Diffusion and Strategic Interaction on Social Networks

Leeat Yariv

Summer School in Algorithmic Game Theory, Part 2, 8.7.2012

The Big Questions

- How does the structure of networks impact outcomes:
  - In different locations within the network and across different network architectures
  - Static and dynamic

- How do networks form to begin with (given the interactions that occur over them)

Games on Networks

- g is network (in $\{0,1\}^{nxn}$):

$$\boldsymbol{g}_{ij} = \begin{cases} 1 & \boldsymbol{i} \neq \boldsymbol{j} \text{ connected} \\ 0 & \text{otherwise} \end{cases}$$

Games on Networks

- g is network (in $\{0,1\}^{nxn}$):

$$g_{ij} = \begin{cases} 1 & i \neq j \text{ connected} \\ 0 & \text{otherwise} \end{cases}$$

- $N_i(g)$ i's neighborhood,

$$N_i(g) = \{ j \mid g_{ij} = 1 \}$$

Games on Networks

- g is network (in $\{0,1\}^{nxn}$):

$$g_{ij} = \begin{cases} 1 & i \neq j \text{ connected} \\ 0 & \text{otherwise} \end{cases}$$

- $N_i(g)$ i's neighborhood,

$$N_i(g) = \{ j \mid g_{ij} = 1 \}$$

- $d_i(g)=|N_i(g)|$ i's degree

Games on Networks

- g is network (in $\{0,1\}^{nxn}$):

$$g_{ij} = \begin{cases} 1 & i \neq j \text{ connected} \\ 0 & \text{otherwise} \end{cases}$$

- $N_i(g)$ i's neighborhood,

$$N_i(g) = \{ j \mid g_{ij} = 1 \}$$

- $d_i(g)=|N_i(g)|$ i's degree

- Each player chooses an action in {0,1}

Payoff Structure: (today) Complements

- Payoffs depend only on the number of neighbors choosing 0 or 1.

Payoff Structure: (today) Complements

- Payoffs depend only on the number of neighbors choosing 0 or 1.

- normalize payoff of all neighbors choosing 0 to 0

Payoff Structure: (today) Complements

- Payoffs depend only on the number of neighbors choosing 0 or 1.

- normalize payoff of all neighbors choosing 0 to 0

- $v(d,x) - c_i$ payoff from choosing 1 if degree is d and a fraction x of neighbors choose 1
  - Increasing in x (positive externalities)

Payoff Structure: (today) Complements

- Payoffs depend only on the number of neighbors choosing 0 or 1.

- normalize payoff of all neighbors choosing 0 to 0

- $v(d,x) - c_i$ payoff from choosing 1 if degree is d and a fraction x of neighbors choose 1
  - Increasing in x

- c_i distributed according to H

Examples (payoff: v(d,x)-c)

- **Average Action**: v(d,x)=v(d)x= x
  (classic coordination games, choice of technology)

- **Total Number**: v(d,x)=v(d)x=dx
  (learn a new language, need partners to use new good or technology, need to hear about it to learn)

- **Critical Mass**: v(d,x)=0 for x up to some M/d and v(d,x)=1 above M/d
  (uprising, voting, …)

- **Decreasing**: v(d,x) declining in d
  (information aggregation, lower degree correlated with leaning towards adoption)

(today) Incomplete information case:

- g drawn from some set of networks G such that:
  - degrees of neighbors are independent
  - Probability of any node having degree d is p(d)
  - probability of given neighbor having degree d is P(d)=dp(d)/E(d)

Equilibrium as a fixed point:

- H(v(d,x)) is the percent of degree d types adopting action 1 if x is fraction of random neighbors adopting.

- Equilibrium corresponds to a fixed point:

$$\mathbf{x = \varphi(x) = \Sigma\, P(d)\, H(v(d,x))}$$

$$\mathbf{= \Sigma\, d\, p(d)\, H(v(d,x)) \,/\, E[d]}$$

Equilibrium as a fixed point:

- H(v(d,x)) is the percent of degree d types adopting action 1 if x is fraction of random neighbors adopting.

- Equilibrium corresponds to a fixed point:

$$\mathbf{x = \varphi(x) = \Sigma\, P(d)\, H(v(d,x))}$$

- **Fixed point exists**

Equilibrium as a fixed point:

- H(v(d,x)) is the percent of degree d types adopting action 1 if x is fraction of random neighbors adopting.
- Equilibrium corresponds to a fixed point:

$$\mathbf{x = \varphi(x) = \Sigma\, P(d)\, H(v(d,x))}$$

- **Fixed point exists**
- **If H(0)=0, x=0 is a fixed point**

Monotone Behavior

Observation 1:

In a game of incomplete information, every symmetric equilibrium is monotone

- nondecreasing in degree if v(d,x) is increasing in d
- nonincreasing in degree if v(d,x) is decreasing in d

Expected payoffs move in the same direction

Monotone Behavior

Intuition

- Symmetric equilibrium – a random neighbor has probability x of choosing 1, probability 1-x of choosing 0.

Monotone Behavior

Intuition

- Symmetric equilibrium – a random neighbor has probability x of choosing 1, probability 1-x of choosing 0.

- Consider agent of degree d+1
  - $v(d,x)$ nondecreasing → payoff from 1 is $v(d+1,x) \geq v(d,x)$.
  - $v(d,x)$ nonincreasing → payoff from 1 is $v(d+1,x) \leq v(d,x)$.

Diffusion

$$\mathbf{x = \varphi(x) = \Sigma\, P(d)\, H(v(d,x))}$$

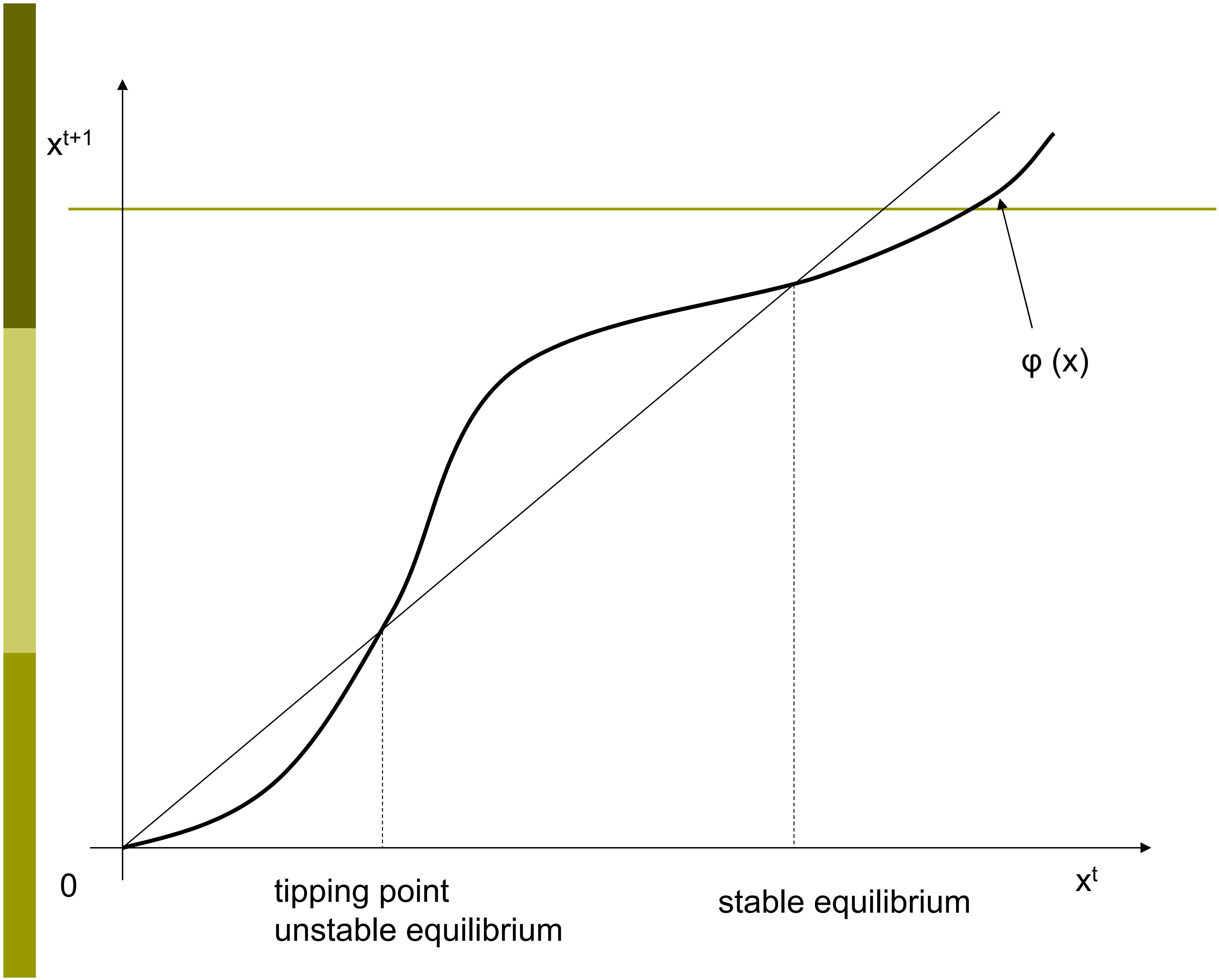
- start with some x^0
- let $x^1 = \varphi(x^0)$, $x^t = \varphi(x^{t-1})$, ...

