- Strategies for i are functions which map the history of is experience in the game to actions in the current period.
- Equilibria: Always play P, always play I are two equilibria.

Learning

Each individual i has a prior belief ρ, about the probability of one's opponent choosing P. The prior distribution is beta with parameters $a^i, b^i > 0$. In more detail,

$$\begin{aligned}\rho^i(x) &= \beta(a_0^i, b_0^i) \\ &= \frac{\Gamma(a_0^i + b_0^i)}{\Gamma(a_0^i)\Gamma(b_0^i)} x^{a_0^i - 1}(1-x)^{b_0^i - 1}.\end{aligned}$$

Let ρ_t^i denote individual i's posterior beliefs after t rounds of matching. The posterior densities ρ_t^i and ρ_t^j will be conditioned on different data, since all information is private. The updating rule for the β distribution has

$$\rho_t^i(h_t) \equiv \beta(a_t^i, b_t^i) = \beta(a_0^i + n, b_0^i + t - n)$$

for histories containing n P's and therefore $t - n$ I's. The posterior mean of the β distribution is $a_t^i/(a_t^i + b_t^i)$.

Optimal Play

$$q^* = (e - r)/(R - r)$$

- Let m_t^i denote i's mean of ρ_t.
- An optimal strategy for individual i is: Choose P if $m_t > q^*$ and choose I otherwise.

Theorem 3: For all initial beliefs $(\rho_0^1, \dots, \rho_0^N)$, almost surely either $\lim_t n_t^P = 0$ or $lim_t n_t^P = N$. The probabilities of both are positive. The limit wealth distributions in both cases is $\Pr\{\lim_t w_t > w\} \sim cw^k$, where k is k_P or k_T, and $k_P < k_T$.

Optimal Play

$$q^* = (e - r)/(R - r)$$

- Let m_t^i denote i's mean of ρ_T.
- An optimal strategy for individual i is: Choose P if $m_t > q^*$ and choose I otherwise.

Theorem 3: For all initial beliefs $(\rho_0^1, \ldots, \rho_0^N)$, almost surely either $\lim_t n_t^P = 0$ or $\lim_t n_t^P = N$. The probabilities of both are positive. The limit wealth distributions in both cases is
$\Pr\{\lim_t w_t > w\} \sim cw^k$, where k is k_P or k_T, and $k_P < k_T$.

- Thicker tails for P. Income inequality.

Social Learning

Averaging the Opinions of Others

- DeGroot (1974)
- X is some event. $p_i(t)$ is the probability that i assigns to the occurance of X at time t.
- A is a stochastic matrix. a_{ij} is the weight i gives to j's opinion.
- $p(t) = Ap(t-1) = \cdots = A^t p(0)$.

Averaging the Opinions of Others

- DeGroot (1974)
- X is some event. $p_i(t)$ is the probability that i assigns to the occurance of X at time t.
- A is a stochastic matrix. a_{ij} is the weight i gives to j's opinion.
- $p(t) = Ap(t-1) = \cdots = A^t p(0)$.

- Slides on herd behavior will be added. Banerjee (1992)

Averaging the Opinions of Others

Example

$$A = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix},$$

$$p(2) = A^2 p(0) = \begin{pmatrix} 5/18 & 8/18 & 5/18 \\ 5/12 & 5/12 & 2/12 \\ 1/4 & 1/2 & 1/4 \end{pmatrix} p(0),$$

$$p(t) = A^t p(0) \to \begin{pmatrix} 3/9 & 4/9 & 2/9 \\ 3/9 & 4/9 & 2/9 \\ 3/9 & 4/9 & 2/9 \end{pmatrix} p(0).$$

For all i,

$$p_i(\infty) = \frac{3}{9} p_1(0) + \frac{4}{9} p_2(0) + \frac{2}{9} p_3(0).$$

$$A = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix}.$$

$$p(2) = A^2 p(0) = \begin{pmatrix} 5/18 & 8/18 & 5/18 \\ 5/12 & 5/12 & 2/12 \\ 1/4 & 1/2 & 1/4 \end{pmatrix} p(0).$$

$$p(t) = A^t p(0) \to \begin{pmatrix} 3/9 & 4/9 & 2/9 \\ 3/9 & 4/9 & 2/9 \\ 3/9 & 4/9 & 2/9 \end{pmatrix} p(0).$$

For all i,

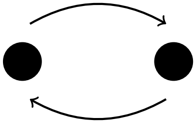
$$p_i(\infty) = \frac{3}{9}p_1(0) + \frac{4}{9}p_2(0) + \frac{2}{9}p_3(0).$$

- Final beliefs depend upon initial beliefs.
- Limit beliefs are identical.
- How general is this?
- Can limit beliefs be independent of the initial condition?

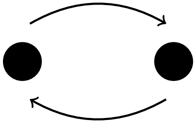
Averaging the Opinions of Others

Distinct Limits

$$A = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 2/3 & 1/3 \end{pmatrix}$$



$$A^t \to \begin{pmatrix} 2/5 & 3/5 & 0 & 0 \\ 2/5 & 3/5 & 0 & 0 \\ 0 & 0 & 3/5 & 2/5 \\ 0 & 0 & 3/5 & 2/5 \end{pmatrix}$$



$$p_i(t) \to \frac{2}{5} p_1(0) + \frac{3}{5} p_2(0) \quad \text{for } i = 1, 2.$$

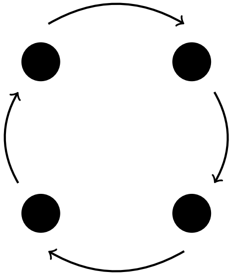
$$p_i(t) \to \frac{3}{5} p_3(0) + \frac{2}{5} p_4(0) \quad \text{for } i = 3, 4.$$

Averaging the Opinions of Others

No Limit

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$A^t = A^{(t-1) \bmod 3 + 1}$$



Theorem: If A is irreducible and aperiodic, then beliefs converge to a limit probability. For all j, $\lim_{t \to \infty} p_j(t) = \sum_i c_i^e p_i(0)$, where c^e is the unit-normalized eigenvalue centrality of A.

