Introduction
“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)
## Where do Social Interactions Appear?

<table>
<thead>
<tr>
<th>Phenomena</th>
<th>Mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor markets</td>
<td>Peer effects</td>
</tr>
<tr>
<td>Career Choices</td>
<td>Stigma</td>
</tr>
<tr>
<td>Retirement</td>
<td>Role models</td>
</tr>
<tr>
<td>Fertility</td>
<td>Social Norms</td>
</tr>
<tr>
<td>Health</td>
<td>Social Learning</td>
</tr>
<tr>
<td>Education Outcomes</td>
<td>Social Capital?</td>
</tr>
</tbody>
</table>
Research Methodologies

- Ethnographies
- Field Experiments
- Large-Scale Experiments, Natural and Real
Questions

- What are appropriate tools for modelling social interactions?

- Models of social interactions: Social norms, group membership, peer effects.

- Describe the peer effects. What goes on at the micro level?

- What are the aggregate effects of interaction on social networks?
Mennis and Harris (2001)

Although other research has investigated deviant peer contagion, and still other research has examined offense specialization among delinquent youths, we have found that deviant peer contagion influences juvenile recidivism, and that contagion is likely to be associated with repeat offending. These findings suggest that juveniles are drawn to specific types of offending by the spatially-bounded concentration of repeat offending among their peers. Research on causes of delinquency within neighborhoods, then, may produce more useful causal models than studies that ignore spatial concentrations of offense patterns.
### Table 4
Logistic regression of person offense recidivism (N = 7166).

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.91*** (13.32)</td>
<td>0.92*** (10.13)</td>
<td>0.92*** (9.94)</td>
</tr>
<tr>
<td>CI 0.87-0.96</td>
<td>CI 0.88-0.97</td>
<td>CI 0.88-0.97</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.85 (1.55)</td>
<td>0.86 (1.41)</td>
<td>0.83 (1.98)</td>
</tr>
<tr>
<td>CI 0.66-1.10</td>
<td>CI 0.67-1.11</td>
<td>CI 0.64-1.08</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.64*** (11.26)</td>
<td>0.65*** (10.01)</td>
<td>0.73* (5.39)</td>
</tr>
<tr>
<td>CI 0.49-0.83</td>
<td>CI 0.50-0.85</td>
<td>CI 0.56-0.95</td>
<td></td>
</tr>
<tr>
<td>Public assistance</td>
<td>0.99 (0.01)</td>
<td>1.00 (0.00)</td>
<td>1.01 (0.01)</td>
</tr>
<tr>
<td>CI 0.84-1.17</td>
<td>CI 0.84-1.18</td>
<td>CI 0.01</td>
<td></td>
</tr>
<tr>
<td>Parental crime</td>
<td>1.35*** (9.25)</td>
<td>1.35*** (8.85)</td>
<td>1.35*** (8.84)</td>
</tr>
<tr>
<td>CI 1.11-1.65</td>
<td>CI 1.11-1.64</td>
<td>CI 1.11-1.64</td>
<td></td>
</tr>
<tr>
<td>Number of prior arrests</td>
<td>1.12*** (15.87)</td>
<td>1.13*** (16.69)</td>
<td>1.12*** (15.74)</td>
</tr>
<tr>
<td>CI 1.06-1.19</td>
<td>CI 1.06-1.19</td>
<td>CI 1.06-1.19</td>
<td></td>
</tr>
<tr>
<td>Prior institutional living arrangement</td>
<td>1.15 (2.44)</td>
<td>1.14 (2.35)</td>
<td>1.13 (1.92)</td>
</tr>
<tr>
<td>CI 0.97-1.36</td>
<td>CI 0.96-1.36</td>
<td>CI 0.95-1.34</td>
<td></td>
</tr>
<tr>
<td>Instant offense type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Person offense</td>
<td></td>
<td>1.23* (6.39)</td>
<td>1.22* (5.58)</td>
</tr>
<tr>
<td>CI</td>
<td>CI 1.05-1.44</td>
<td>CI 1.03-1.43</td>
<td></td>
</tr>
<tr>
<td>Contagion effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area person recidivism rate (×10)</td>
<td></td>
<td></td>
<td>2.77*** (84.05)</td>
</tr>
<tr>
<td>CI</td>
<td>CI 2.23-3.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.40* (5.91)</td>
<td>0.31*** (8.99)</td>
<td>0.63***</td>
</tr>
<tr>
<td>AUC</td>
<td>0.58***</td>
<td>0.59***</td>
<td></td>
</tr>
</tbody>
</table>

A gray box indicates a variable that was excluded from that model run. Cell values indicate odds ratios. Wald statistic is shown in parentheses. “CI” indicates confidence interval at 95% confidence. *p < 0.05, **p < 0.01, ***p < 0.005.
Aggregate Analysis

Glaeser Sacerdote and Scheinkman 1996.