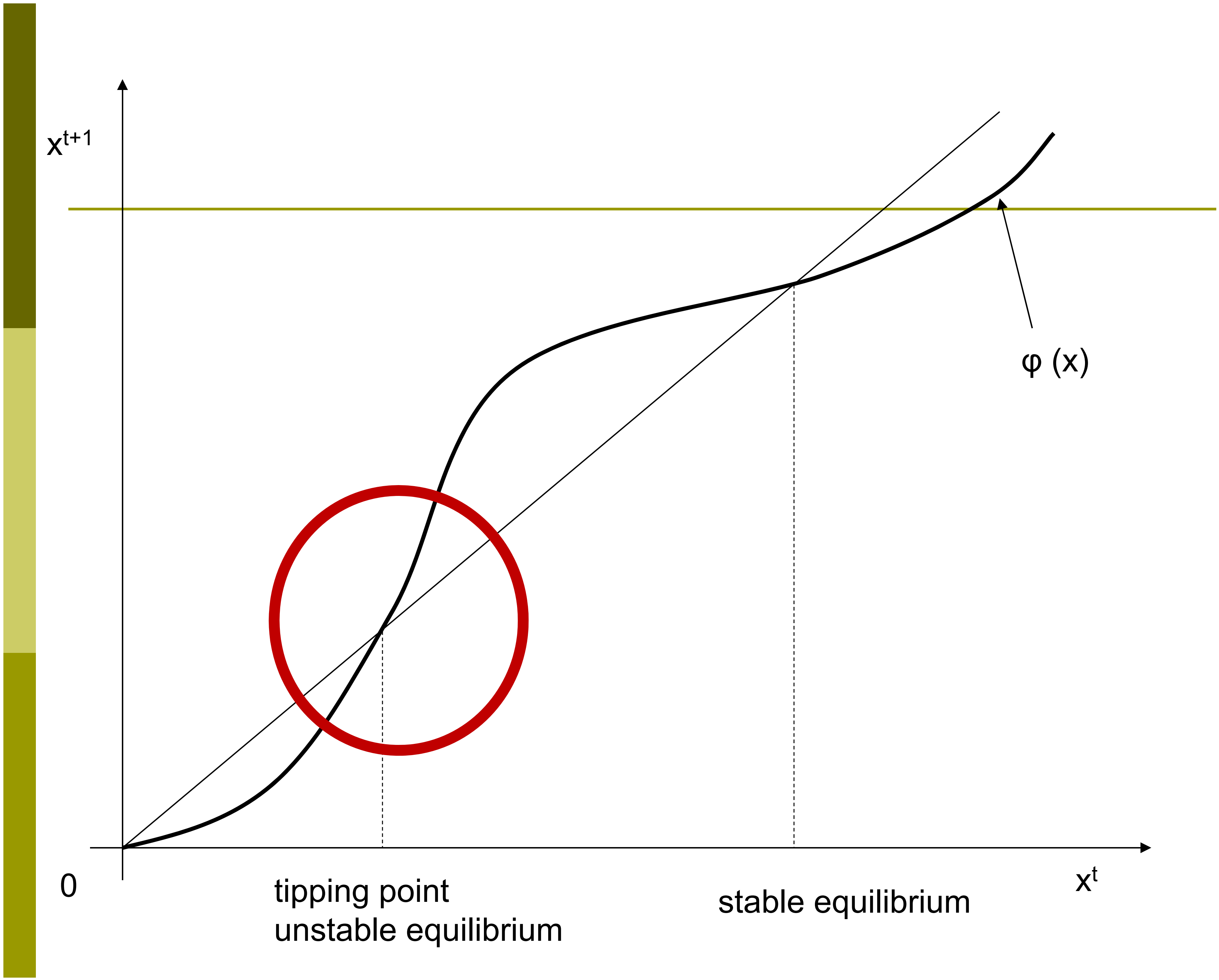
Diffusion

$$\mathbf{x = \varphi(x) = \Sigma\, P(d)\, H(v(d,x))}$$

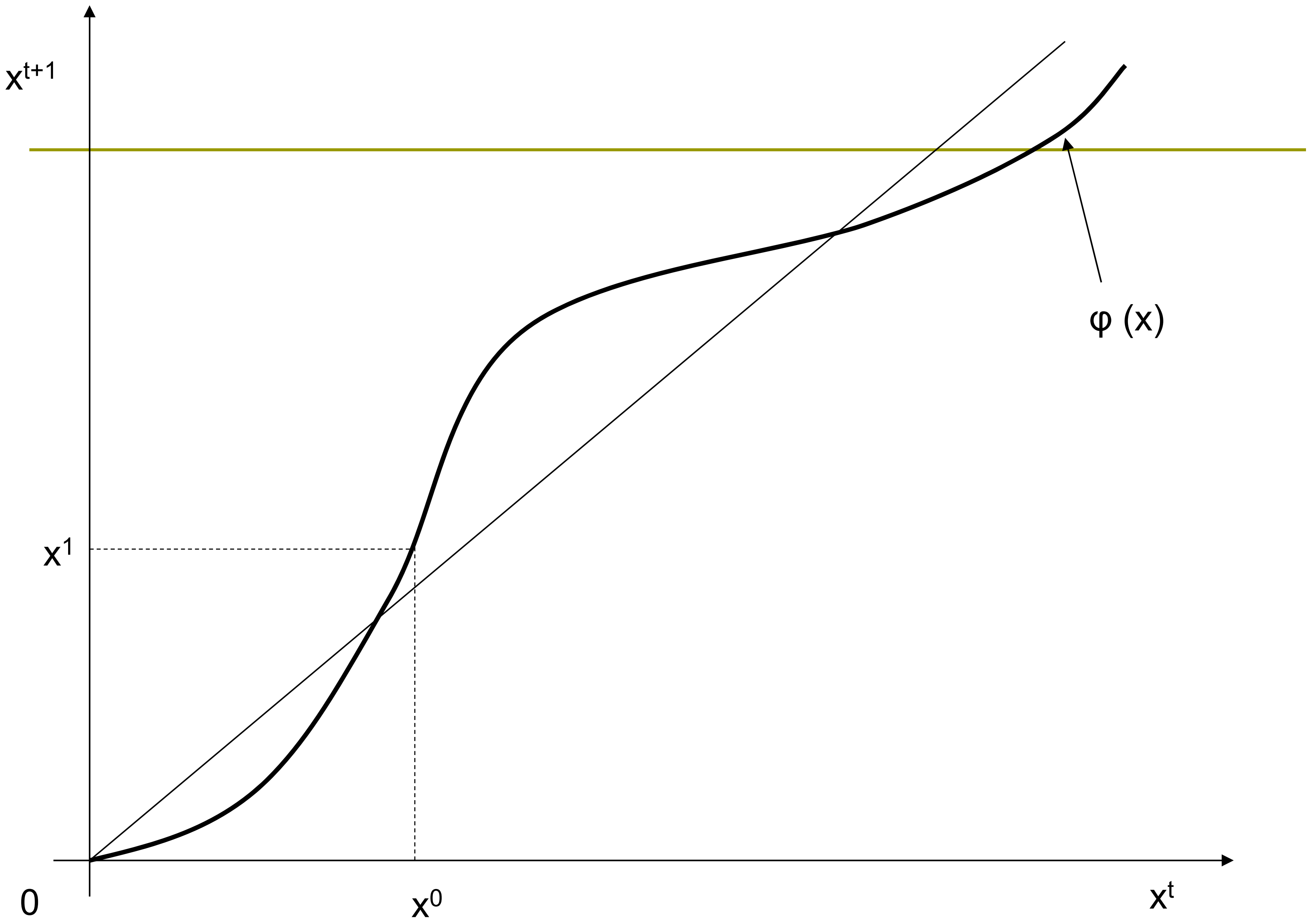
- start with some x^0
- let $x^1 = \varphi(x^0)$, $x^t = \varphi(x^{t-1})$, ...

- **Interpretations**
  - examining equilibrium set with incomplete information
    - Stable equilibria are converged to from above and below
  - looking at diffusion: complete information best response dynamics on "large, well-mixed" social network

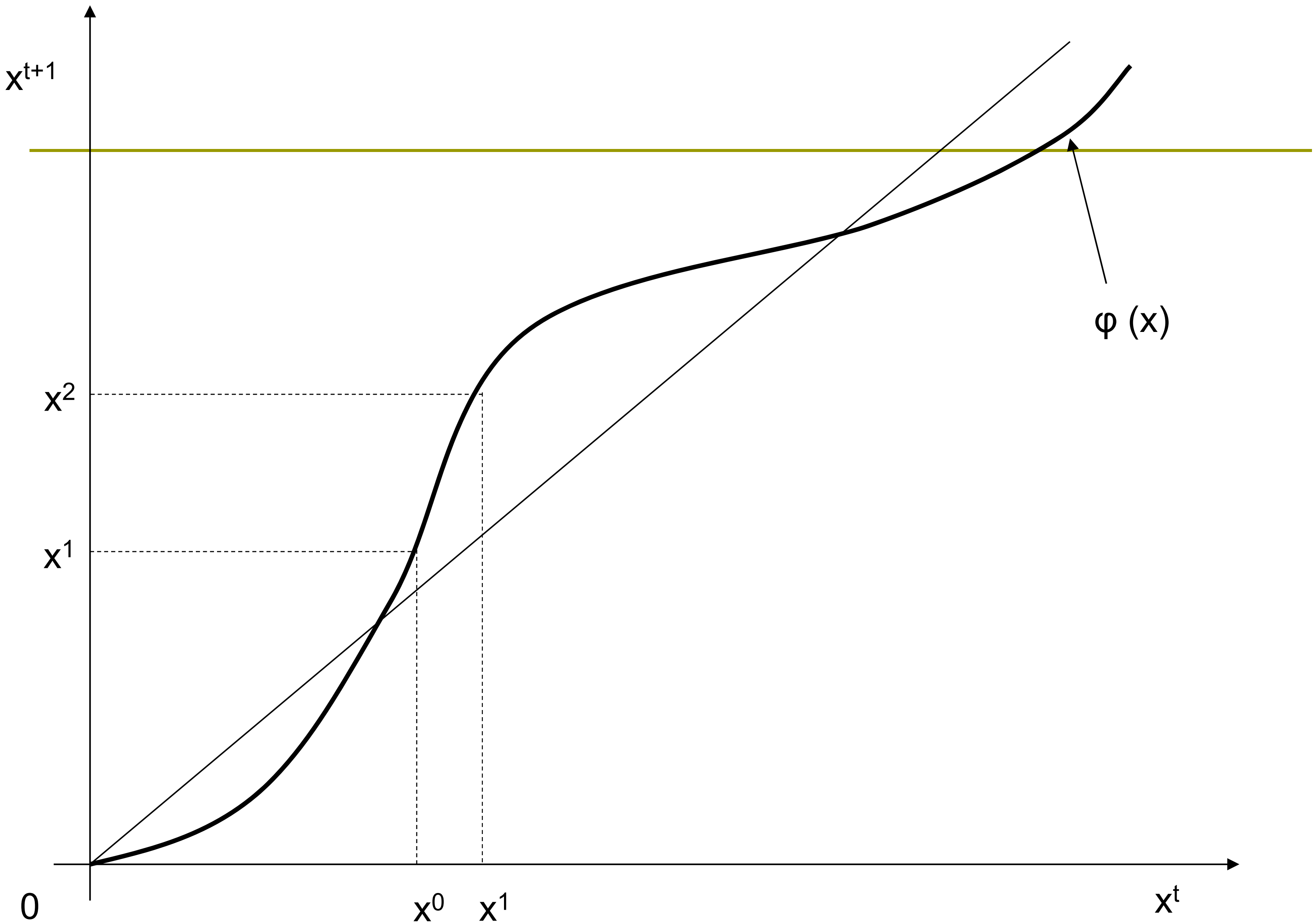

x^{t+1}
φ (x)
0
tipping point
unstable equilibrium
stable equilibrium
x^t



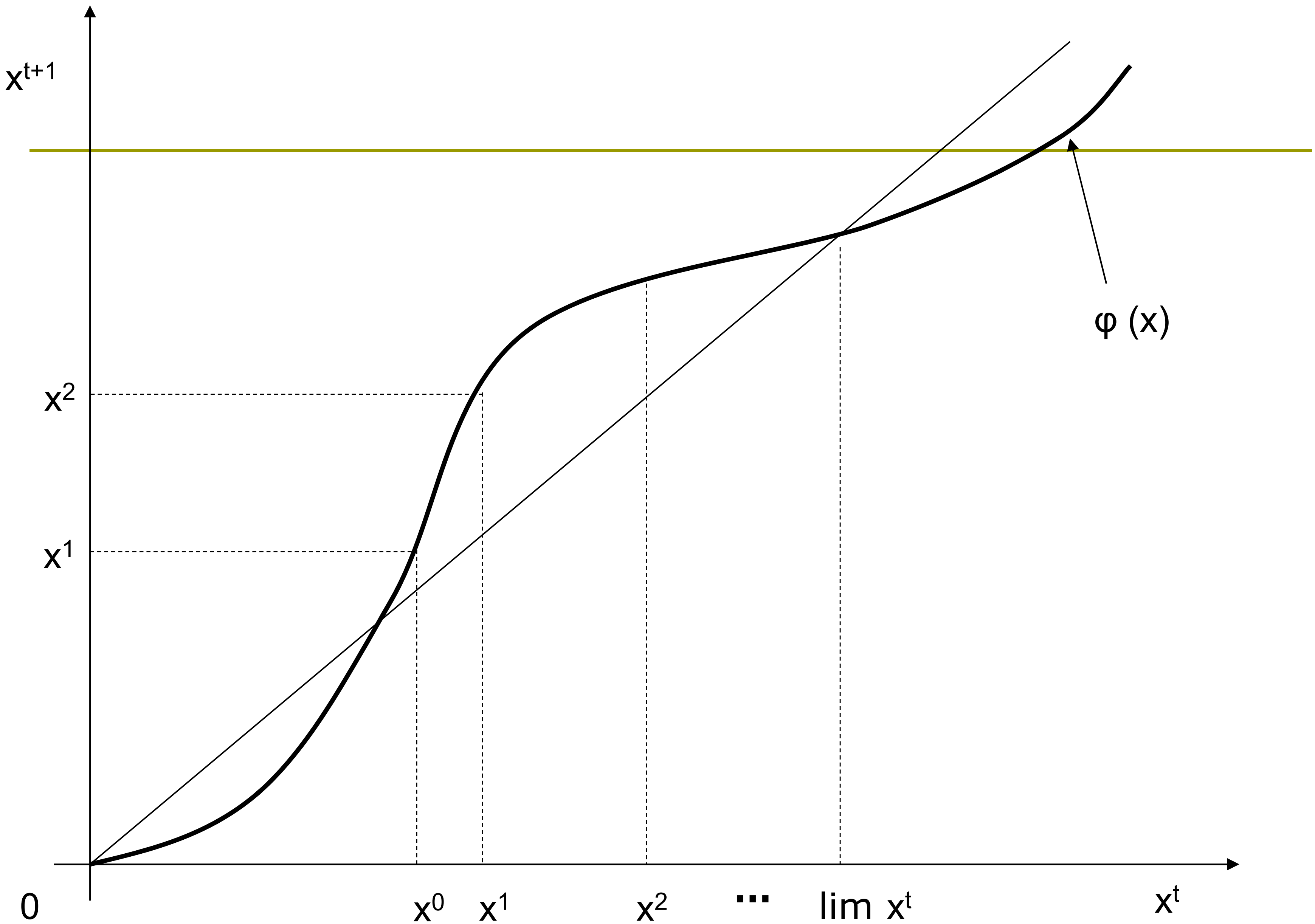
x^{t+1}
φ (x)
0
tipping point
unstable equilibrium
stable equilibrium
x^t