Theorem: If A is irreducible and aperiodic, then beliefs converge to a limit probability. For all j, $\lim_{t\to\infty} p_j(t) = \sum_i c_i^e p_i(0)$, where c^e is the unit-normalized eigenvalue centrality of A.

- Make comparisons to Markov processes.
- Right versus left eigenvectors.

Speed of Convergence

How long does it take for an individual's belief to get near to the limit belief?

$$|p_i(t) - p_i(\infty)| = \left| \sum_j \left(A_{ij}^t - \sum_j c_j^e \right) p_j(0) \right|$$

For each j $0 < p_j(0) < 1$,

$$\sup_{p(0)} \left| \sum_j \left(A_{ij}^t - \sum_j c_j^e \right) p_j(0) \right| =$$

$$\max \left\{ \sum_{j: A_{ij}^t \geq c_j^e} \left(A_{ij}^t - \sum_j c_j^e \right), - \sum_{j: A_{ij}^t \leq c_j^e} \left(A_{ij}^t - \sum_j c_j^e \right) \right\}$$

$$= \left|\left| A_{ij}^t - c^e \right|\right|_{TV}$$

Speed of Convergence

For x and y in the non-negative unit simplex,

$$||x - y||_{TV} = \sup_A |\sum_{i \in A}(x_i - y_i)|.$$

We want to max this over individuals, so

$$d(t) = sup_i \left|\left|A^t_{ij} - c^e\right|\right|_{TV}.$$

Define

$$t(\epsilon) = \min\{t : d(t) < \epsilon\}$$

$$t^* = t(1/4).$$

Speed of Convergence

For x and y in the non-negative unit simplex,

$$||x - y||_{TV} = \sup_A |\sum_{i \in A}(x_i - y_i)|.$$

We want to max this over individuals, so

$$d(t) = sup_i \left|\left|A^t_{ij} - c^*\right|\right|_{TV}.$$

Define

$$t(\epsilon) = \min\{t : d(t) < \epsilon\}$$
$$t^* = t(1/4).$$

- $|| \cdot ||_{TV}$ makes sense if you think of probability measures.
- for all integers $m > 0$, $d(mt^*) < 2^{-m}$.

A Lower Bound for t^*

$$Q(i,j) = c_i^e A_{ij}, \qquad Q(A,B) = \sum_{i \in A\, j \in B} Q(i,j).$$

$Q(A,B)$ is the amount of influence B inherits from A.

$$\Phi(S) = \frac{Q(S,S^c)}{\sum_{i \in S} c_i^e}, \qquad \Phi^* = \inf_{S : \sum_{i \in S} c_i^e \le 1/2} \Phi(S).$$

$\Phi(S)$ is the share of S's influence that is inherited by S^c.

Theorem: $$t^* \geq \frac{1}{4\Phi^*}.$$

A Lower Bound for t^*

$$Q(i,j) = c_i^e A_{ij}, \qquad Q(A,B) = \sum_{i \in A, j \in B} Q(i,j).$$

$Q(A,B)$ is the amount of influence B inherits from A.

$$\Phi(S) = \frac{Q(S,S^c)}{\sum_{i \in S} c_i^e}, \qquad \Phi^* = \inf_{S: \sum_{i \in S} c_i^e \le 1/2} \Phi(S).$$

$\Phi(S)$ is the share of S's influence that is inherited by S^c.

Theorem: $t^* \geq \frac{1}{4\Phi^*}$.

- If Φ^* is small, there is some set of individuals who is influenced mostly from itself. Information leaks slowly from S^c to S, and this inhibits learning.
- $Q(S, S^c) = Q(S^c, S)$.

Limit Beliefs and the "Wisdom of Crowds"

- Suppose that $p_i(0) = p + \epsilon_i$. The ϵ_i are all independent, have mean 0, and variances are bounded.
- What is the relationship between $p_i(\infty)$ and p?
- A sequence of networks $(V_n, E_n)_{n=1}^{\infty}$, $|V_n| = n$, with centrality vectors s^n, and belief sequences $p^n(t)$.

Definition: The sequence learns if for all $\epsilon > 0$,
$\Pr\{|\lim_{n\to\infty}\lim_{t\to\infty} p^n(t) - p| > \epsilon\} = 0$.

Theorem: If there is a $B > 0$ such that for all i, each individual's normailzed centrality is less than B/n, then the sequence learns.

- What conditions on the networks guarantee this?

Limit Beliefs and the "Wisdom of Crowds"

- Suppose that $p_i(0) = p + \epsilon_i$. The ϵ_i are all independent, have mean 0, and variances are bounded.
- What is the relationship between $p_i(\infty)$ and p?
- A sequence of networks $(V_n, E_n)_{n=1}^{\infty}$, $|V_n| = n$, with centrality vectors s^n, and belief sequences $p^n(t)$.

Definition: The sequence learns if for all $\epsilon > 0$,
$\Pr\{|\lim_{n\to\infty} \lim_{t\to\infty} p^n(t) - p| > \epsilon\} = 0$.

Theorem: If there is a $B > 0$ such that for all i, each individual's normailzed centrality is less than B/n, then the sequence learns.

- What conditions on the networks guarantee this?

- Normalized means the indicessum to 1.
- A finite network will only be correct by accident. So …
- a sequence of graphs, all irreducible aperiodic.
- Law of large numbers.