The most puzzling aspect of crime is not its overall level nor the relationships between it and either deterrence or economic opportunity. Rather, following Quetelet [1835], we believe that the most intriguing aspect of crime is its astoundingly high variance across time and space.

Positive covariance across agents’ decisions about crime is the only explanation for variance in crime rates higher than the variance predicted by difference in local conditions.
A Model

- $2N + 1$ 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 left.
- Type of individual $i$ is $p_i$. 

![Diagram of lattice with individuals of different types]
Model (of sorts)

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

\[
E\{a_i\} = \frac{p_1}{p_0 + p_1} \equiv p, \quad S_n = \sum_{|i|\leq n} \frac{a_i - p}{2n + 1}.
\]

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

where

\[
\pi = p_0 + p_1, \quad f(\pi) = \frac{2 - \pi}{\pi}.
\]
### TABLE IIA
**ESTIMATES OF $f(\pi)$**

<table>
<thead>
<tr>
<th>Data series</th>
<th>Crimes per capita ($p$) times 1-crimes per capita ($1 - p$)</th>
<th>Sample variance $p(1 - p)$</th>
<th>Estimated $\lambda^2$</th>
<th>Estimated $f(\pi)$</th>
<th>Estimated $f(\pi)$ $\lambda^2 = .008$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serious crime</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>0.073</td>
<td>1313.8</td>
<td>0.013</td>
<td>754.6</td>
<td>604.7</td>
</tr>
<tr>
<td>N = 658</td>
<td></td>
<td></td>
<td>(0.003)</td>
<td>(118.2)</td>
<td></td>
</tr>
<tr>
<td>1970</td>
<td>0.042</td>
<td>1045.5</td>
<td>0.004</td>
<td>475.1</td>
<td>284.3</td>
</tr>
<tr>
<td>N = 617</td>
<td></td>
<td></td>
<td>(0.001)</td>
<td>(42.5)</td>
<td></td>
</tr>
<tr>
<td>NYC</td>
<td>0.053</td>
<td>575.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N = 70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1986</td>
<td>0.078</td>
<td>1500.0</td>
<td>0.0003</td>
<td>155.0</td>
<td>73.2</td>
</tr>
<tr>
<td>N = 631</td>
<td></td>
<td></td>
<td>(0.0015)</td>
<td>(58.5)</td>
<td></td>
</tr>
</tbody>
</table>
Empirical problems

- Unobserved correlated shocks
- Endogeneity of the network
- Distinguishing endogenous and contextual effects
Plan

- Network Science
- Labor Markets — Weak and Strong Ties
- Peer Effects and Complementarities — Games on Networks
- Matching and Network Formation
- Social Capital
- Social Learning
- Diffusion
Network Science
A directed graph $G$ is a pair $(V, E)$ where $V$ is a set of vertices, or nodes, and $E$ is a set of Edges. An edge is an ordered pair $(v, w)$, meaning that there is a connection from $v$ to $w$. If $(w, v) \in E$ whenever $(v, w) \in E$, then $G$ is an undirected graph.

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

A path of $G$ 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 $v_n$. 
A subset of vertices is connected if there is a path between every two of them. A component of $G$ 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 $G$ 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 $E$. 
Graphs

- 3 Components, \{A, B\}, \{C, D, E\}, \{F, \ldots, M\}.
- Min degree = 1.
- Max deg = 4.
- Diam Large Comp. = 3.
- Degree Dist. Large Comp.
  - 1 : 4 / 13
  - 2 : 4 / 13
  - 3 : 4 / 13
  - 4 : 1 / 13.
Common Network Measurements

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

Panel A: Core Infection Model

Panel B: Inverse Core Model

Panel C: Bridge Between Disjoint Populations

Panel D: Spanning Tree
Some Social Networks
Some Social Networks


\[ n \rightarrow \text{# nodes}, \quad m \rightarrow \text{# edges}, \quad z \rightarrow \text{mean degree,} \]
\[ l \rightarrow \text{mean geodesic length,} \quad \alpha \rightarrow \text{exponent of degree dist.,} \]
\[ C^{(k)} \rightarrow \text{clustering coeff.s,} \quad r \rightarrow \text{degree corr. coeff.} \]
Comparison: Erdös-Rényi Random Graphs

Every possible \((v, w)\) edge is assigned to \(E\) with probability \(p\).