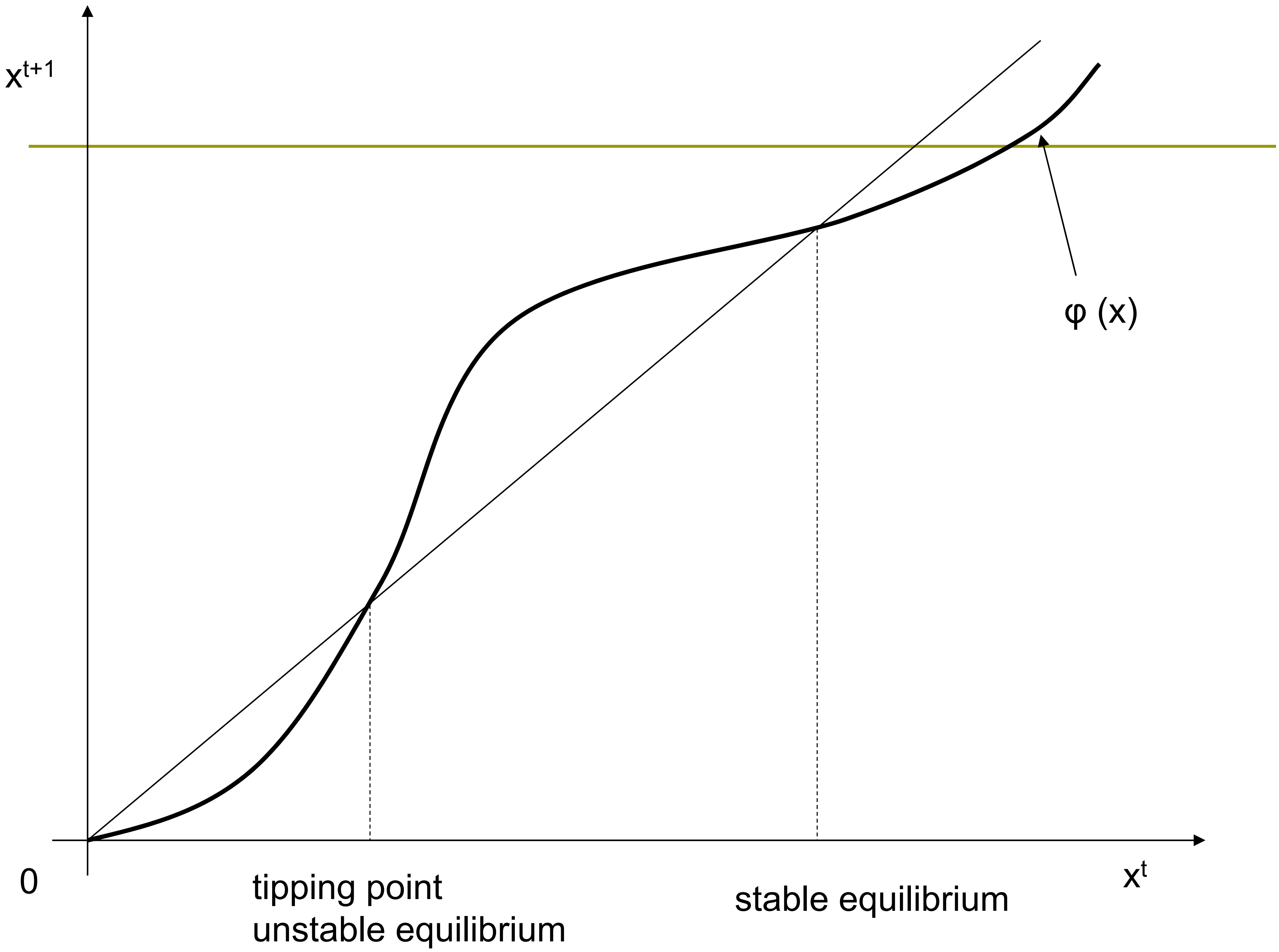
x^{t+1}
x^1
0
x^0
x^t
φ (x)



x^{t+1}
x^2
x^1
0
x^0
x^1
x^t
φ (x)



x^{t+1}
x^2
x^1
0
x^0
x^1
x^2
...
lim x^t
x^t
φ (x)

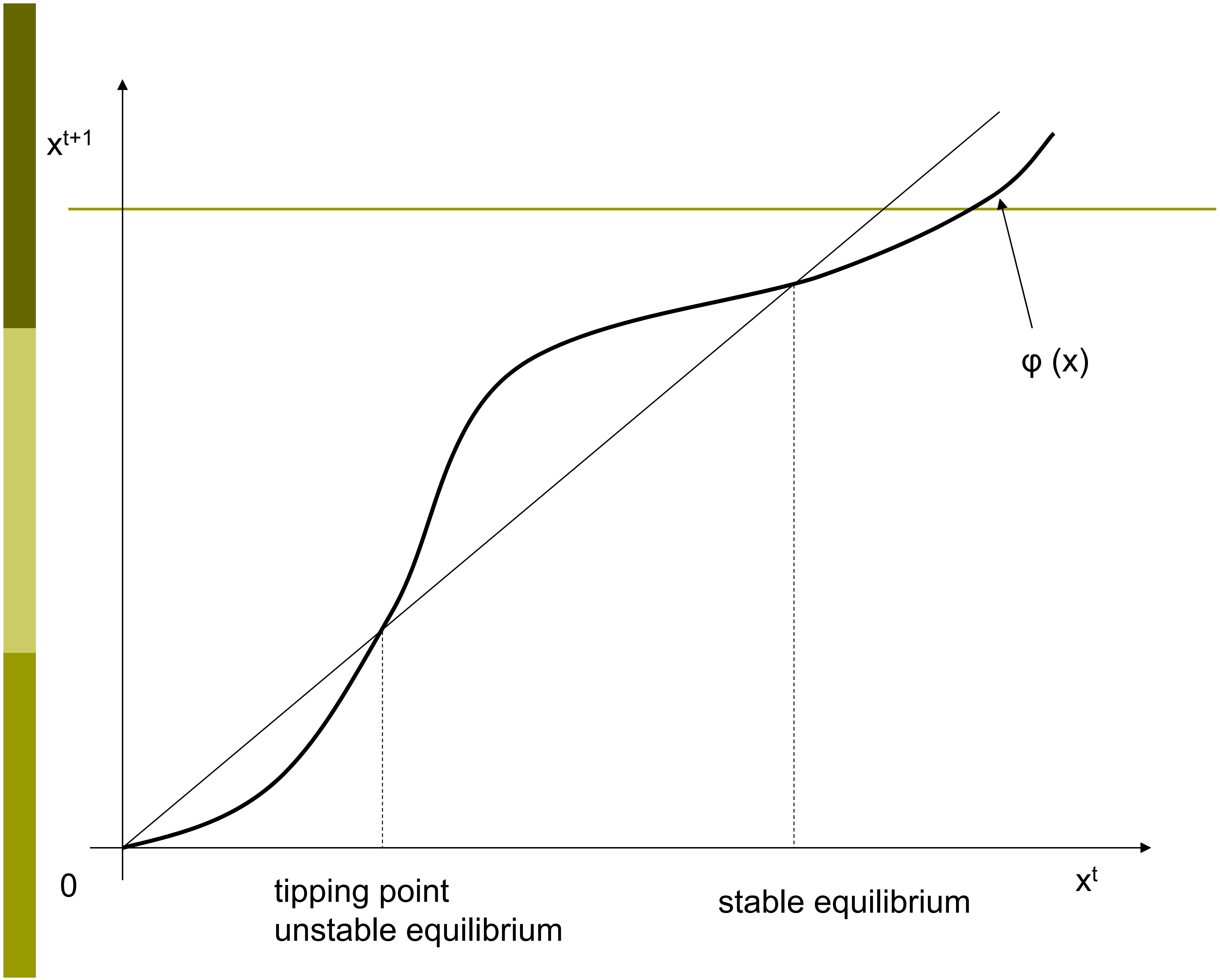


x^{t+1}
φ (x)
0
x^t
tipping point
unstable equilibrium
stable equilibrium

Stability at 0

$\varphi(x)<x$ in a neighborhood around 0 (joint condition on H, v(d,x), P(d))

If H is continuous, and 0 is stable, then "generically": next unstable (first tipping point, where volume of adopters grows), next is stable, etc.

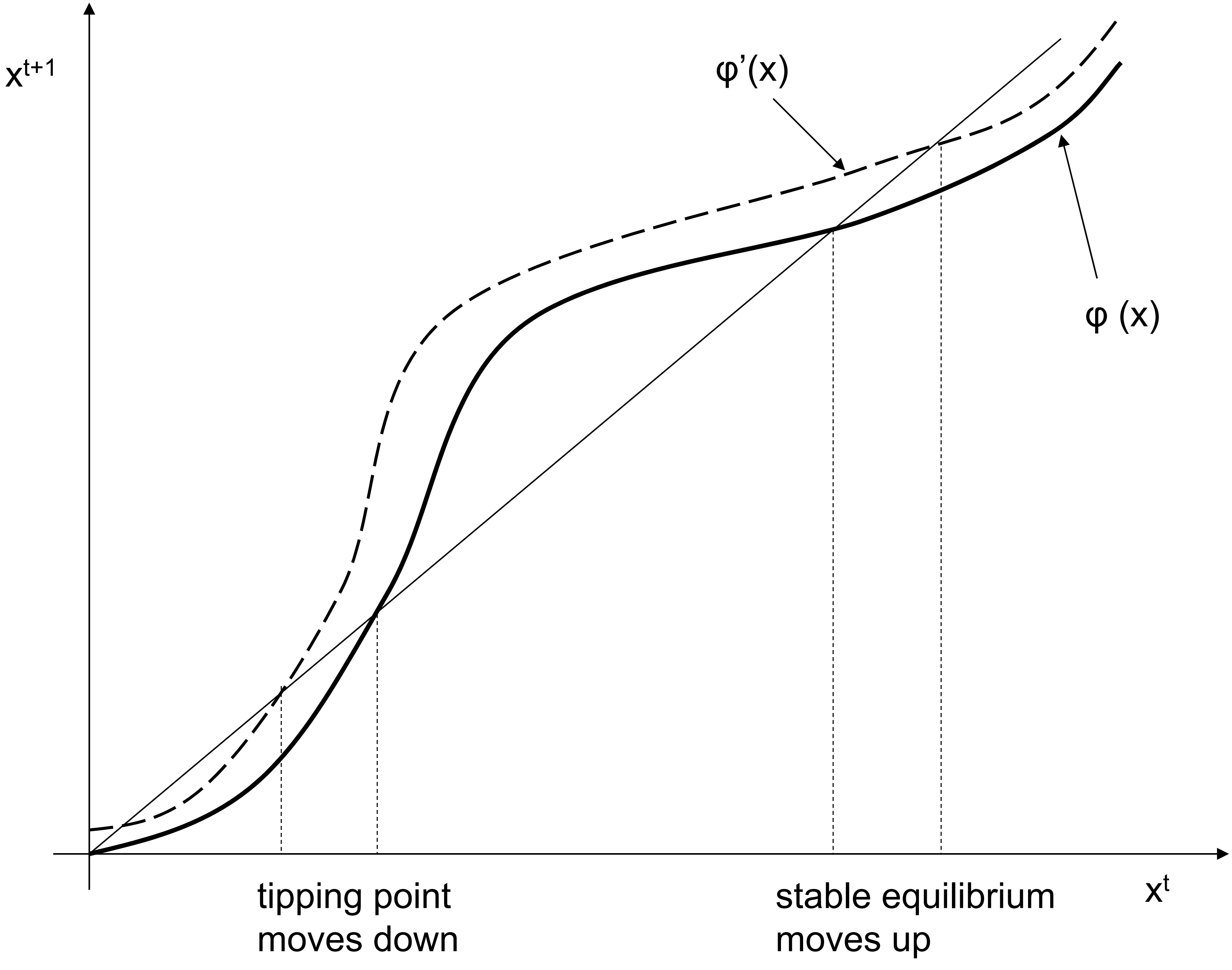


x^{t+1}
φ (x)
0
tipping point
unstable equilibrium
stable equilibrium
x^{t}

How can we relate structure (network or payoff) to diffusion?