Bayesian Learning on Networks

Multi-armed bandit problem

- An undirected network $\mathcal{G}$.
- Two actions, A and B. A pays off 1 for sure. B pays off 2 with probability p and 0 with probability $1 - p$.
- At times $t = \{1, 2, \ldots\}$, each individual makes a choice, to maximize $\mathrm{E}\left\{\sum_{\tau=t}^{\infty} \beta^{\tau} \pi_{i\tau} | h_t\right\}$, the expected present value of the discounted payoff stream given the information.
- $p \in \{p_1, \ldots, p_K\}$. W.l.o.g. $p_j \neq p_k$ and $p_k \neq 1/2$.
- Each individual has a full-support prior belief μ_i on the p_k.
- Individuals see the choices of his neighbors, and the payoffs.

Bayesian Learning on Networks

Multi-armed bandit problem

- If the network contains only one member, this is the classic multi-armed bandit problem.
- How does the network change the classic results?
- What does one learn from the behavior of others?

Theorem: With probability one, there exists a time such that all individuals in a component play the same action from that time on.

- In one-individual problem, it is possible to lock into A when B is optimal. How does the likelihood of this change in a network?

- If the network contains only one member, this is the classic multi-armed bandit problem.
- How does the network change the classic results?
- What does one learn from the behavior of others?

Theorem: With probability one, there exists a time such that all individuals in a component play the same action from that time on.

- In one-individual problem, it is possible to lock into A when B is optimal. How does the likelihood of this change in a network?

- Bala and Goyal (1996)
- Compare to trust model — learning about exogenous versus endogenous things.
- Review classic Bandit problem results.
- BG simplify by not allowing individuals to infer what others have *seen*.
- Why the theorem? If everyone plays B only finitely often, then the component locks into A in finite time. Suppose one individual plays B infinitely often. Then he learns p, and $p > 1/2$. SInce he learns, there is time beyond which he choose only B. If an individual choosing A is connected to the B-player, then he sees B's payoff stream and learns p. Thus at some point he switches to p, and at some point locks in. Etc.
- Intuitively, the greater the degree, the better the learning. Is this true?

$(\Omega, \mathcal{F}, p)$ A probability space.

X A finite set of actions.

Y_i A finite set of signals observed by i. $y_i : \Omega \to Y_k$ is $\mathcal{F}$-measurable.

$\sigma(f)$ If f is a measurable mapping of Ω into any measure space, σf is the σ-algebra generated by f. Define $\sigma(y_k) = \mathcal{Y}_k$.

Definition: A decision function maps states Ω to actions X. A decision rule maps σ-fields on Ω to decision rules, that is, $d(\mathcal{G}) : \Omega \to X$. For any σ-field $\mathcal{G}$, $d(\mathcal{G})$ is $\mathcal{G}$-measurable. That is, $\sigma d(\mathcal{G}) \subset \mathcal{G}$.

- Updating of beliefs:

$$\mathcal{F}_k(t+1) = \mathcal{F}_k(t) \vee \bigvee_{j \neq k} \sigma d\left(\mathcal{F}_j(t)\right),$$

$$\mathcal{F}_k(0) = \mathcal{Y}_k.$$

Key Property: If $\sigma d(\mathcal{G}) \subset \mathcal{H} \subset \mathcal{G}$, then $d(\mathcal{G}) = d(\mathcal{H})$.

- Updating of beliefs:

$$\mathcal{F}_k(t+1) = \mathcal{F}_k(t) \vee \bigvee_{j \neq k} \sigma d\left(\mathcal{F}_j(t)\right),$$

$$\mathcal{F}_k(0) = \mathcal{Y}_k.$$

Key Property: If $\sigma d(\mathcal{G}) \subset \mathcal{H} \subset \mathcal{G}$, then $d(\mathcal{G}) = d(\mathcal{H})$.

Join $A \vee B$ is the smallest σ-field containing A and B.

Theorem: Suppose d has the key property. Then there are σ-algebras $\mathcal{F}_k \subset \bigvee_k Y_k$ and $T \geq 0$ such that $\mathcal{F}_k(t) = \mathcal{F}_k$ for all $t \geq T$, and for all k and j,

$$d(\mathcal{F}_k) = d(\mathcal{F}_j) = d\left(\bigwedge_i \mathcal{F}_i\right).$$

If the decision functions for all individuals are common knowledge, then they agree.

Theorem: Suppose d has the key property. Then there are σ-algebras $\mathcal{F}_k \subset \bigvee_k Y_k$ and $T \geq 0$ such that $\mathcal{F}_k(t) = \mathcal{F}_k$ for all $t \geq T$, and for all k and j,

$$d(\mathcal{F}_k) = d(\mathcal{F}_j) = d\left(\bigwedge_i \mathcal{F}_i\right).$$

If the decision functions for all individuals are common knowledge, then they agree.

Proof: Washburn and Teneketzis (1984).

Now given is a connected undirected network (V, E).

- Individuals i and k communicate directly if there is an edge connecting them.
- Individuals i and k communicate indirectly if there is a path connecting them.

Key Network Property: For any sequence of individuals $k = 1, 2, \ldots, n$, if $\sigma d(\mathcal{F}_k) \subset \mathcal{F}_{k+1}$ and $\sigma d(\mathcal{F}_n) \subset \mathcal{F}_1$, then $d(\mathcal{F}_k) = d(\mathcal{F}_1)$ for all k.