Poisson random graphs: A sequence of graphs \(G_n\) with \(|V_n| = n\) such that \(p \cdot (n - 1) \to z\).

Large \(n\) facts:

- Phase transition at \(z = 1\).
- Low-density: Exponential component size distribution with a finite limit mean.
- High-density: a giant connected component of size \(O(n)\). Remainder size distribution exponential . . . .
- Clustering coefficient is \(C^2 = O(n^{-1})\).

Simulation of Erdös-Rényi random sets on 300 nodes.
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?
Centrality

Which nodes are important?

- Degree Centrality: The centrality of a node is its (in/out) degree.

- Katz (1953) Centrality: How many nodes can node \( i \) reach?

\[
c_i(\alpha) = \sum_k \sum_j \alpha^k (A^k)_{ij}.
\]

\( A^k_{ij} \) is the number of paths of length \( k \) from \( i \) to \( j \). The parameter \( \alpha \) discounts longer paths.
Eigenvector Centrality: Suppose that in the adjency matrix, $a_{ij} = 1$ if $j$ influences $i$, and 0 otherwise. The centrality index of $j$ is proportional to the sum of the centrality indices of the people she influences. so

$$c_j = \mu \sum_i c_i a_{ij}$$

where $\mu > 0$. If the network is connected, then (Perron Frobenius Theorem) there is a unique scalar $\mu$ and a one-dimensional set of vectors $c \geq 0$ that solve this. $\mu$ is the inverse of the Perron eigenvalue, and $c$ is in the corresponding eigenspace. (Bonacich 1972a,b, 1987).
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 beliefs.

- **Opportunity Homophily**: We mostly meet people like us.
Sources of Homophily

- Fixed attributes
  - Selection

- Variable attributes
  - Social influence

- Identification
Consider a network with $N$ individuals: Fraction $p$ are males, fraction $q = 1 - p$ are females.

- Assign nodes to gender randomly, each node male with probability $p$.
- What is the probability of a “cross-gender” edge?
Measuring Homophily

Consider a network with \( N \) individuals: Fraction \( p \) are males, fraction \( q = 1 - p \) are females.

- Assign nodes to gender randomly, each node male with probability \( p \).
- What is the probability of a “cross-gender” edge?
- A fraction of cross-gender edges less than \( 2pq \) is evidence for homophily.
“Arbitrarily selected individuals (N=296) in Nebraska and Boston are asked to generate acquaintance chains to a target person in Massachusetts, employing “the small world method” (Milgram, 1967). Sixty-four chains reach the target person. Within this group the mean number of intermediaries between starters and targets is 5.2. Boston starting chains reach the target person with fewer intermediaries than those starting in Nebraska; subpopulations in the Nebraska group do not differ among themselves. The funneling of chains through sociometric “stars” is noted, with 48 per cent of the chains passing through three persons before reaching the target. Applications of the method to studies of large scale social structure are discussed.”

Travers and Milgram (1969)
Small Worlds

Watts-Strogatz Model

Homophily

+ 

Weak Ties
Small Worlds

Watts-Strogatz Model

Homophily

+ 

Weak Ties
Small Worlds

Watts-Strogatz Model

Homophily

+ 

Weak Ties
My Wife: “What a suprise meeting you here. The world is indeed small.”

Friend: “No, it’s very stratified.”
Labor Markets
Inequality in Labor Markets

The chart illustrates the percent wage gap between men and women from 1960 to 2010. The graph shows a steady increase in the wage gap over time, with men consistently having a higher wage gap compared to women.
Inequality in Labor Markets