- Keep track of how φ shifts with changes

  [concentrating on regular environments]



x^{t+1}
φ'(x)
φ (x)
tipping point
moves down
stable equilibrium
moves up
x^t

FOSD Shifts

- P(d) First Order Stochastically Dominates P'(d) if:

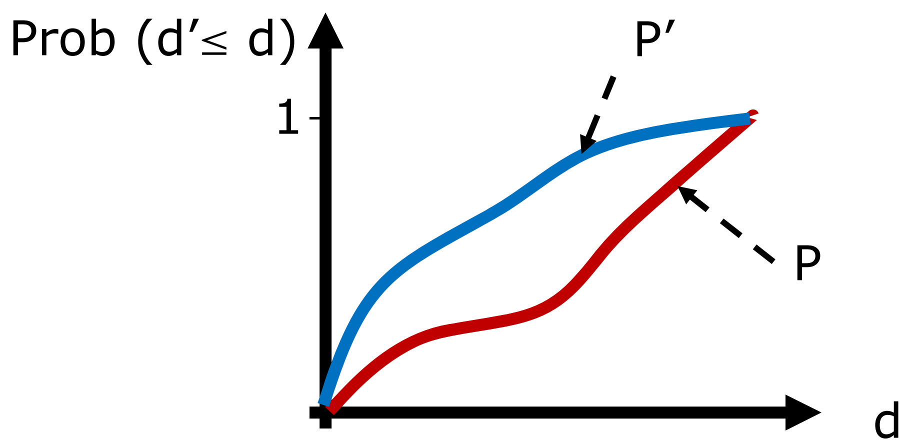
$$P(d) \leq P'(d) \quad \text{for all } d$$

FOSD Shifts

- P(d) First Order Stochastically Dominates P′(d) if:

$$P(d) \leq P'(d) \quad \text{for all } d$$

- P puts more weight on higher degrees.



FOSD Shifts

- P(d) First Order Stochastically Dominates P'(d) if:

$$P(d) \leq P'(d) \quad \text{for all } d$$

- For any increasing function f(d):

$$\sum_d f(d)P(d) \geq \sum_d f(d)P'(d)$$

Adding Links

- Consider a FOSD shift in distribution P(d)
  - More weight on higher degrees
  - v(d,x) nondecreasing in d ⇒ Higher expectations of higher actions (Observation 1)
  - More likely to take higher action

Adding Links

- Consider a FOSD shift in distribution P(d)
  - More weight on higher degrees
  - v(d,x) nondecreasing in d ⇒ Higher expectations of higher actions (Observation 1)
  - More likely to take higher action

- If v(d,x) is nondecreasing in d, then this leads to a pointwise increase of

$$\varphi(x) = \sum P(d)\, H(v(d,x))$$

Adding Links

- Consider a FOSD shift in distribution P(d)
  - More weight on higher degrees
  - v(d,x) nondecreasing in d ⇒ Higher expectations of higher actions (Observation 1)
  - More likely to take higher action

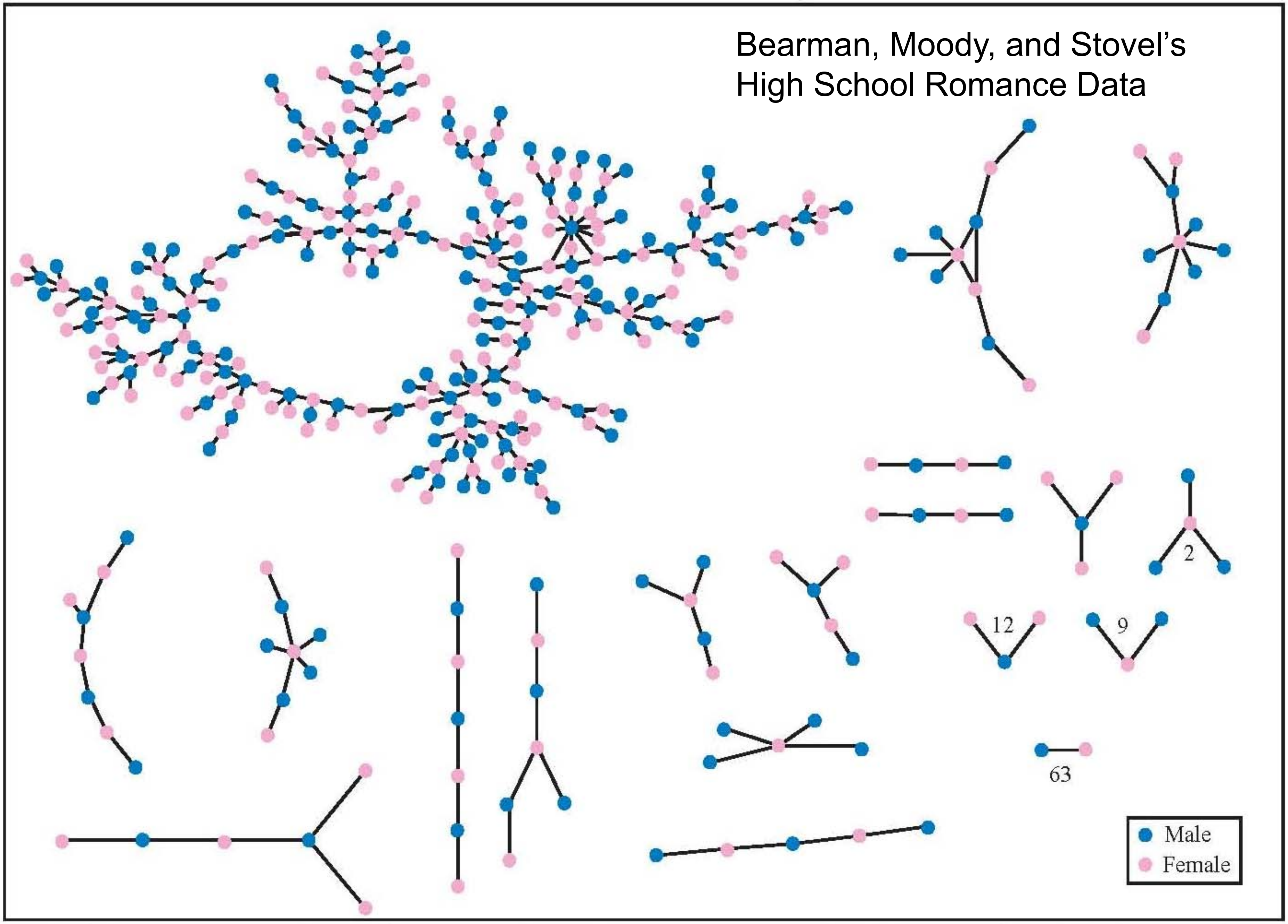
- If v(d,x) is nondecreasing in d, then this leads to a pointwise increase of

$$\varphi(x) = \sum P(d)\, H(v(d,x))$$

- lower tipping point and higher stable equilibrium

The Structure of Romantic and Sexual Relations at "Jefferson High School"

Bearman, Moody, and Stovel's High School Romance Data



Coauthorships and Poisson

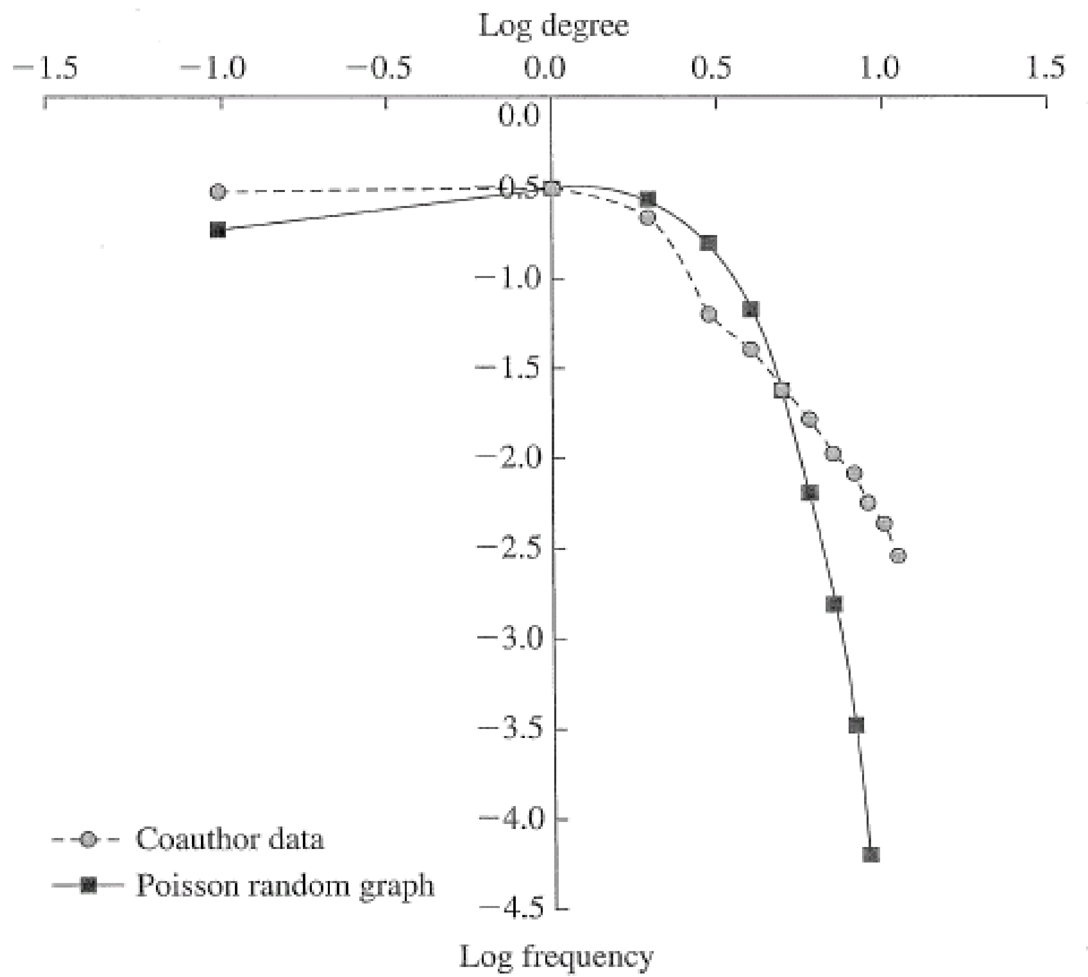
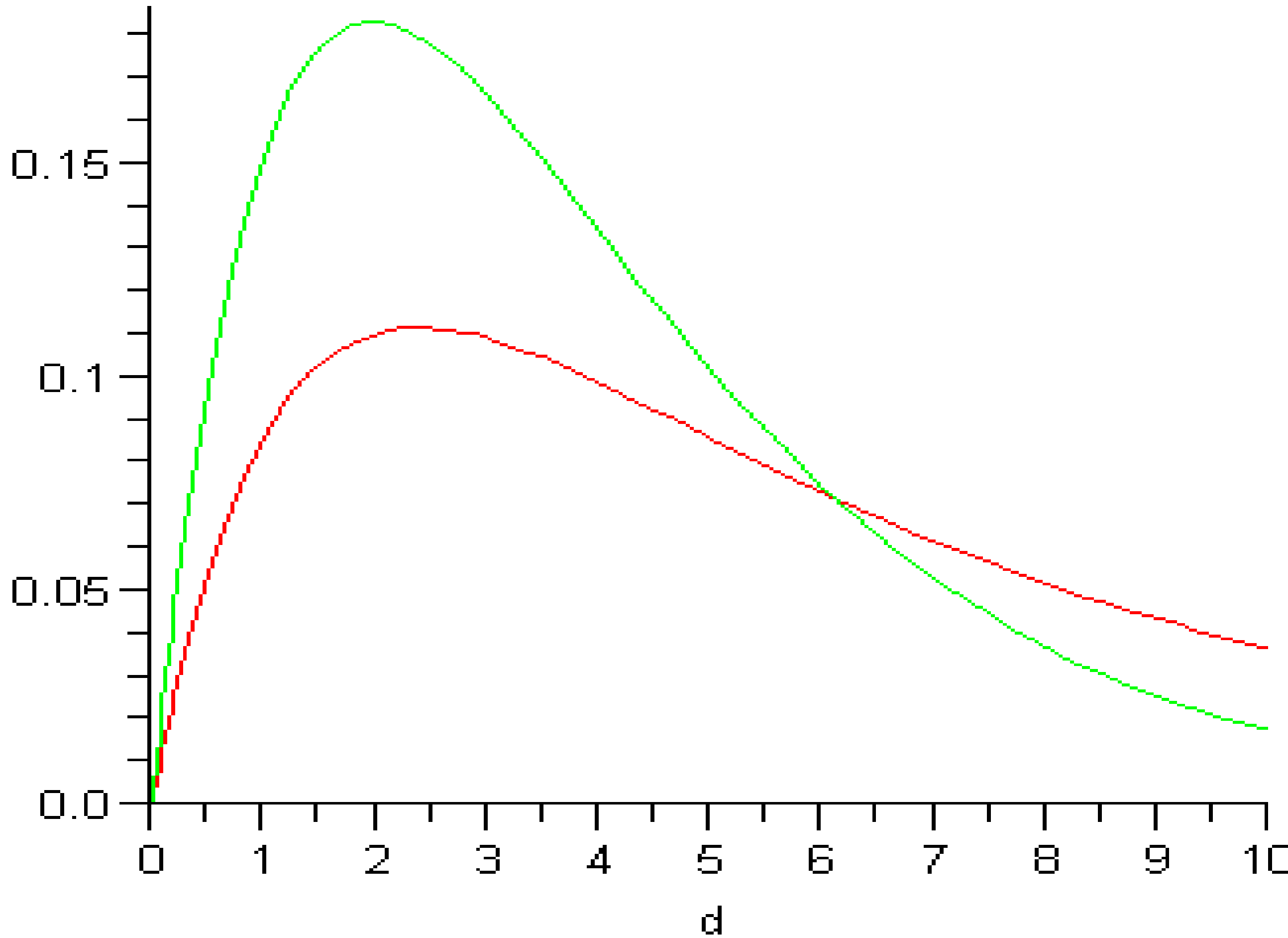


FIGURE 3.1 Comparison of the degree distributions of a coauthorship network and a Poisson random network with the same average degree.