Bayesian Learning on Networks

Updating of beliefs:

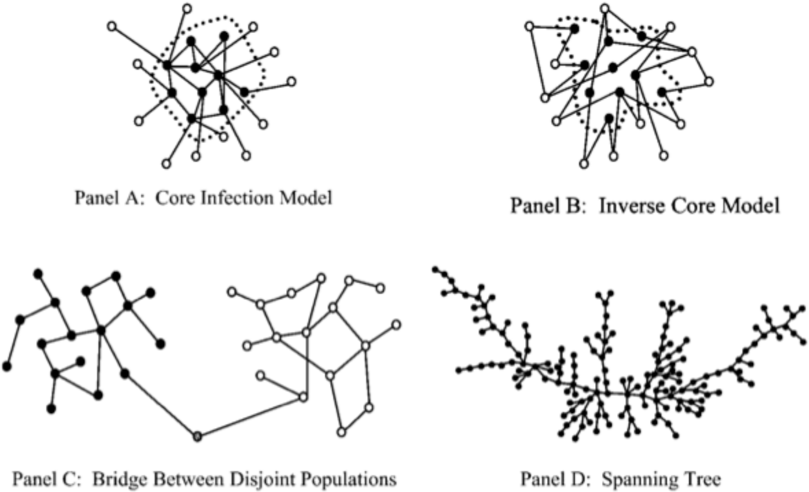
$$\mathcal{F}_k(t+1) = \mathcal{F}_k(t) \vee \bigvee_{j \sim k} \sigma d\left(\mathcal{F}_j(t)\right),$$
$$\mathcal{F}_k(0) = \mathcal{Y}_k.$$

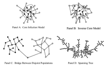
Theorem: Suppose d has the key network property. Then there are σ-algebras $\mathcal{F}_k \subset \bigvee_k Y_k$ and $T \geq 0$ such that $\mathcal{F}_k(t) = \mathcal{F}_k$ for all $t \geq T$, and for all k and j,

$$d(\mathcal{F}_k) = d(\mathcal{F}_j) = d\left(\bigwedge_i \mathcal{F}_i\right).$$

Diffusion

Network Effects and Diffusion

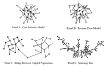




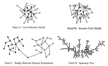
The simplest epidemiological models assume random mixing among all members of the population. Under random mixing, the number of new infections at time t is easily calculated as the number of susceptibles times the number of infecteds times the proportion of contacts between susceptibles and infecteds that result in infection. The result of a random mixing model is the classical S-shaped diffusion curve, where one observes a slow start, followed by exponential growth, and then a decline, either from recovery or death

For transmission of STDs, partner selection matters. Preferred-partner mixing models are based on the homophily principle. Most preferred-mixing models do not explicitly look at the network structure of relationships, but this can be important, and is changing.

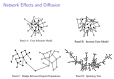
Three stylized images of sexual networks can be derived from the literature on the diffusion of STDs.



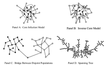
A core is a group of high activity-level actors (e.g., those with multiple partners or who are frequent drug users) who interact frequently and pass infection to one another (often causing reinfection for treatable STDs), and diffuse infection out to a less densely connected population. Cores are predicted to sustain endemic pockets of disease, since the pattern of intense interaction among members of the core pushes R_0 in the core above 1. In a core, it is likely that an individual’ĂŹs past partner is tied through multiple chains to his or her current or future partner. Thus if cores exist in a population, cyclicity will be extremely high in the network, and the length of chains connecting pairs of indi- viduals in the population (geodesics) will be low.



A key mode of transmission in developing countries is male long-distance truck drivers having sex with female commercial sex workers (CSWs), members of the groups that constitute possible infection reservoirs (like CSWs) are structurally disconnected from one another and do not transmit infection directly to one another. Capturing such dynamics — which may be more characteristic of two-sex diffusion processes — requires more complex switching models, often called inverse core models (Garnett et al., 1996). In an inverse core, a central group of infected persons pumps disease out to others but does not pass infection directly among themselves. The key difference between a core and an inverse core stems from the social organization of sexual relations, since truck-drivers are more likely than other potential carriers to spread infection to individuals not in the graph (specifically, their regular sex partners). Compared to core models, an increased focus on heirarchy.



A third model in the epidemiological literature describes disease diffusion dynamics as driven by bridging processes (Aral, 2000; Gorbach et al., 2000; Morris et al., 1996). These models posit two populations of persons engaged in different behaviors (i.e., a high-risk and a low-risk population) linked by a few individuals who bridge the boundary between each world (e.g., an IV drug user who shares needles with his drug partners and who has sex with non-IV-drug users).



A spanning tree is a long chain of interconnections that stretches across a population, like rural phone wires running from a long trunk line to individual houses (Hague and Harary, 1983, 1996). The global structure of a chainlike spanning tree is characterized by a graph with few cycles, low redundancy, and consequently very sparse overall density. The shortest distance between any two randomly selected individuals (geodesic) is significantly higher than that observed in either the core or inverse core structures.

Random-mixing dynamics and positive preferences for partners do not produce spanning tree structures. Rather, this network structure appears when formal or informal rules preclude the enactment of specific relations. In the language of kinship structures, spanning trees are the product of negative proscriptions: sets of rules about whom one cannot be in a relationship with.

Varieties of Action

- Graphical Games — Diffusion of action
  - Blume (1993, 1995) — Lattices
  - Morris (2000) — General graphs
  - Young and Kreindler (2011) — Learning is fast
- Social Learning — Diffusion of knowledge
  - Banerjee (1992)
  - Bikhchandani et al. (1992)
  - Rumors

Coordination Games

	A	B
A	a,a	0,0
B	0,0	b,b

$a, b > 0$

Pure coordination game

Three equilibria:

$$\langle a, a \rangle, \quad \langle b, b \rangle, \quad \text{and} \quad \langle (\frac{b}{a+b}, \frac{a}{a+b}), (\frac{b}{a+b}, \frac{a}{a+b}) \rangle$$

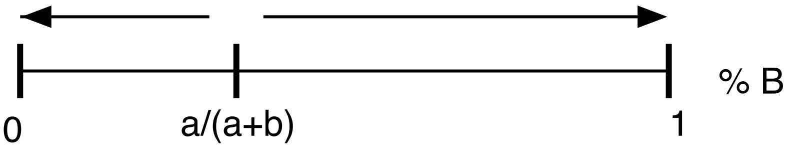
Coordination Games

	A	B
A	a,a	0,0
B	0,0	b,b

$a, b > 0$

Pure coordination game

Best response dynamics



Coordination Games

	A	B
A	a,a	d,c
B	c,d	b,b

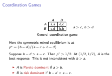
$a > c,\ b > d$

General coordination game

Here the symmetric mixed equilibrium is at
$p^* = (b - d)/(a - c + b - d)$.