Table 2: Changes in the 90-50 and 50-10 Wage Gaps

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th></th>
<th>Male</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>90-50</strong></td>
<td><strong>50-10</strong></td>
<td><strong>90-50</strong></td>
<td><strong>50-10</strong></td>
</tr>
<tr>
<td>Total change</td>
<td>.1035</td>
<td>.0886</td>
<td>.0660</td>
<td>-.0774</td>
</tr>
<tr>
<td>Due to wage dispersion within occupations</td>
<td>.0778</td>
<td>.0507</td>
<td>.0342</td>
<td>-.0537</td>
</tr>
<tr>
<td>Due to wage changes between occupations</td>
<td>.0588</td>
<td>.0257</td>
<td>.0184</td>
<td>-.0154</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td></td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>90-50</strong></td>
<td><strong>50-10</strong></td>
<td><strong>90-50</strong></td>
<td><strong>50-10</strong></td>
</tr>
<tr>
<td>Total change</td>
<td>.0592</td>
<td>.1526</td>
<td>.0368</td>
<td>.0153</td>
</tr>
<tr>
<td>Due to wage dispersion within occupations</td>
<td>.0292</td>
<td>.1074</td>
<td>.0343</td>
<td>.0044</td>
</tr>
<tr>
<td>Due to wage changes between occupations</td>
<td>.0417</td>
<td>.0660</td>
<td>.0103</td>
<td>-.0045</td>
</tr>
</tbody>
</table>
Table 1—Job-Finding Methods Used by Workers

<table>
<thead>
<tr>
<th>Source/data</th>
<th>Percentage of jobs found using each method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Friends/relatives</td>
</tr>
<tr>
<td>Myers and Shultz (1951)/sample of displaced textile workers:</td>
<td></td>
</tr>
<tr>
<td>First job</td>
<td>62</td>
</tr>
<tr>
<td>Mill job</td>
<td>56</td>
</tr>
<tr>
<td>Present job</td>
<td>36</td>
</tr>
<tr>
<td>Rees and Shultz (1970)/Chicago labor-market study, 12 occupations:*</td>
<td></td>
</tr>
<tr>
<td>Typist</td>
<td>37.3</td>
</tr>
<tr>
<td>Keypunch operator</td>
<td>35.3</td>
</tr>
<tr>
<td>Accountant</td>
<td>23.5</td>
</tr>
<tr>
<td>Tab operator</td>
<td>37.9</td>
</tr>
<tr>
<td>Material handler</td>
<td>73.8</td>
</tr>
<tr>
<td>Janitor</td>
<td>65.5</td>
</tr>
<tr>
<td>Janitress</td>
<td>63.6</td>
</tr>
<tr>
<td>Fork-lift operator</td>
<td>66.7</td>
</tr>
<tr>
<td>Punch-press operator</td>
<td>65.4</td>
</tr>
<tr>
<td>Truck operator</td>
<td>56.8</td>
</tr>
<tr>
<td>Maintenance electrician</td>
<td>57.4</td>
</tr>
<tr>
<td>Tool and die maker</td>
<td>53.6</td>
</tr>
<tr>
<td>Granovetter (1974)/sample of residents of Newton, MA:</td>
<td></td>
</tr>
<tr>
<td>Professional</td>
<td>56.1</td>
</tr>
<tr>
<td>Technical</td>
<td>43.5</td>
</tr>
<tr>
<td>Managerial</td>
<td>65.4</td>
</tr>
<tr>
<td>Corcoran et al. (1980)/Panel Study of Income Dynamics, 11th wave:</td>
<td></td>
</tr>
<tr>
<td>White males</td>
<td>52.0</td>
</tr>
<tr>
<td>White females</td>
<td>47.1</td>
</tr>
<tr>
<td>Black males</td>
<td>58.5</td>
</tr>
<tr>
<td>Black females</td>
<td>43.0</td>
</tr>
</tbody>
</table>

*aMost of these workers were rehired at a previous mill job or hired at a new mill established in one of the vacated mills.

*bIn 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.

*cAgencies and ads are combined under the heading "formal means."

*dGate 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)
Two cliques.
Why do Weak Ties Matter?

Two cliques.

$A-B$ is a bridge.
Two cliques.

$A - B$ is a **bridge**.

*Local bridge’s endpoints have no common friends.*
Why do Weak Ties Matter?

Two cliques.

A–B is a **bridge**.

*Local bridge*’s endpoints have no common friends.

**Triadic closure**: A length-2 path containing only strong edges is a closed triad.
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. $\pi = \text{employee productivity} - \text{wage}$. Free entry, risk-neutral entrepreneurs.

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

- $\tau$ — “network density”
- $\alpha$ — “inbreeding bias”
Firms hire period 1 workers through the anonymous market, clears at wage $w_{m1}$.

Production occurs. 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

Comparative Statics

\[ \alpha, \tau \uparrow \quad \implies \quad \begin{cases} w_{m2} \downarrow \\ w_R \uparrow \\ \pi_2 \uparrow \\ w_{m1} \uparrow \end{cases} \]
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.