Example - Coauthor versus Romance



Prob given neighbor has degree

Green – romance

Red - coauthor

Co-author versus Romance

- Example: adopt if chance that at least one neighbor adopts exceeds .95 $(1-(1-x)^d \geq c=.95)$
- Romance stable equilibrium:
  - degree 3 and above adopt
  - Prob given neighbor adopts x = .65
  - Percent adopting = .29
- Coauthor stable equilibrium:
  - degree 2 and above adopt
  - Prob given neighbor adopts x = .91
  - Percent adopting = .55
  - Utility higher

Raising Costs

- Raising of costs of adoption of action 1 (FOSD shift of H) lowers φ(x) pointwise
  - raises tipping points, lowers stable equilibria

MPS Shifts

- P(d) is a Mean Preserving Spread of P′(d) if P and P′ correspond to identical means and:

$$\sum_{d=0}^{d^*} P(d) \geq \sum_{d=0}^{d^*} P'(d) \;\; for\ all\ d^*$$

MPS Shifts

- P(d) is a Mean Preserving Spread of P’(d) if P and P’ correspond to identical means and:

$$\sum_{d=0}^{d^*} P(d) \geq \sum_{d=0}^{d^*} P'(d) \;\; for\ all\ d^*$$

- For any convex function f(d):

$$\sum_{d} f(d)P(d) \geq \sum_{d} f(d)P'(d)$$

Increasing Variance of Degrees

- $v(d,x)$ increasing convex in d, H convex
  - e.g., $v(d,x)=dx$, H uniform[0,C] (with high C)

- p′ is MPS of p implies $\varphi(x)$ is pointwise higher under p′

- Roughly, increasing variance leads to lower tipping points and higher stable equilibria

Intuition:

- MPS increases number of high degree nodes. With increasing v, they adopt in greater numbers and thus decrease tipping point

Intuition:

- MPS increases number of high degree nodes. With increasing v, they adopt in greater numbers and thus decrease tipping point
- Convexity in v and H: the increases of adoption rates from higher degrees more than offset the decrease in rates from lower degrees; leads to higher overall equilibrium

Can we relate the payoff structure to equilibrium?

- Assume v(d,x)=v(d)x
- Vary v(d)
- If we can influence v, whom should we target to shift equilibrium?

Proposition: impact of v(d)

Consider changing v(d) by rearranging its ordering.

Proposition: impact of v(d)

Consider changing v(d) by rearranging its ordering.

If p(d)d increasing, then v(d) increasing raises φ(x) pointwise (raises stable equilibria, lowers unstable)

[e.g., p is uniform]

Proposition: impact of v(d)

Consider changing v(d) by rearranging its ordering.

If p(d)d increasing, then v(d) increasing raises φ(x) pointwise (raises stable equilibria, lowers unstable)

[e.g., p is uniform]

If p(d)d decreasing, then v(d) decreasing raises φ(x) pointwise (lowers stable equilibria, raises unstable)

[e.g., p is power]

Optimal Targeting

- Goes against idea of “targeting” high degree nodes
- Want the most probable neighbors to have the best incentives to adopt

What about adoption rates?

- Does adoption speed up or slow down?
- How does this depend on payoff/network structure?
- How does it differ across d?

Adoption varied across d

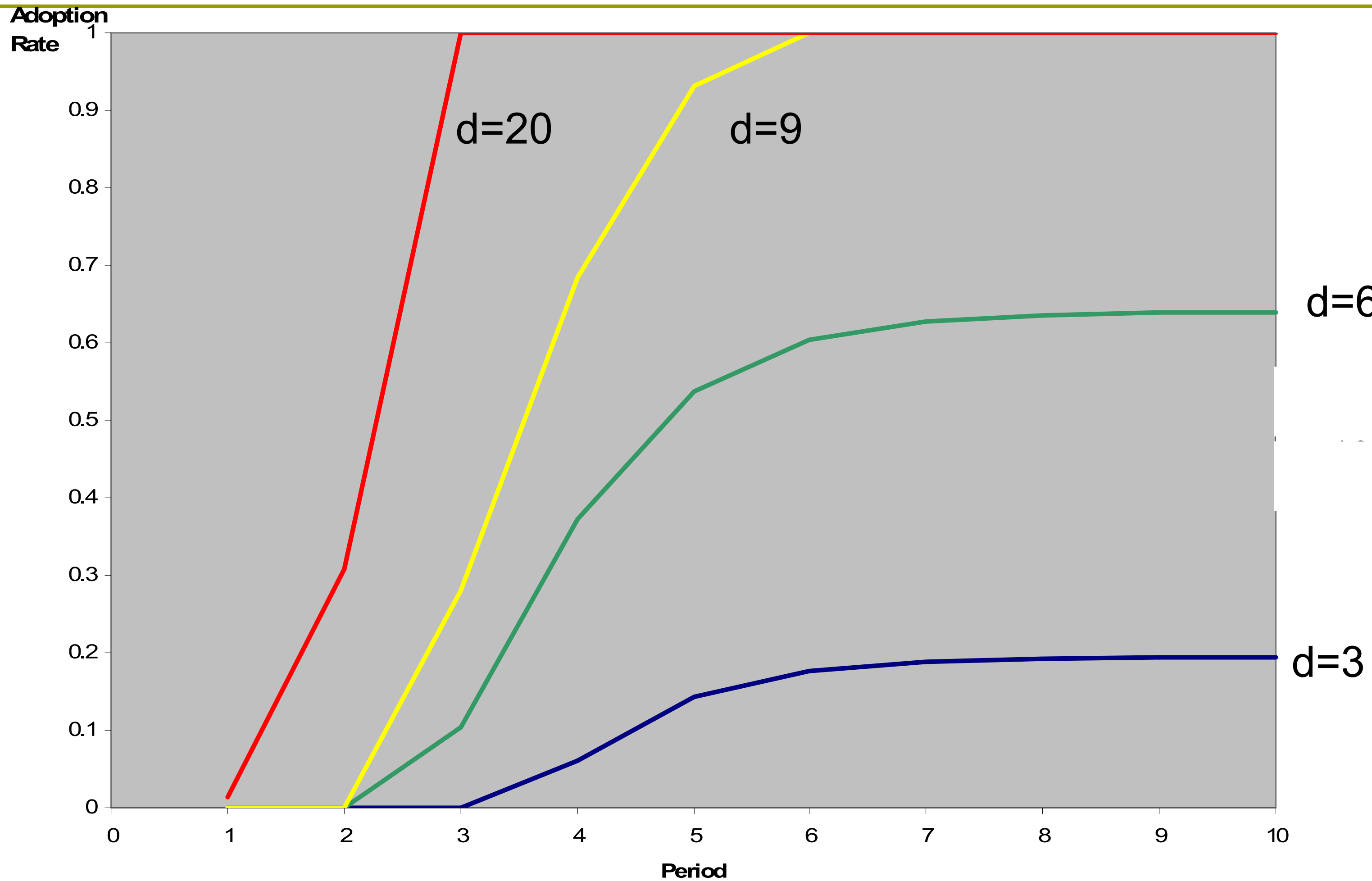
- if v(d,x) is increasing in d, then clearly higher d adopt in higher percentage for each x
  - adoption fraction is H(v(d,x)) which is increasing
- Patterns over time?

Speed of adoption over time

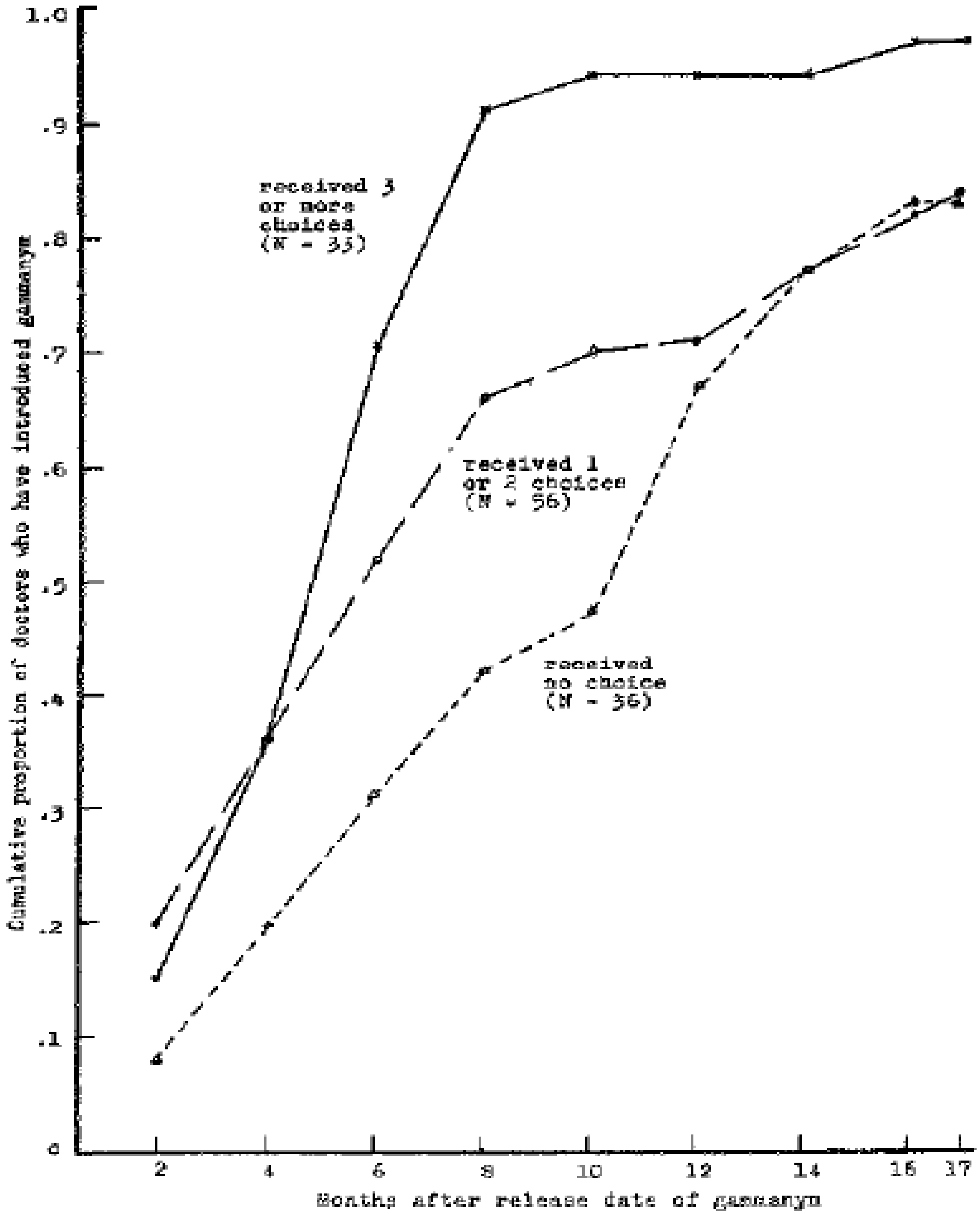
If $H(0)=0$ and H is C^2 and increasing

- If H is concave, then $\varphi(x)/x$ is decreasing
  - Convergence upward slows down, convergence downward speeds up
- If H is convex, then $\varphi(x)/x$ is increasing
  - Convergence upward speeds up, convergence downward slows down