Suppose $b - d > a - c$. Then $p^* > 1/2$. At $(1/2, 1/2)$, A is the best response. This is not inconsistent with $b > a$.

- A is Pareto dominant if $a > b$.
- B is risk dominant if $b - d < a - c$.



We have theories that favor risk dominant selection over payoff dominant selection: Blume (1993), Kandori, Mailath and Robb (1993), Young (1993).

Coordination Games — Stochastic Stability

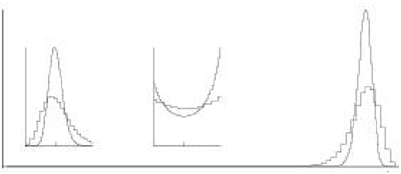
Continuous time stochastic process

- Each player has an alarm clock. When it goes off, she makes a new strategy choice. The interval between rings has an exponential distribution.
- Strategy revision:
  - Each individual best-responds with prob. $1 - \epsilon$, Kandori, Mailath and Robb (1993); Young (1993)

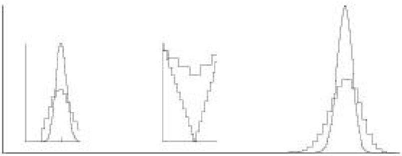
    or
  - The log-odds of choosing A over B is proportional to the payoff difference — logit choice, Blume (1993, 1995).

The Stochastic Process

This is a Markov process on the state space $[0, \ldots, N]$, where the state is the number of players choosing B.



Logit Choice



Mistakes

In both cases, as $\text{Prob}\{\text{best response} \uparrow 1\}$, $\text{Prob}\{N\} \uparrow 1$.

The Stochastic Process

This is a Markov process on the state space $[0, \ldots, N]$, where the state is the number of players choosing B.

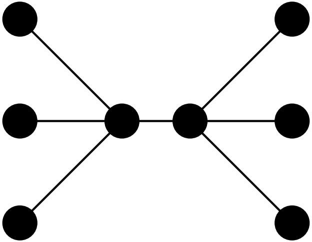
Logit Choice Mistakes

In both cases, as Prob{best response $\uparrow 1$}, Prob{N} $\uparrow 1$.

- One big peak on right. Very small peak on left, bimodel, with U in the middle.

Coordination on Networks

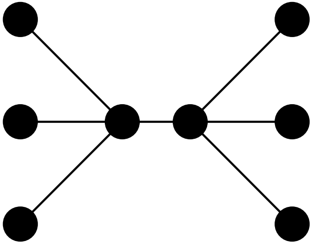
- Is the answer the same on every graph?



.

Coordination on Networks

- Is the answer the same on every graph?



Mistake: $0 : 0.5\, N : 0.5$.   Logit: $N : 1$.

General Analysis

- In general, the strategy revision process is an ergodic Markov process.
- There is no general characterization of the invariant distribution.
- The answer is well-understood for potential games and logit updating.

General Analysis

- In general, the strategy revision process is an ergodic Markov process.
- There is no general characterization of the invariant distribution.
- The answer is well-understood for potential games and logit updating.

- Gibbs states
- Will be in the final version of the notes.

A General Diffusion Model

- Best response strategy revision. If fraction q or more of your neighbors choose A, then you choose A.
- Two obvious equilibria: All A and All B.
- How easy is it to "tip" from one to the other? What about intermediate equilibria?

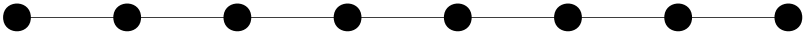
A General Diffusion Model

- Imagine that everyone initially uses B.
- Now a small group adopts A.
- When does it spread, when does it stop?
- The answer should depend on the network structure, who are the initial adopters, and the threshold p^*.

Diffusion of Coordination — Line

When the Poisson alarm clock rings, the player best responds to his neighbors. $p^* < 1/2$. Questions:

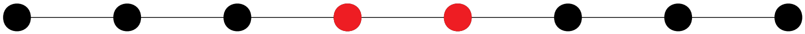
- Are islands of risk dominance stable?
- Can risk dominance spread?



Diffusion of Coordination — Line

When the Poisson alarm clock rings, the player best responds to his neighbors. $p^* < 1/2$. Questions:

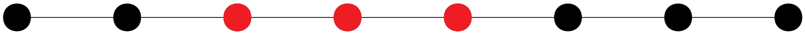
- Are islands of risk dominance stable?
- Can risk dominance spread?



Diffusion of Coordination — Line

When the Poisson alarm clock rings, the player best responds to his neighbors. $p^* < 1/2$. Questions:

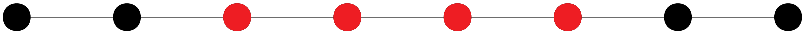
- Are islands of risk dominance stable?
- Can risk dominance spread?



Diffusion of Coordination — Line

When the Poisson alarm clock rings, the player best responds to his neighbors. $p^* < 1/2$. Questions:

- Are islands of risk dominance stable?
- Can risk dominance spread?



Diffusion of Coordination — Line

When the Poisson alarm clock rings, the player best responds to his neighbors. $p^* < 1/2$. Questions:

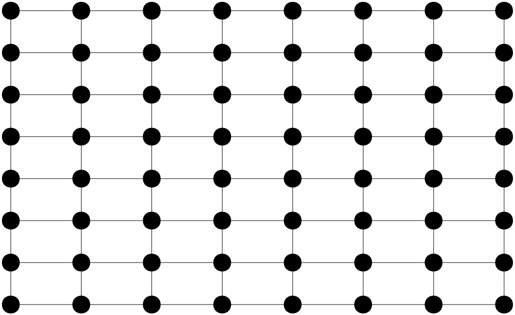
- Are islands of risk dominance stable?
- Can risk dominance spread?