...Using pooled data from three cross-sectional surveys in urban China, the results show a steady increase in the use of weak ties and an increasing and persistent use of strong ties in finding jobs between 1978 and 2008. The results also show no systematic difference between the use of weak ties for finding jobs in the market sector versus the state sector. However, they show faster growth in the use of strong ties for finding jobs in the state sector, compared to the market sector.
Network Structure and Inequality

- Dynamic Markov model
- Illustrate how network structure matters
Network Structure and Inequality Model

- Discrete time.

- $N$ individuals.

- Symmetric adjacency matrix $A$.

- A configuration of the model is a map $s : \{1, \ldots, N\} \rightarrow \{0, 1\}$. Interpretation: 0 is unemployed, 1 is employed.

- $p$ is the probability that an individual learns about a job opening.
1. With probability $p + q \leq 1$, a *job event* happens.
   - With probability $qk / N$ one of the $k$ employed individuals loses her job.
   - With probability $p$ a single randomly chosen individual learns about a job.

2. If she is unemployed, she takes the job.

3. If she is employed, she passes the offer on to an unemployed neighbor, chosen at random.

4. If all neighbors are employed, the referral dies.
In any period, the configuration can change in one of three ways:

- A 0 can change to a 1;
- A 1 can change to a 0;
- The configuration can remain unchanged.

\[
\Pr\{s_{t+1}(i) = 1 \mid s_t(i) = 0, s_t(-i)\} = \frac{p}{N} \left(1 + \sum_j a_{ij} s_t(j) \frac{1}{\sum_k a_{jk} s_t(k)}\right)
\]

\[
\Pr\{s_{t+1}(i) = 0 \mid s_t(i) = 1, s_t(-i)\} = \frac{q}{N}
\]
\[ \text{Cov}(s_{t+1}(1), s_{t+1}(3) \mid s_t = (0, 1, 0)) = \] 

\[ E\{s_{t+1}(1) \cdot s_{t+1}(3) \mid (0, 1, 0)_t\} \]

\[ - E\{s_{t+1}(1) \mid (0, 1, 0)\} \cdot E\{s_{t+1}(3) \mid (0, 1, 0)_t\} = -\frac{p^2}{N^2} \]
Equilibrium is an invariant distribution of the Markov chain.

The transition matrix is irreducible, so the invariant distribution $\mu$ is unique!

$\text{Cov}_\mu(s(i), s(j)) \geq 0.$

$\text{Cov}_\mu(s(i), s(j)) > 0$ if and only if $i$ and $j$ are in the same connected component.
Because of symmetry, this is a Markov process on the number of employed. $m_{ij}$ is the probability that $j$ workers will be employed tomorrow if $i$ workers are employed today.

\[
M = \begin{pmatrix}
1 - p & p & 0 \\
\frac{q}{2} & 1 - p - \frac{q}{2} & p \\
0 & q & 1 - q
\end{pmatrix}
\]

The invariant distribution is a probability distribution that solves

\[\rho M = \rho.\]

\[
\rho(0) = \frac{q^2}{\Delta} \quad \rho(1) = \frac{2pq}{\Delta} \quad \rho(2) = \frac{2p^2}{\Delta}.
\]
Network Structure and Inequality

\[
M = \begin{pmatrix}
1 - p & p & 0 \\
\frac{q}{2} & 1 - \frac{q}{2} - \frac{p}{2} & \frac{p}{2} \\
0 & q & 1 - q
\end{pmatrix}
\]

\[
\rho(0) = \frac{q^2}{\Delta} \quad \rho(1) = \frac{2pq}{\Delta} \quad \rho(2) = \frac{p^2}{\Delta}.
\]
Suppose that $emp = k$ out of $N$ individuals are employed after $t$ events.

$$\Pr\{emp_{t+1} = k + 1|emp_t = k\} = p,$$

$$\Pr\{emp_{t+1} = k - 1|emp_t = k\} = \frac{kq}{N}.$$

$$\frac{\rho(k + 1)}{\rho(k)} = \frac{Np}{kq} \quad \frac{\rho(k)}{\rho(0)} = \frac{N^k}{k!} \left(\frac{p}{q}\right)^k$$
Network Structure and Inequality

Pair of Cliques

Product distribution
Peer Effects and Complementarities
Behaviors on Networks
Three Types of Network Effects

- Information and social learning.
- Network externalities.
- Social norms.
A Common Regression

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

Where

- \( \omega_i \) 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
- \( \varepsilon_i \) is an unobserved individual effect.
LIM Model

The Reflection Problem

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

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

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

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

The reduced form is

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

\[ \omega_i = \beta' x_i + \delta' \sum_j c_{ij} x_j + \gamma' \sum_j a_{ij} E\{\omega_j|x\} + \eta_i \]

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 $(x_i, z_i) \in \mathbb{R}^2$. $x_i$ is publicly observable, $z_i$ is private.
- There is a Harsanyi prior $\rho$ on the space of types $\mathbb{R}^{2I}$.
- Actions are $\omega_i \in \mathbb{R}$.
- 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 $i$. 