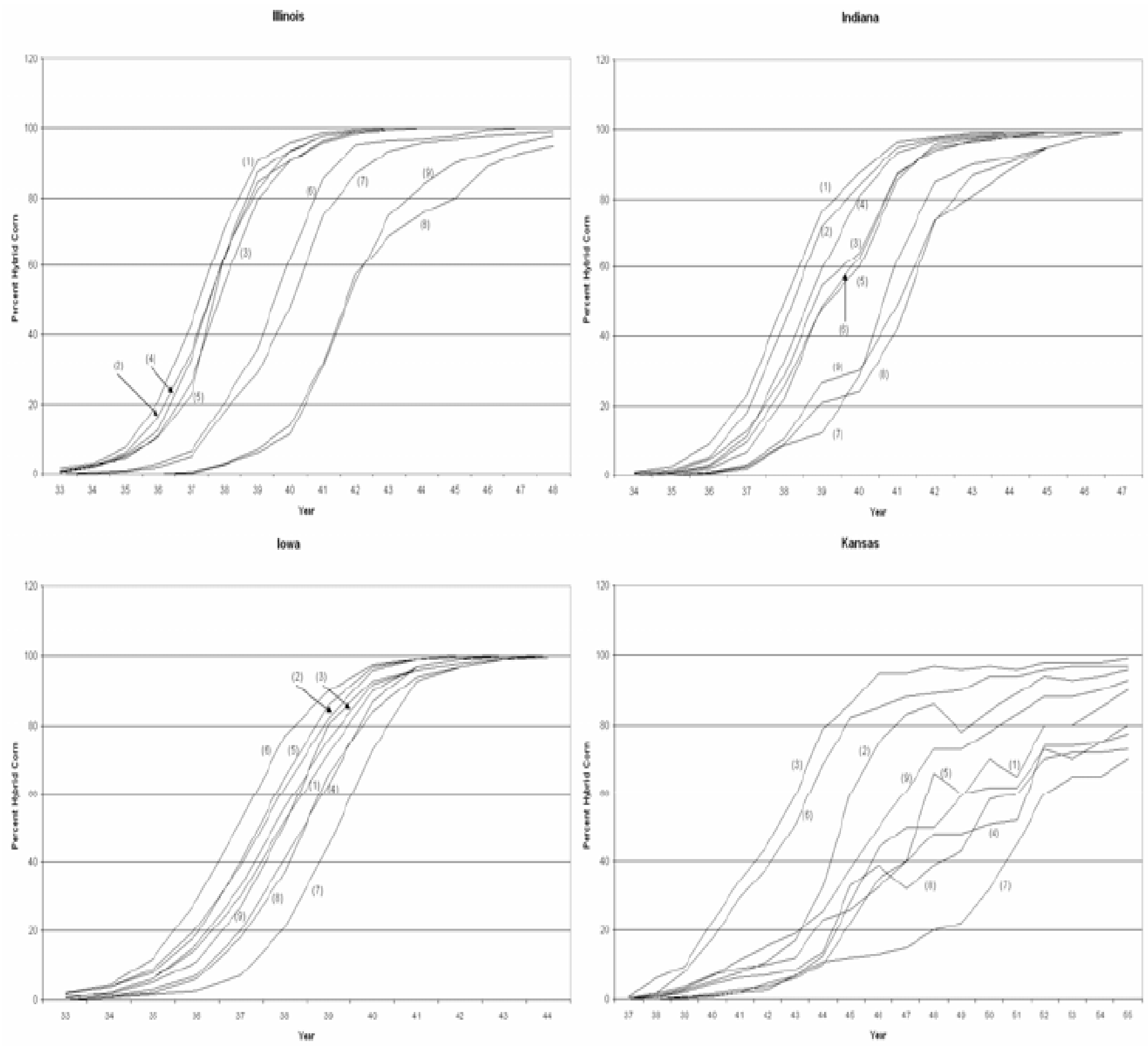
Diffusion Across Degrees



fraction adopting over time, power distribution exponent -2, initial seed x=.03, costs Uniform[1,5], v(d)=d



Tetracycline Adoption
(Coleman, Katz, and Menzel, 1966)



Hybrid Corn, 1933-1952
(Griliches, 1957, and Young, 2006)

Summary:

- Networks differ in structure – Capture some aspects by degree distribution

Summary:

- Networks differ in structure – Capture some aspects by degree distribution

- Location matters:
  - v(d,x) increasing in d
    - more connected adopt "earlier," at higher rate
    - have higher expected payoffs

Summary:

- Networks differ in structure – Capture some aspects by degree distribution

- Location matters:
  - v(d,x) increasing in d
    - more connected adopt "earlier," at higher rate
    - have higher expected payoffs

- Structure matters:
  - Lower tipping points, raise stable equilibria if:
    - lower costs (downward shift FOSD of H)
    - increase in connectedness (FOSD shift of P)
    - MPS of p if v, H (weakly) convex
    - match higher propensity v(d) to more prevalent degrees p(d)d (want *decreasing* v for power laws)
  - adoption speeds vary over time depending on curvature of the cost distribution

Network Formation

- Two simple (mechanical) models generating Poisson and Power-like distributions

- One simple (strategic) model generating similarity between connected nodes (homophily)

Uniform Randomness

- Index nodes by birth time: node i born at i=0,1,2,…

- $d_i(t)$ – degree of node i at time t

Uniform Randomness

- Index nodes by birth time: node i born at i=0,1,2,…
- $d_i(t)$ – degree of node i at time t
- $d_i(i)$ – number of links formed at birth
- $d_i(t) - d_i(i)$ – number of links formed between i and nodes born between i+1 and t

Dynamic Connections

- Suppose we start with m+1 nodes all connected (born in periods 0,…,m)
- From m+1 and on, each newborn node connects to m random nodes.

Dynamic Connections

- Suppose we start with m+1 nodes all connected (born in periods 0,…,m)
- From m+1 and on, each newborn node connects to m random nodes.
- Consider expected degrees

Continuous Time Approximation

- Initial condition: $d_i(i) = m$

- Approximate change over time:

$$\frac{dd_i(t)}{dt} = \frac{m}{t} \quad \text{for all } t > i$$

- This ODE has the solution:

$$d_i(t) = m + m * log(\frac{t}{i})$$

Deducing Degree Distribution

- For any d, t, find i(d) such that:

$$d_{i(d)}(t) = d$$

Deducing Degree Distribution

- For any d, t, find i(d) such that:

$$d_{i(d)}(t) = d$$

- Then,

$$F_t(d) = 1 - \frac{i(d)}{t}$$

Deducing Degree Distribution

- For any d, t, find i(d) such that:

$$d_{i(d)}(t) = d$$

- Then,

$$F_t(d) = 1 - \frac{i(d)}{t}$$

- Solving we get:

$$d = m + m * log\left(\frac{t}{i(d)}\right) \leftrightarrow \frac{i(t)}{d} = e^{-\frac{d-m}{m}}$$

Deducing Degree Distribution

- For any d, t, find i(d) such that:

$$d_{i(d)}(t) = d$$

- Then,

$$F_t(d) = 1 - \frac{i(d)}{t}$$

- Solving we get:

$$d = m + m * log\left(\frac{t}{i(d)}\right) \leftrightarrow \frac{i(t)}{d} = e^{-\frac{d-m}{m}}$$

$$\rightarrow \boxed{F_t(d) = 1 - e^{-\frac{d-m}{m}}}$$

Preferential Attachment

- Price (1976), Barabasi and Albert (1999)

- As before, nodes attach randomly, but with probabilities proportional to degrees

Preferential Attachment

- Price (1976), Barabasi and Albert (1999)

- As before, nodes attach randomly, but with probabilities proportional to degrees

- At t, probability i receives a new link to the newborn is:

$$m\frac{d_i(t)}{\sum_{j=1}^{t} d_j(t)}$$

Preferential Attachment

- There are tm links overall → $\sum_{j=1}^{t} d_j(t)$=2tm

Preferential Attachment

- There are tm links overall → $\sum_{j=1}^{t} d_j(t)$=2tm
- So probability i receives a new link:

$$m \frac{d_i(t)}{\sum_{j=1}^{t} d_j(t)} = \frac{d_i(t)}{2t}$$

Preferential Attachment

- There are tm links overall → $\sum_{j=1}^{t} d_j(t)$=2tm

- So probability i receives a new link:

$$m\frac{d_i(t)}{\sum_{j=1}^{t} d_j(t)} = \frac{d_i(t)}{2t}$$

- The continuous-time approximation is then:

$$\frac{dd_i(t)}{dt} = \frac{d_i(t)}{2t}$$

Preferential Attachment

- Can replicate analysis before to get:

$$F_t(d) = 1 - m^2 d^{-2}$$

- The density is then:

$$f_t(d) = 2m^2 d^{-3}$$

- Power distribution with degree 3!

Homophily in Peer Groups

- Homophily = love for the same (Lazarsfeld and Merton, 1954):
  - Socially connected individuals tend to be similar

Homophily in Peer Groups

- Homophily = love for the same (Lazarsfeld and Merton, 1954):
  - Socially connected individuals tend to be similar
- Evidence across the board and across fields (mostly correlational): Politics, Sociology, Economics

Homophily

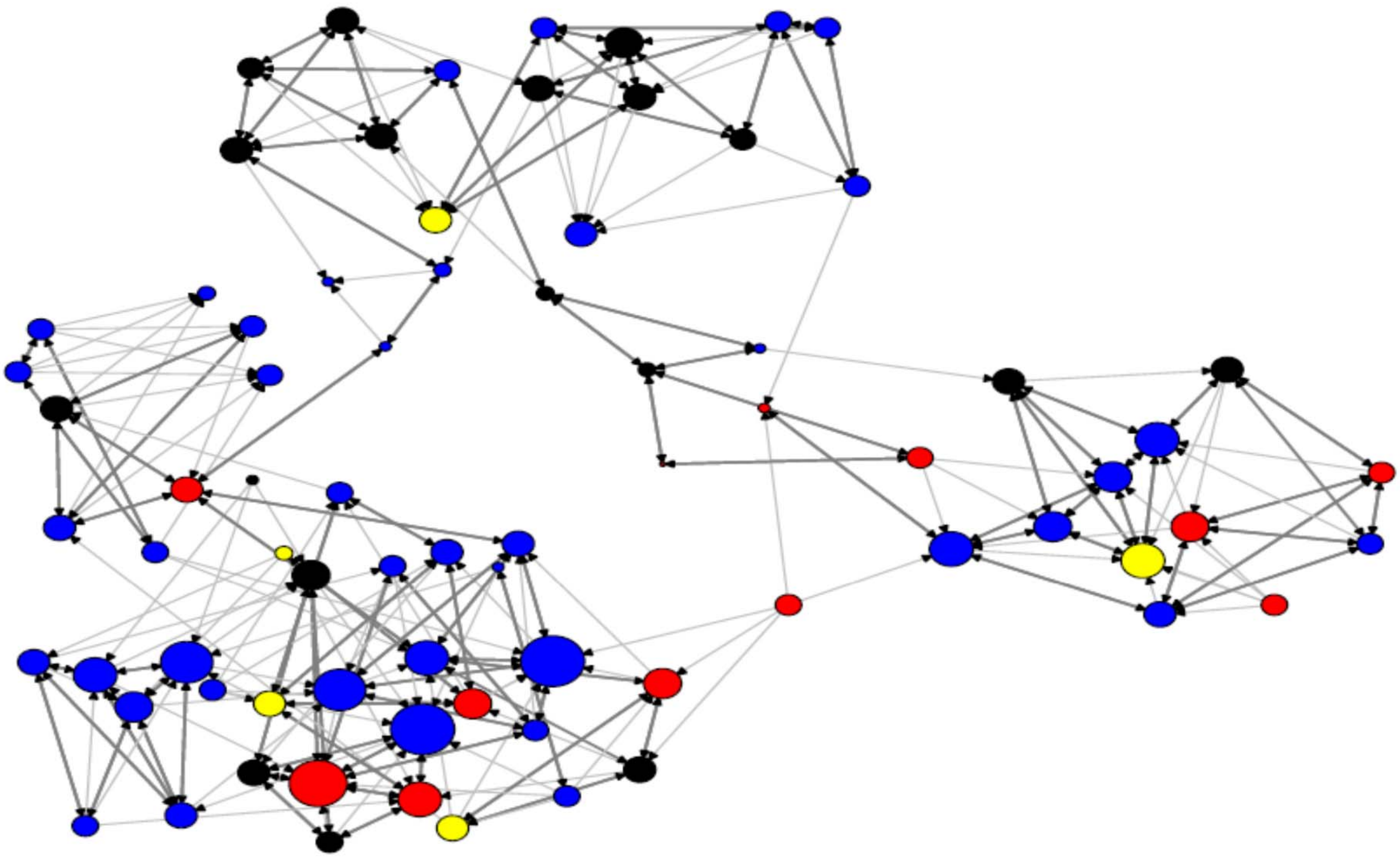
TABLE 3.4
Friendship frequencies (in percent) compared to population percentages by ethnicity in a Dutch high school

	Ethnicity of students				
	Dutch ($n = 850$)	Moroccan ($n = 62$)	Turkish ($n = 75$)	Surinamese ($n = 100$)	Other ($n = 230$)
Percentage of the population (rounded)	65	5	6	8	17
Percentage of friendships with own ethnicity	79	27	59	44	30

Source: Based on data from Baerveldt et al. [27].