Contagion threshold 1/2, stability threshold 1/2.

Diffusion of Coordination — Lattices

When the Poisson alarm clock rings, the player best responds to his neighbors. $p^* < 1/4$. Questions:

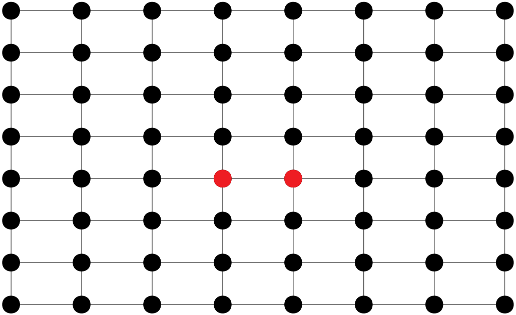
- Are islands of risk dominance stable?
- Can risk dominance spread?



Diffusion of Coordination — Lattices

When the Poisson alarm clock rings, the player best responds to his neighbors. $p^* < 1/4$. Questions:

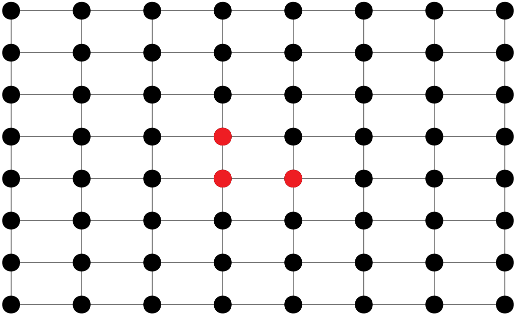
- Are islands of risk dominance stable?
- Can risk dominance spread?



Diffusion of Coordination — Lattices

When the Poisson alarm clock rings, the player best responds to his neighbors. $p^* < 1/4$. Questions:

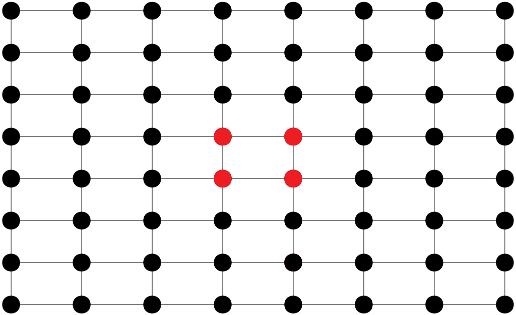
- Are islands of risk dominance stable?
- Can risk dominance spread?

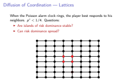


Diffusion of Coordination — Lattices

When the Poisson alarm clock rings, the player best responds to his neighbors. $p^* < 1/4$. Questions:

- Are islands of risk dominance stable?
- Can risk dominance spread?





Take threshold $p^* < 1/4$, nearest neighbor interaction.

Use Language of cascade, cascade threshold.

Cascade threshold: A property of the graph — the maximum vertex threshold such that a cascade can take place.

Nodes, once switched, will never switch back.

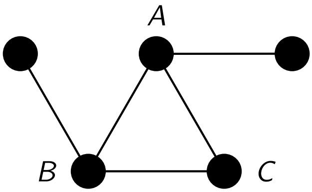
What happens if $1/4 < p^* < 1/2$? $1, 2, 3$ unstable, 4 is stable but cannot grow.

Why isn't this good theory?

Diffusion of coordination — General Graphs

- A cluster of density p is a set of vertices C such that for each $v \in C$, at least fraction p of v's neighbors are in C.

The set $C = \{A, B, C\}$ is a cluster of density $2/3$.



General Graphs

Two observations:

- Every graph will have a cascade threshold.
- If the initial adoptees are a cluster of density at least p^*, then diffusion can only move forward.

General Graphs: Clusters Stop Cascades

Consider a set S of initial adopters in a network with vertices T, and suppose that remaining nodes have threshold q.

Claim: If S^c contains a cluster with density greater than $1 - q$, then S will not cause a complete cascade.

Proof: If there is a set $T \subset S^c$ with density greater than $1 - q$, then even if S/T chooses A, every member of T has fraction more than $1 - q$ choosing B, and therefore less than fraction q are choosing A. Therefore no member of T will switch.

General Graphs: Clusters Stop Cascades

Claim: If a set $S \subset V$ of initial adopters of an innovation with threshold q fails to start a cascade, then there is a cluster $C \in V/S$ of density greater than $1 - q$.

Proof: Suppose the innovation spreads from S to T and then gets stuck. No vertex in T^c wants to switch, so less than a fraction q of its neighbors are in T, more than fraction $1 - q$ are out. That is T^c has density greater than $1 - q$.

Networks and Optimality

- Networks make it easier for cascades to take place.
  - In the fully connected graph, a cascade from a small group never takes place. With stochastic adjustment in the mistakes model, the probability of transiting from all A to all B is $O(\epsilon^{qN})$, where q is the indifference threshold. On a network, the probability of transiting from all A to all B is on the order of ϵ^{K}, where K is the size of a group needed to start a cascade, and this is independent of N.

Networks and Optimality

- Networks make it easier for cascades to take place.
  - In the fully connected graph, a cascade from a small group never takes place. With stochastic adjustment in the mistakes model, the probability of transiting from all A to all B is $O(\epsilon^{qN})$, where q is the indifference threshold. On a network, the probability of transiting from all A to all B is on the order of ϵ^K, where K is the size of a group needed to start a cascade, and this is independent of N.
- This is not always optimal!
  - Risk dominance and Pareto dominance can be different. This can be understood as a robustness question. If the population has correlated on the efficient action, how easy is it to undo? Hard if the efficient action is risk dominant. If the efficient action is not risk-dominant, it is easier to undo on sparse networks than on nearly completely connected networks.