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 \(i\).
Equilibrium

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

\[\Gamma \omega + \Delta x = \eta.\]

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 \((\gamma, \delta)\) 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 \(\gamma\) and \(\delta\).
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.
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 $x \succ_{y'} x'$.
Transferable Utility

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 $\pi_i$ and $\pi_j$.

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

$$w_i + \pi_g \geq v_{ig} \quad \text{and} \quad w_j + \pi_f \geq v_{jf}.$$
Find the optimal match by maximizing total surplus:

\[ v(L \cup F) = \max_x \sum_{l,f} v_{lf} x_{lf} \]

s.t. \[ \sum_f x_{lf} \leq 1 \quad \text{for all } l, \]

\[ \sum_l x_{lf} \leq 1 \quad \text{for all } f, \]

\[ x \geq 0 \]

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

The dual has variables for each individual and firm.

\[
\min_{w, \pi} \sum_{l,f} w_l + \pi_f \\
\text{s.t.} \quad \pi_f + w_l \geq v_{lf} \quad \text{for all pairs } l, f, \\
\pi \geq 0, \ w \geq 0.
\]

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 $w_l + \pi_f = v_{lf} = w'_l + \pi'_f$. 
Assortative Matching Increasing Differences

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

**Definition:** $v : X \times Y \to \mathbb{R}$ 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 $\mathbb{R}$, 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 $D_{xy} v(x, y) \geq 0$. 

The roommate problem.

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.” (Bourdieu 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)
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)
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 Model

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

$$c + e = y$$ parental budget constraint
Intergenerational Transfers Model

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 $\alpha$.

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

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

$$V^*(y) = \max_{0 \leq c \leq y} \mathbb{E}\left\{ U\left( c, V^*(h(\tilde{\alpha}, y - c)) \right) \right\}.$$
The Bellman equation has a unique solution, and there is a \( \bar{y} \) such that \( y \leq \bar{y} \) for all \( \alpha \).

The solution defines a Markov process of income.

\[
y \xrightarrow{e, \alpha} e, \alpha \xleftarrow{y} \nu
\]

\[
y \xrightarrow{h} \hat{y}
\]

\[
y \xrightarrow{\ldots}
\]

If education is a normal good, then the Markov process is ergodic, and the invariant distribution \( \mu \) has support on \([0, \hat{y}]\), where \( \hat{y} \) solves \( h(1, e^*(y)) = y \).
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 $\tau_1$ is more egalitarian than tax policy $\tau_2$ iff the distribution of income under $\tau_2$ is riskier than that of $\tau_1$ conditional on the education level $e$.

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

Three Stories about Trust:

**Reciprocity:** Reputation games, folk theorems, . . .

**Social Learning:** Generalized trust.

**Behavioral Theories:** Evolutionary Psychology, prosocial preferences, . . .
Inequality and Trust

- Evidence for a correlation between trust and income inequality

- Trust is correlated with optimism about one’s own life chances
  - Uslaner (2002)
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.
Knaak and Keefer (1997). “Does social capital have a payoff”.

\[ g_i = X_i \gamma + Z_i \pi + \text{CIVIC}_i \alpha + \text{TRUST}_i \beta + \epsilon_i \]

- \( g_i \) real per-capita growth rate.
- \( X_i \) control variables — Solow.
- \( Z_i \) control variables — “endogenous” growth models.
- \( \text{CIVIC}_i \) index of the level of civic cooperation.
- \( \text{TRUST}_i \) 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’.
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 $\pi$ is realized from the following payoff matrix:

<table>
<thead>
<tr>
<th></th>
<th>partner</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P$</td>
<td>$\tilde{R}$</td>
</tr>
<tr>
<td>$I$</td>
<td>$\tilde{e}$</td>
</tr>
</tbody>
</table>

Gross Returns
A Model of Trust

- $E\tilde{R} > E\tilde{e} > E\tilde{r}$.

- Individuals reinvest a constant fraction $\beta$ of their wealth.