Westridge

- Goeree, McConnell, Mitchell, Tromp, Yariv, 2009

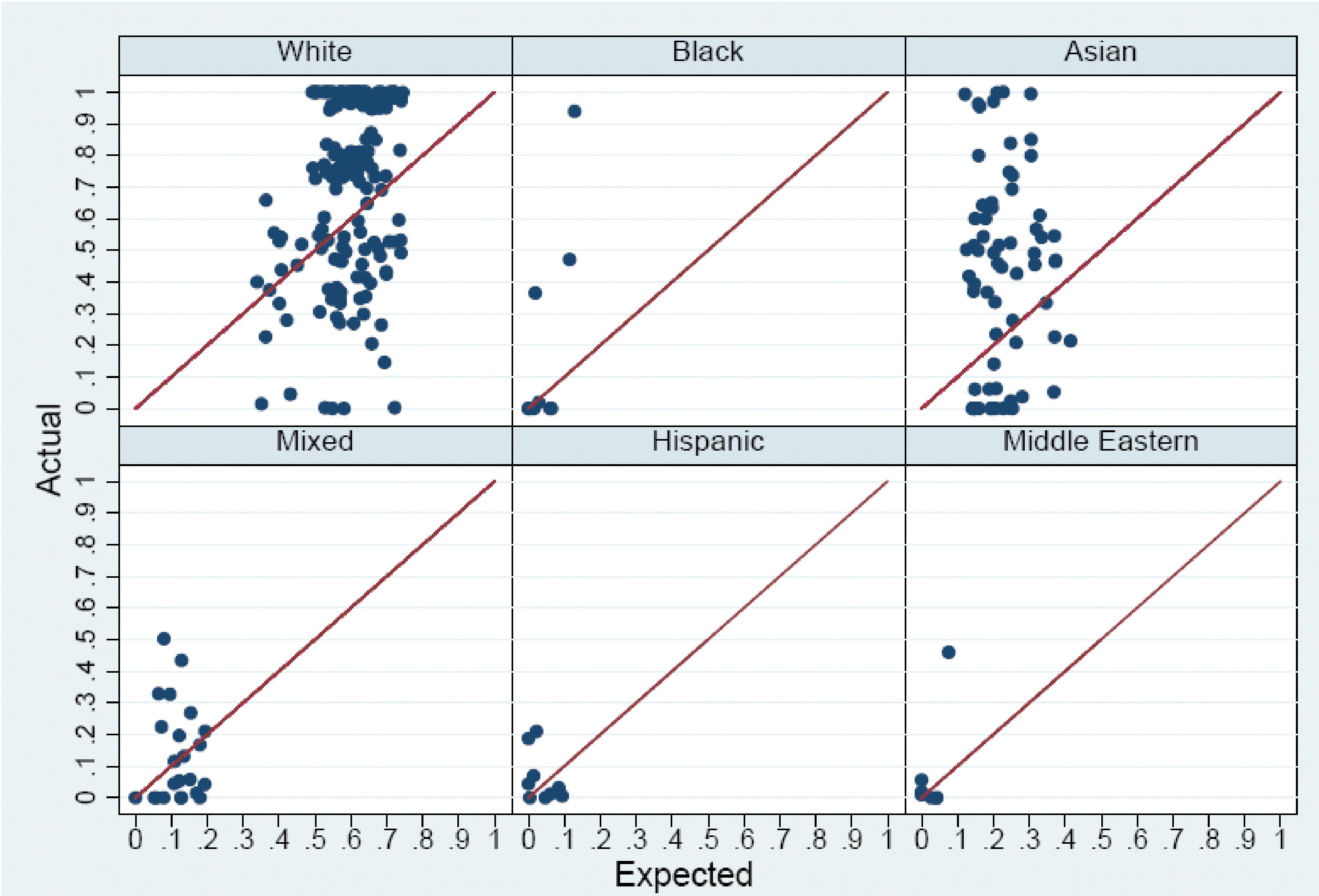


Homophily – Westridge

- 53% of direct friends are of the same race while 41% of all other friends are of the same race

Race	60%
Confidence	53%
Popularity	53%
Height	55%

Homophily – Westridge (2)



The Question

The Question

- In many realms agents choose whom to interact with, socially or strategically
  - Examples: political parties, clubs, internet forums, neighborhoods, etc.
  - New technologies tend to remove geographical constraints in many interactions

The Question

- In many realms agents choose whom to interact with, socially or strategically
  - Examples: political parties, clubs, internet forums, neighborhoods, etc.
  - New technologies tend to remove geographical constraints in many interactions
- Yet, in the literature, group of players is commonly exogenous
  - It is often considered how endowments (demographics, preferences, etc.) of players affect outcomes

The Question

- In many realms agents choose whom to interact with, socially or strategically
  - Examples: political parties, clubs, internet forums, neighborhoods, etc.
  - New technologies tend to remove geographical constraints in many interactions
- Yet, in the literature, group of players is commonly exogenous
  - It is often considered how endowments (demographics, preferences, etc.) of players affect outcomes
- Now: endowments determine friendships that, in turn, affect outcomes
  - Study the structure of (endogenous) groups, predicting both friendships and outcomes

The Goal (Baccara and Yariv, 2012)

The Goal (Baccara and Yariv, 2012)

- Provide a simple, information-based model to analyze how individuals *choose* peer groups prior to a strategic interaction

The Goal (Baccara and Yariv, 2012)

- Provide a simple, information-based model to analyze how individuals *choose* peer groups prior to a strategic interaction
  - individuals differ in how much they care about each of two dimensions (e.g., savings and education, food and music, etc.)
  - individuals in a group play a public good (i.e. information) game

The Goal (Baccara and Yariv, 2012)

- Provide a simple, information-based model to analyze how individuals *choose* peer groups prior to a strategic interaction
  - individuals differ in how much they care about each of two dimensions (e.g., savings and education, food and music, etc.)
  - individuals in a group play a public good (i.e. information) game
- Understand the elements determining the emergence of homophily (or heterophily)
  - information gathering cost
  - group size (communication costs)
  - population attributes

A Model of Information Sharing among Friends

A Model of Information Sharing among Friends

- Two issues, $A, B \in \{0, 1\}$ determined at the outset. For simplicity: $P(A = 1) = P(B = 1) = \frac{1}{2}$

A Model of Information Sharing among Friends

- Two issues, $A, B \in \{0, 1\}$ determined at the outset. For simplicity: $P(A = 1) = P(B = 1) = \frac{1}{2}$
- n agents in a group. Each agent makes a decision on each dimension

$$v = (v_A, v_B) \in V = \{0, 1\} \times \{0, 1\}$$

A Model of Information Sharing among Friends

- Two issues, $A, B \in \{0, 1\}$ determined at the outset. For simplicity: $P(A = 1) = P(B = 1) = \frac{1}{2}$
- n agents in a group. Each agent makes a decision on each dimension

$$v = (v_A, v_B) \in V = \{0, 1\} \times \{0, 1\}$$

- Each agent i characterized by taste $t_i \in [0, 1]$. The utility of agent i from choosing v when the realized states are A and B :

$$u_i(v, A, B) = t_i 1_A(v_A) + (1 - t_i) 1_B(v_B)$$

Information Structure

- Before choosing $v \in V$, agents have access to information.

Information Structure

- Before choosing $v \in V$, agents have access to information.
- Each agent i selects simultaneously an information source $x_i \in \{\alpha, \beta\}$. Source α provides a signal $s \in \{0, 1, \varnothing\}$

$$\Pr(s = A) = q_\alpha > 1/2, \; \Pr(s = \varnothing) = 1 - q_\alpha$$

  Similarly, source β provides the realized state B with probability $q_\beta > 1/2$.
- Signals are conditionally i.i.d.

Information Structure

- Before choosing $v \in V$, agents have access to information.
- Each agent i selects simultaneously an information source $x_i \in \{\alpha, \beta\}$. Source α provides a signal $s \in \{0, 1, \varnothing\}$

$$\Pr(s = A) = q_\alpha > 1/2, \ \Pr(s = \varnothing) = 1 - q_\alpha$$

  Similarly, source β provides the realized state B with probability $q_\beta > 1/2$.
- Signals are conditionally i.i.d.
- **What makes a group?** After information sources are selected, all signals are realized and made public within the group.

$\Rightarrow$ If k agents choose $x = \alpha$,

- probability that state A is revealed is $1 - (1 - q_\alpha)^k$
- probability of making the right decision on A is $1 - \frac{1}{2}(1 - q_\alpha)^k$
- Similarly for $x = \beta$

Our Methodology

Our Methodology

1. *Given a group of agents* $t_1 \geqslant t_2 \geqslant \ldots \geqslant t_n$, we characterize equilibrium information collection

Our Methodology

1. *Given a group of agents* $t_1 \geqslant t_2 \geqslant \ldots \geqslant t_n$, we characterize equilibrium information collection
2. We step backward: *given a group size* n, an agent of type $t \in [0, 1]$ can *choose* the other $n - 1$ agents in her group
   - We characterize the optimal group choice for each agent t

Our Methodology

1. *Given a group of agents* $t_1 \geqslant t_2 \geqslant \dots \geqslant t_n$, we characterize equilibrium information collection
2. We step backward: *given a group size* n, an agent of type $t \in [0, 1]$ can *choose* the other $n - 1$ agents in her group
   - We characterize the optimal group choice for each agent t
3. We characterize the stable groups:
   - *A group is stable if it is optimal for all its members*

Our Methodology

1. *Given a group of agents* $t_1 \geqslant t_2 \geqslant \ldots \geqslant t_n$, we characterize equilibrium information collection
2. We step backward: *given a group size* n, an agent of type $t \in [0, 1]$ can *choose* the other $n - 1$ agents in her group
   - We characterize the optimal group choice for each agent t
3. We characterize the stable groups:
   - *A group is stable if it is optimal for all its members*
4. We characterize optimal group choice and stable groups when:
   - *Information is free*
   - *Information is costly*: every signal costs $c > 0$

Our Methodology

1. *Given a group of agents* $t_1 \geqslant t_2 \geqslant \ldots \geqslant t_n$, we characterize equilibrium information collection
2. We step backward: *given a group size* n, an agent of type $t \in [0, 1]$ can *choose* the other $n - 1$ agents in her group
   - We characterize the optimal group choice for each agent t
3. We characterize the stable groups:
   - *A group is stable if it is optimal for all its members*
4. We characterize optimal group choice and stable groups when:
   - *Information is free*
   - *Information is costly*: every signal costs $c > 0$
5. We consider a finite population and we consider the stable allocations on this population into groups

Free Information: Information Collection Equilibrium

- Consider a group of agents $(t_1, .., t_n)$, $t_1 \geqslant t_2 \geqslant ... \geqslant t_n$
- Equilibrium sources: $(x_1, .., x_n) \in \{\alpha, \beta\}^n$

Free Information: Information Collection Equilibrium

- Consider a group of agents $(t_1, .., t_n)$, $t_1 \geqslant t_2 \geqslant \ldots \geqslant t_n$
- Equilibrium sources: $(x_1, .., x_n) \in \{\alpha, \beta\}^n$

Lemma 1 *If there exist $i < j$ such that $x_i = \beta$ and $x_j = \alpha$, then $(y_1, .., y_n) \in \{\alpha, \beta\}^n$, where $y_l = x_l$ for all $l \neq i, j$, $y_i = \alpha$ and $y_j = \beta$ is an equilibrium as well*

$$\overrightarrow{\underbrace{t_n \leqslant t_{n-1} \leqslant \ldots \leqslant t_{\kappa+1}}_{\text{source } \beta} \leqslant \underbrace{t_\kappa \leqslant t_{\kappa-1} \leqslant \ldots \leqslant t_1}_{\text{source } \alpha}}^{t}$$

$\Longrightarrow$ The equilibrium number of α-signals (κ) and β-signals $(n-\kappa)$ is uniquely determined

Free Information: Optimal Group Choice for Type t

Free Information: Optimal Group Choice for Type t

Let $n_f^{\alpha}(t)$ be the optimal number of α-signals for type t when group size is n

$n_f^{\alpha}(t)$ equates marginal contribution of an α-signal and a β-signal

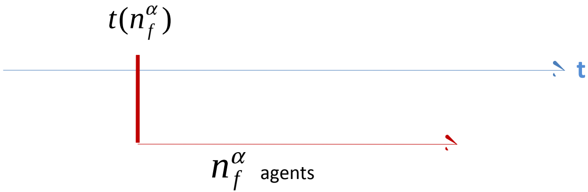
Free Information: Optimal Group Choice for Type t