References: Introduction I

Asch, Solomon E. 1951. "Effects of group pressure on the modification and distortion of judgements." In *Groups, Leadership and Men*, edited by H. Guetzkow. Pittsburgh: Carnegie Press.

Baker, Wayne E. 1984. "The Social Structure of a National Securities Market." *American Journal of Sociology* 89 (4):775–811.

Duesenberry, James S. 1960. "Comment on G. Becker, 'An economic analysis of fertility'." In *Demographic and Economic Change in Developed Countries*, edited by George B. Roberts. New York: Columbia University Press for the National Bureau of Economic Research, 231–34.

Milgram, Stanley. 1967. "The Small World Problem." *Psychology Today* 2: 60–67.

References: Introduction II

Polanyi, Karl. 1944. *The Great Transformation*. New York: Farrar and Rinehart.

Reiss Jr., Albert J. 1986. "Co-offending influences on criminal careers." In *Criminal Careers and 'Career Criminals'*, vol. 2, edited by Alfred Blumstein, Jacqueline Cohen, Jeffrey A. Roth, and Christy A. Visher. Washington, DC: National Academy Press, 121–160.

Sacerdote, Bruce I. 2001. "Peer effects with random assignment results for Dartmouth roommates." *Quarterly Journal of Economics* 116 (2):681–704.

Sherif, M. *et al.* 1954/1961. *Intergroup Conflict and Cooperation: The Robbers Cave Experiment.* Norman: University of Oklahoma Book Exchange.

References: Introduction III

Zimbardo, Philip. 1969. "The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos." In W.D. Arnold & D. Levine (Eds.) *Nebraska Symposium on Motivation*, 237–307. Lincoln: University of Nebraska Press.

References: Network Science I

Eggenberger, F. and G. Polya. 1923. "Über die Statistik verketteter Vorgänge". *Zeitschrift für Angewandte Mathematik und Mechanik* 3 (4): 279–289.

Gladwell, Malcom. 1999. "Six degrees of Lois Weisberg." *New Yorker*, January 11.

Korte, Charles and Stanley Milgram. 1970. "Acquaintance networks between racial groups: Application of the small world method." *Journal of Personality and Social Psychology* 15 (2):101–08.

R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins and E. Upfal. "Stochastic models for the web graph." FOCS 2009.

Milgram, Stanley. 1967. "The small world problem." *Psychology Today* 2:60–67.

Newman, Mark E. J. 2003. "The structure and function of complex networks." *SIAM Review* 45 (2):167–256.

References: Network Science II

Herbert A. Simon (1955). "On a class of skew distribution functions". 55 (3–4): 425–440.

Travers, Jeffrey and Stanley Milgram. 1969. "An experimental study of the small world problem." *Sociometry* 32 (4):425–43.

G. K. Zipf. 1949. *Human Behavior and the Principle of Least Effort* Addison-Wesley, Cambridge, MA.

Consequences of Social Networks

Besley, Timothy and Anne Case. 1993. "Modeling Technology Adoption in Developing Countries." *American Economic Review* :396–402.

Conley, Timothy G. and Christopher R. Udry. 2010. "Learning About a New Technology: Pineapple in Ghana." *American Economic Review*

Foster, Andrew D and Mark R Rosenzweig. 1995. "Learning by Doing and Learning from Oth- ers: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 1176–1209.

Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman. 1996. "Crime and Social Interaction." *Quarterly Journal of Economics* 111 (2):507–48.

Munshi, Kaivan. 2004. "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution." *Journal of Development Economics* 73:185–213.

References: Properties of Social Networks

Amaral, L. A. N., A. Scala, M. Barthélémy, and H. E. Stanley. 2000. "Classes of small-world networks." *Proc. Natl. Acad. Sci. USA.* 97:11149–52.

Bearman, Peter, James Moody, and Katherine Stovel. 2004. "Chains of affection: The structure of adolescent romantic and sexual networks." *American Journal of Sociology* 110 (1):44–99.

Bonacich, Phillip. 1987. "Power and centrality: A family of measures." *American Journal of Sociology* 92 (5): 1170–1182.

Bonacich, Phillip and Paulette Lloyd. 2001. "Eigenvector-like measures of centrality for asymmetric relations." *Social Networks* 23 (3): 191–201.

Christakis, Nicholas A. and James H. Fowler. 2007. "The spread of obesity in a large social network over 32 years." *New England Journal of Medicine* 357:370–9.

Cohen-Cole, Ethan and Jason M. Fletcher. 2008. "Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic." *Journal of Health Economics* 27 (5):1382–87.

Davis, Gerald F., Mina Yoo, and Wayne E. Baker. 2003. "The small

References: Labor Markets

Calvó-Armengol, Antoni and Matthew O. Jackson. 2004. "The Effects of Social Networks on Employment and Inequality." *American Economic Review* 94 (3):426–54.

Granovetter, M. S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78 (6):1360–80.

Montgomery, James D. 1991. "Social networks and labor-market outcomes: Towards an economic analysis." *American Economic Review* 81 (5):1408–18.

Rapoport, A. and W. Horvath. 1961. "A study of a large sociogram." *Behavioral Science* 6:279–91.

Scotese, Carol A. 2012. "Wage inequality, tasks and occupations." Unpublished, Virginia Commonwealth University.

Tian, Felicia and Nan Lin. 2016. "Weak ties, strong ties, and job mobility in urban China." *Social Networks* 44:117–29.

Yakubovich, Valery. 2005. "Weak ties, information, and influence: How workers find jobs in a local Russian labor market." *American Sociological Review* 70 (3):408–21.

References: Peer Effects and Complementarities I

Blume, L., W. Brock, S. Durlauf, and Y. Ioannides. 2011. "Identification of Social Interactions." In *Handbook of Social Economics*, vol. 1B, edited by J. Benhabib, A. Bisin, and M. Jackson. Amsterdam: North Holland, 853–964.