- Strategies for $i$ are functions which map the history of $i$'s 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 $\rho$, about the probability of one’s opponent choosing $P$. The prior distribution is beta with parameters $a^i, b^i > 0$. In more detail,

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

Let $\rho_t^i$ denote individual $i$’s posterior beliefs after $t$ rounds of matching. The posterior densities $\rho_t^i$ and $\rho_t^j$ will be conditioned on different data, since all information is private. The updating rule for the $\beta$ 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 $\beta$ distribution is $a_t^i / (a_t^i + b_t^i)$. 

Optimal Play

\[ q^* = \frac{(e - r)}{(R - r)} \]

- Let \( m_t^i \) denote \( i \)'s mean of \( \rho_t \).

- An optimal strategy for individual \( i \) is: Choose \( P \) if \( m_t > q^* \) and choose \( I \) otherwise.

**Theorem 3:** For all initial beliefs \( (\rho_1^0, \ldots, \rho_N^0) \), 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 \( k_I \), and \( k_P < k_I \).
Social Learning
Averaging the Opinions of Others

- DeGroot (1974)

- $X$ is some event. $p_i(t)$ is the probability that $i$ assigns to the occurrence of $X$ at time $t$.

- $M$ is a stochastic matrix. $m_{ij}$ is the weight $i$ gives to $j$’s opinion.

- $p(t) = Mp(t-1) = \cdots = M^t p(0)$. 
Averaging the Opinions of Others

Example

\[
M = \begin{pmatrix}
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\
\frac{1}{2} & \frac{1}{2} & 0 \\
0 & \frac{1}{2} & \frac{1}{2}
\end{pmatrix},
\]

\[
p(2) = M^2 p(0) = \begin{pmatrix}
\frac{5}{18} & \frac{8}{18} & \frac{5}{18} \\
\frac{5}{12} & \frac{5}{12} & \frac{2}{12} \\
\frac{1}{4} & \frac{1}{2} & \frac{1}{4}
\end{pmatrix} p(0),
\]

\[
p(t) = M^t p(0) \rightarrow \begin{pmatrix}
\frac{3}{9} & \frac{4}{9} & \frac{2}{9} \\
\frac{3}{9} & \frac{4}{9} & \frac{2}{9} \\
\frac{3}{9} & \frac{4}{9} & \frac{2}{9}
\end{pmatrix} p(0).
\]

\[
p_i(\infty) = \left(\frac{1}{9}\right)\left(3p_1(0) + 4p_2(0) + 2p_3(0)\right).
\]
Averaging the Opinions of Others

Distinct Limits

\[ M = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 2/3 & 1/3 \end{pmatrix} \]

\[ M^t \rightarrow \begin{pmatrix} 2/5 & 3/5 & 0 & 0 \\ 2/5 & 3/5 & 0 & 0 \\ 0 & 0 & 3/5 & 2/5 \\ 0 & 0 & 3/5 & 2/5 \end{pmatrix} \]

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

\[ p_i(t) \rightarrow \frac{1}{5}(3p_3(0) + 2p_4(0)) \quad \text{for } i = 3, 4. \]
Averaging the Opinions of Others

No Limit

\[ M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \]

\[ M^t = M^{(t-1) \text{mod} 3 + 1} \]
Averaging the Opinions of Others

Convergence

**Theorem**: If $M$ is irreducible and aperiodic, then beliefs converge to a limit probability. $\lim_{t \to \infty} p(t) = \sum_i \pi_i p_i(0)$, where $\pi$ is the left Perron eigenvector of $M$.

Connection to Markov processes.
Averaging the Opinions of Others

Social influence

Influential individuals are those who influence other influential individuals. We want to measure this by a scalar $s_i$ for each individual $i$.

**Definition:** The Bonacich (eigenvector) centrality of individual $j$ is the average of the social influences of those he influences, weighted by the amount he influences them (Bonacich, 1987).

Then $s$ solves

$$s_j = \sum_i m_{ij} s_i,$$

Thus $s$ is the left Perron eigenvector of $M$, and so $s = \pi$. 

Suppose that $p_i(0) = p + \epsilon_i$. The $\epsilon_i$ 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 $i$, $s^n_i \leq B/n$, then the sequence learns.

What conditions on the networks guarantee this?
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 $E \left\{ \sum_{\tau=t}^{\infty} \beta^{\tau} \pi_{i\tau} | h_t \right\}$, 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 $\mu_i$ 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?
A probability space.

A finite set of actions.