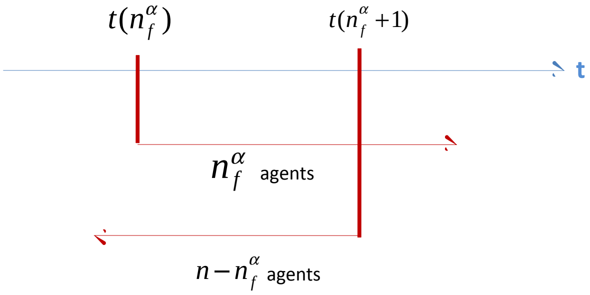
Let $n_f^\alpha(t)$ be the optimal number of α-signals for type t when group size is n

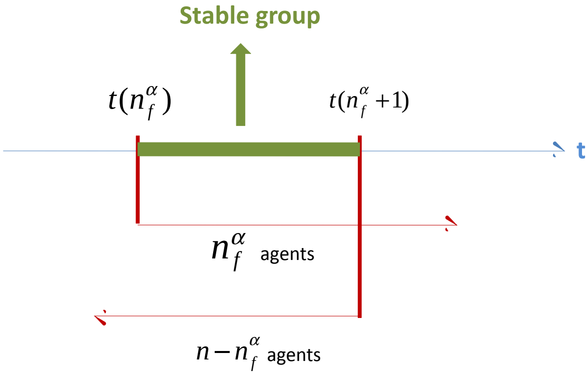
$n_f^\alpha(t)$ equates marginal contribution of an α-signal and a β-signal

For any t, the class of optimal groups $t_1 \geqslant ... \geqslant t_n$ (one of which is t) entails:

- $n_f^\alpha(t)$ agents getting α signals (above the threshold $t(n_f^\alpha(t))$) and
- $n - n_f^\alpha(t)$ agents getting β signals (below the threshold $t^n(n_f^\alpha(t) + 1)$)


$t(n_f^\alpha)$
t
n_f^α agents


$t(n_f^\alpha)$
$t(n_f^\alpha + 1)$
t
n_f^α agents
$n - n_f^\alpha$ agents



Stable group
$t(n_f^\alpha)$
$t(n_f^\alpha+1)$
t
n_f^α agents
$n-n_f^\alpha$ agents

Free Information Case: Stability

Proposition 1

(i) *There exist* $0 = t^n(0) < t^n(1) < ... < t^n(n) < t^n(n+1) = 1$ *such that a group* $(t_1, ..., t_n)$ *is stable if and only if there exists* $k = 0, .., n$ *such that for all* i,

$$t_i \in [t^n(k), t^n(k+1)] \equiv T_k^n$$

Free Information Case: Stability

Proposition 1

(i) *There exist* $0 = t^n(0) < t^n(1) < ... < t^n(n) < t^n(n+1) = 1$ *such that a group* $(t_1, ..., t_n)$ *is stable if and only if there exists* $k = 0, .., n$ *such that for all* i,

$$t_i \in [t^n(k), t^n(k+1)] \equiv T_k^n$$

(ii) *The intervals* T_k^n *are wider for moderate types and narrower for extreme types*

Free Information Case: Stability

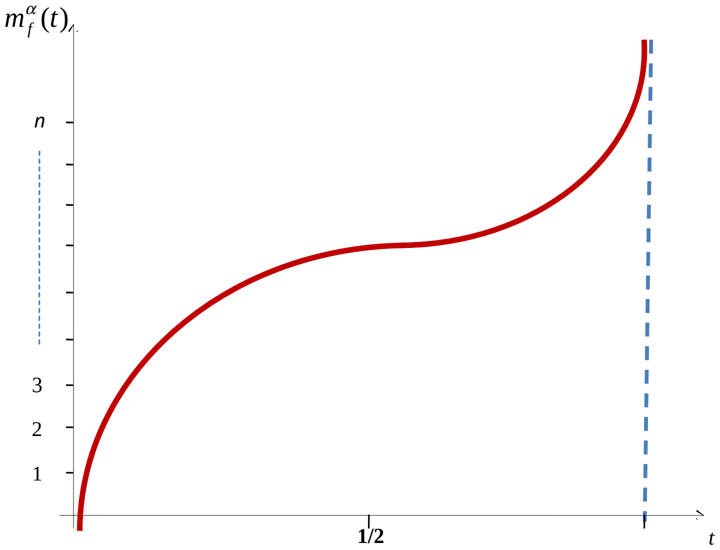
Proposition 1

(i) *There exist* $0 = t^n(0) < t^n(1) < ... < t^n(n) < t^n(n+1) = 1$ *such that a group* $(t_1, ..., t_n)$ *is stable if and only if there exists* $k = 0, .., n$ *such that for all* i,

$$t_i \in [t^n(k), t^n(k+1)] \equiv T_k^n$$

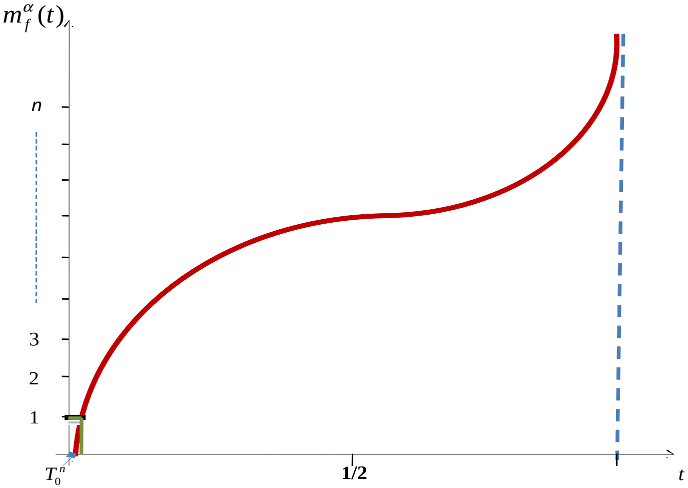
(ii) *The intervals* T_k^n *are wider for moderate types and narrower for extreme types*

$\Rrightarrow$ Note: Same characterization if each agent acquires $h \geq 1$ signals: in stable groups agents agree on allocation of $n \times h$ signals across α and β

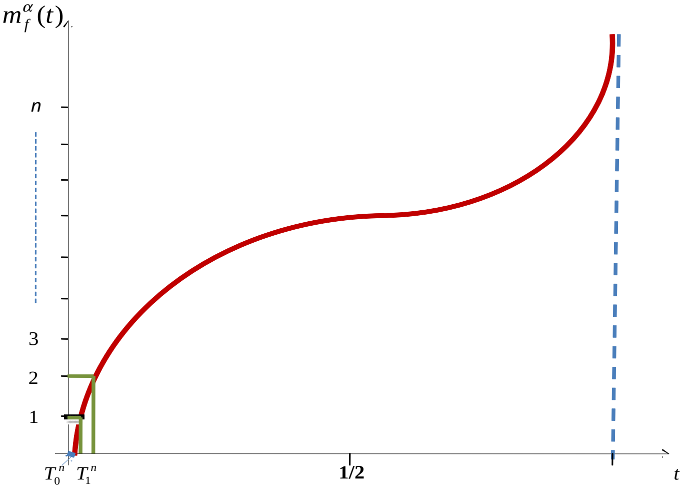


$m_f^\alpha(t)$
n
3
2
1
1/2
t

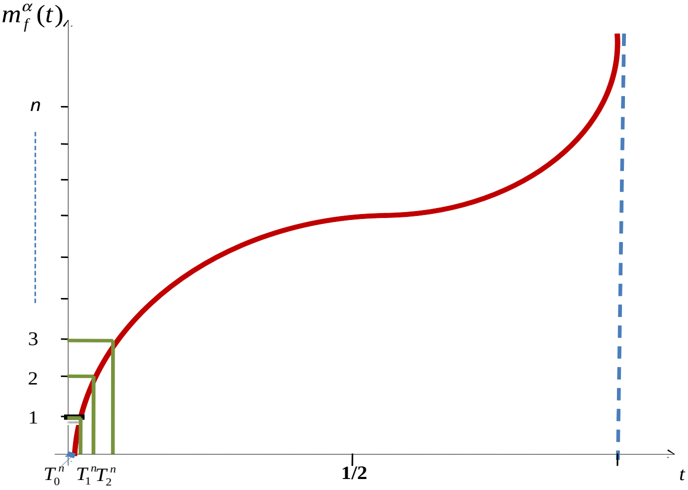
Intuition - Constructing Stable Sets



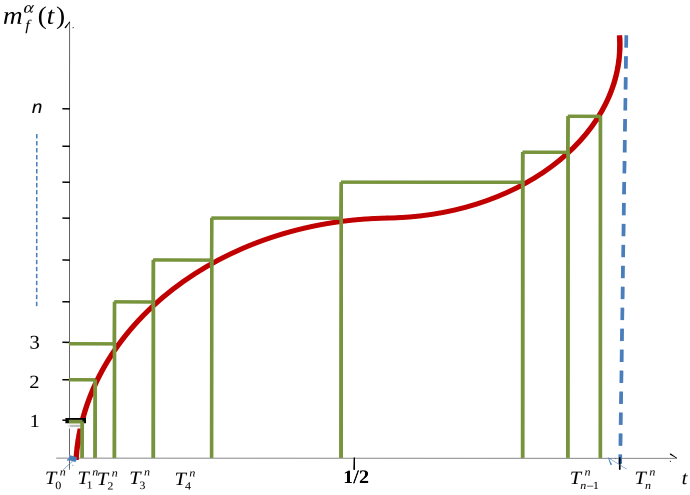
Intuition - Constructing Stable Sets



Intuition - Constructing Stable Sets



Intuition - Constructing Stable Sets



Convergence for Large n

Convergence for Large n

Proposition 2

Consider two agents of taste parameters t, t'.

1. *If they belong to a non-extreme stable group of size $n \geq 2$, they belong to a non-extreme stable group of some size $n' > n \Rightarrow$ Non-extreme intervals do not converge to points*

Convergence for Large n

Proposition 2

Consider two agents of taste parameters t, t'.

1. *If they belong to a non-extreme stable group of size $n \geq 2$, they belong to a non-extreme stable group of some size $n' > n$* $\Rightarrow$ *Non-extreme intervals do not converge to points*
2. *Extreme stable groups become fully homogeneous (containing only $t = 0, 1$) as n diverges* $\Rightarrow$ *The extreme intervals converge to the extremes* 0 and 1

Convergence for Large n

Proposition 2

Consider two agents of taste parameters t, t'.

1. *If they belong to a non-extreme stable group of size $n \geq 2$, they belong to a non-extreme stable group of some size $n' > n \Rightarrow$ Non-extreme intervals do not converge to points*
2. *Extreme stable groups become fully homogeneous (containing only $t = 0, 1$) as n diverges $\Rightarrow$ The extreme intervals converge to the extremes* 0 and 1

$\Rightarrow$ Implication: As group size increases, more homophily for extreme types, stable for moderate types

Stable Groups with Free Information

Stable Groups with Free Information

- Cohesive, larger intervals for moderates, narrower for extremists.
  - **Implication**: As geographical constraints decrease (e.g. automobile, Internet, etc.) $\Longrightarrow$ Homogeneity increases (consistent with Lynd and Lynd (1929)). Stronger for extreme than for moderates

Stable Groups with Free Information

- Cohesive, larger intervals for moderates, narrower for extremists.
  - **Implication**: As geographical constraints decrease (e.g. automobile, Internet, etc.) $\Longrightarrow$ Homogeneity increases (consistent with Lynd and Lynd (1929)). Stronger for extreme than for moderates
- For moderates, heterogeneity persists for large group size
  - **Implication**: As connection costs decrease (e.g., email, chats, sms, online social networking, etc.) $\Longrightarrow$ Homogeneity increases more for extreme individuals than for moderate ones

Stable Groups with Free Information

- Cohesive, larger intervals for moderates, narrower for extremists.
  - **Implication**: As geographical constraints decrease (e.g. automobile, Internet, etc.) $\Longrightarrow$ Homogeneity increases (consistent with Lynd and Lynd (1929)). Stronger for extreme than for moderates
- For moderates, heterogeneity persists for large group size
  - **Implication**: As connection costs decrease (e.g., email, chats, sms, online social networking, etc.) $\Longrightarrow$ Homogeneity increases more for extreme individuals than for moderate ones
- **Empirically**, deducing preferences directly from individual actions is problematic $\Longrightarrow$ Important to account for public goods obtained from friendships

The End