Blume, Lawrence E., William Brock, Steven N. Durlauf, and Rajshri Jayaraman. 2013. "Linear Social Interaction Models." *Journal of Political Economy* 123 (2):444–96.

Datcher, Linda. 1982. "Effects of Community and Family Background on Achievement." *Review of Economics and Statistics* 64: 32–41.

Durlauf, Steven N. 2004. "Neighborhood effects." In *Handbook of Regional and Urban Economics*, edited by J. V. Henderson and J. F. Thisse. Amsterdam: Elsevier, 2173–2242.

References: Peer Effects and Complementarities II

Hoxby, Caroline M. and Gretchen Weingarth. 2005. "Taking race out of the equation: School reassignment and the structure of peer effects." NBER Working Paper.

Ioannides, Yannis M. and Giorgio Topa. 2010. "Neighborhood effects: Accomplishments and looking beyond them." *Journal of Regional Science* 50 (1):343–62.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60:531–42.

Sacerdote, B. 2011. "Peer effects in education: How might they work, how big are they and how much do we know thus far?" *Handbook of the Economics of Education* 3:249–277.

References: Matching I

Blume, Lawrence, David Easley, Jon Kleinberg, Bobby Kleinberg and Éva Tardos. 2013. "Network formation in the presence of contagious risk." *ACM Transactions on Economics and Computation* 1 (2).

Jackson, Matthew O. and Asher Wolinsky. 1996. "A strategic model of network formation." *Journal of Economic Theory* 71:44–74.

Gale, David and Lloyd Shapley. 1962. "College admissions and the stability of marriage." *American Mathematical Monthly* 69:9–14.

Shapley, Lloyd and Martin Shubik. 1972. "The assignment game I: The core." *International Journal of Game Theory* 1: 111–130.

References: Social Capital I

Bourdieux, P. and L. J. D. Wacquant. 1992. *An Invitation to Reflexive Sociology*. Chicago, IL: University of Chicago Press.

Fukuyama, Francis. 1996. *Trust: Social Virtues and the Creation of Prosperity*. New York: Simon and Schuster.

Knaak, Stephen and Philip Keefer. 1997. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics* 112 (4):1251–88.

Lin, Nan. 2001. *Social Capital*. Cambridge UK: Cambridge University Press.

Loury, Glenn. 1981. "Intergenerational Transfers and the Distribution of Earnings." *Econometrica* 49:843–867.

Loury, Glenn. 1992. "The economics of discrimination: Getting to the core of the problem." *Harvard Journal for African-Americvan Public Policy* 1:91–110.

References: Social Capital II

Portes, Alejandro. 1998. "Social capital: Its origins and applications in modern sociology." *Annual Review of Sociology* 24:1–24.

Putnam, Robert D. 1995. "Bowling alone: America's declining social capital." *Journal of Democracy* 6:65–78.

Putnam, Robert. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster.

Rothstein, B. and Eric M. Uslaner. 2005. "All for all: Equality, corrpution, and social trust." *World Politics* 58 (1):41–72.

Uslaner, Eric M. 2002. *The Moral Foundations of Trust*. Cambridge UK: Cambridge University Press.

Uslaner, Eric M. and M. Brown. 2005. "Inequality, trust, and civic engagement." *American Politics Research* 33 (6):868–94.

References: Social Learning

Bala, Venkatesh and Sanjeev Goyal. 1996. "Learning from neighbors." *Review of Economic Studies* 65:595-621.

Banerjee, Abhijit V. 1992. "A simple model of herd behavior." *Quarterly Journal of Economics* 107 (3):797–817.

Bonacich, Phillip. 1987. "Power and centrality: A family of measures." *American Journal of Sociology* 92 (5):1170–82.

Bonacich, Phillip and Paulette Lloyd. 2001. "Eigenvector-like measures of centrality for asymmetric relations." *Social Networks* 23:191–201.

DeGroot, Morris H. 1974. "Reaching a consensus." *Journal of the American Statistical Association* 69:118–21.

Jackson, Matthew O. and Benjamin Golub. 2010. "Naive learning in social networks and the wisdom of crowds." *American Economic Journal: Microeconomics* 2 (1):112-49.

Washburn, R. B. and D. Teneketzis. 1984. "Asymptotic Agreement Among Communicating Decisionmakers." *Stochastics* 13:103–29.

References: Diffusion

Aral, S. O. 2000. "Behavioral Aspects of Sexually Transmitted Diseases: Core Groups and Bridge Populations." *Sexually Transmitted Diseases* 27:327?28.

Banerjee, Abijit. 1992. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics* 107 (3):797–817.

Bikhchandani, Sushil, David Hirshleifer and Ivo Welch "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." *Journal of Political Economy* 100 (5):992–1026.

Blume, Lawrence E. 1993. "The statistical mechanics of strategic interaction." *Games and Economic Behavior* 5 (5):387–424.

———. 1995. "The statistical mechanics of best-response strategy revision." *Games and Economic Behavior* 11 (2):111–145.

Garnett G. P., J. P. Hughes, R. M. Anderson, B. P. Stoner, S. O. Aral, W. L. Whittington, H. H. Handsfield, and K. K. Holmes. 1996. "Sexual Mixing Patterns of Patients Attending Sexually Transmitted Disease Clinics." *Sexually Transmitted Diseases* 23 (3): 248?57.

Gorbach P. M., H. Sopheab, T. Phalla, H. B. Leng, S. Mills, A.

References: Community Structure

Massey, Douglas and Nancy Denton. 1988. "The dimensions of residential segregation." *Social Forces* 67:281–315.

References: General

Easley, David A. and Jon Kleinberg. 2010. *Networks, Crowds and Markets: Reasoning About a Highly Connected World*. Cambridge UK: Cambridge University Press.

Jackson, Matthew O. 2008. *Social and Economic Networks*. Princeton University Press.