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

If $f$ is a measurable mapping of $\Omega$ into any measure space, $\sigma f$ is the $\sigma$-algebra generated by $f$. Define $\sigma(y_k) = Y_k$.

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

\[ F_k(t + 1) = F_k(t) \lor \bigvee_{j \neq k} \sigma d\left(F_j(t)\right), \]

\[ F_k(0) = \mathcal{Y}_k. \]

**Key Property:** If \( \sigma d(\mathcal{G}) \subset \mathcal{H} \subset \mathcal{G} \), then \( d(\mathcal{G}) = d(\mathcal{H}) \).
Theorem: Suppose $d$ has the key property. Then there are $\sigma$-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.
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 \(k\).
Updating of beliefs:

\[ F_k(t + 1) = F_k(t) \lor \bigvee_{j \sim k} \sigma d(F_j(t)), \]

\[ F_k(0) = Y_k. \]

**Theorem:** Suppose \( d \) has the key network property. Then there are \( \sigma \)-algebras \( F_k \subset \bigvee_k Y_k \) and \( T \geq 0 \) such that \( F_k(t) = F_k \) for all \( t \geq T \), and for all \( k \) and \( j \),

\[ d(F_k) = d(F_j) = d\left( \bigwedge_i F_i \right). \]
Diffusion
Network Effects and Diffusion

Panel A: Core Infection Model

Panel B: Inverse Core Model

Panel C: Bridge Between Disjoint Populations

Panel D: Spanning Tree
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, QJE (1992)
  - Bikchandani, Hershleifer and Welch (1992)
  - Rumors
Coordination Games

\[
\begin{array}{c|cc}
 & A & B \\
\hline
A & a,a & 0,0 \\
B & 0,0 & b,b \\
\end{array}
\]

Pure coordination game

Three equilibria:

\[\langle a, a \rangle, \langle b, b \rangle, \text{ and } \langle \left( \frac{b}{a+b}, \frac{a}{a+b} \right), \left( \frac{b}{a+b}, \frac{a}{a+b} \right) \rangle\]
Coordination Games

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>a,a</td>
<td>0,0</td>
</tr>
<tr>
<td>B</td>
<td>0,0</td>
<td>b,b</td>
</tr>
</tbody>
</table>

$A, B > 0$

Pure coordination game

Best response dynamics

$\frac{a}{a+b}$
Coordination Games

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>a,a</td>
<td>d,c</td>
</tr>
<tr>
<td>B</td>
<td>c,d</td>
<td>b,b</td>
</tr>
</tbody>
</table>

$a > c, b > d$

General coordination game

Here the symmetric mixed equilibrium is at

$$p^* = \frac{(b - d)}{(a - c + b - d)}.$$  

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 $b - d < a - c$. 

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 \(B\).

In both cases, as \(\text{Prob}\{\text{best response} \uparrow 1\}\), \(\text{Prob}\{N\} \uparrow 1\).
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.
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?
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 — 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?
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?
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?
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?
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.
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.
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 $O(\epsilon^q N)$, where $q$ is the indifference threshold. On a network, the probability of transiting from all $A$ to all $B$ is on the order of $\epsilon^K$, 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.
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 $O(\epsilon^q N)$, where $q$ is the indifference threshold. On a network, the probability of transiting from all $A$ to all $B$ is on the order of $\epsilon^K$, 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.
Community Structure

Under Construction
Imagine a social network, such as a friendship network in a school or network of information sharing in a village. Suppose the network participants represent several ethnic groups, races or tribes.

▶ How “integrated” is the network with respect to predefined communities?

▶ What are the implicit “communities” of highly mutually interactive neighbors?

▶ How do these community structures map onto each other?
Measuring Segregation

Attributes of physical segregation.

- **Evenness** — Differential distribution of two groups across the network.
- **Exposure** — The degree to which different groups are in contact.
- **Concentration** — Relative concentration of physical space occupied by different groups.
Measuring Segregation

Attributes of physical segregation.

- Centralization — Extent to which a group is near the center.
- Clustering — Degree to which group members are connected to others in the group.
A city is divided into $N$ areas. Area $i$ has minority population $m_i$ and majority population $M_i$. Total populations are $m$ and $M$, respectively.

\[
\text{dissimilarity index} = \frac{1}{2} \sum_{i=1}^{N} \left| \frac{m_i}{m} - \frac{M_i}{M} \right|.
\]
Incomplete


References: Network Science II


References: Social Learning


References: Diffusion


