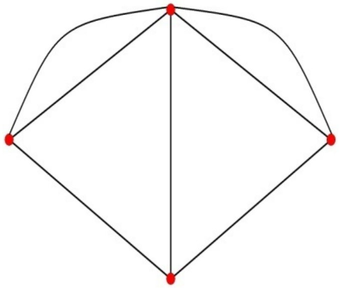
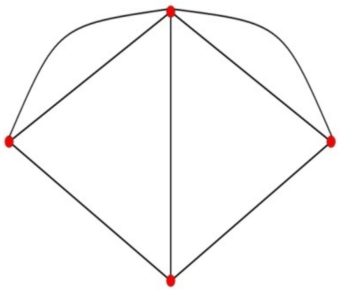
Lecture 3
Diffusion, Identification, Network Formation



Matthew O. Jackson
NBER July 22, 2014

www.stanford.edu\~jacksonm\Jackson-NBER-slides2014.pdf

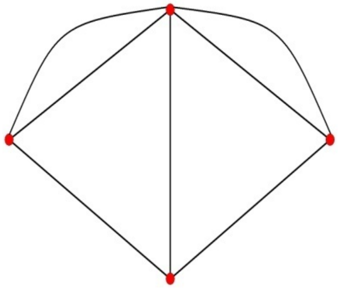
Lecture 3



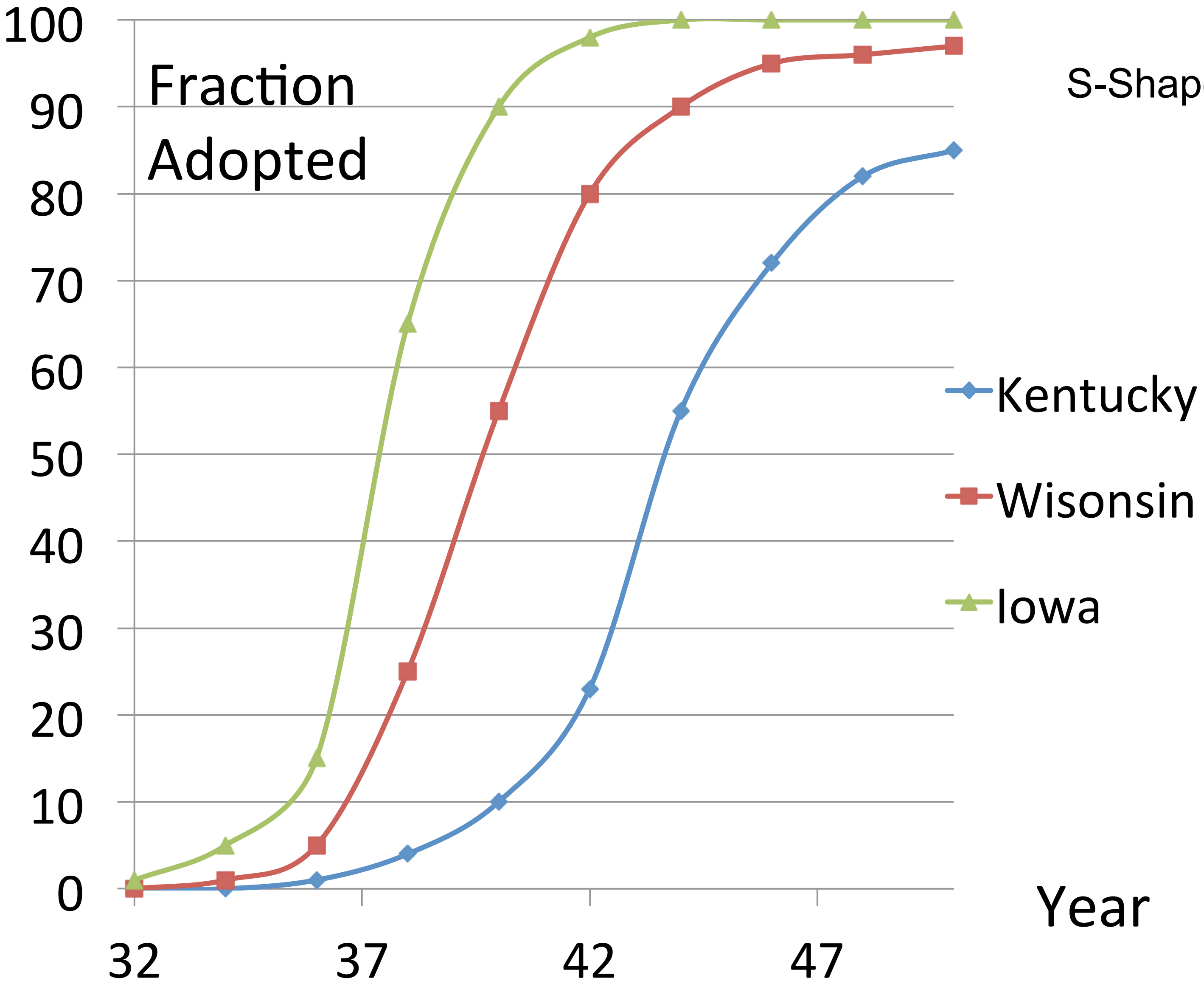
- Diffusion

- More on issues of identification, endogeneity of networks, network formation

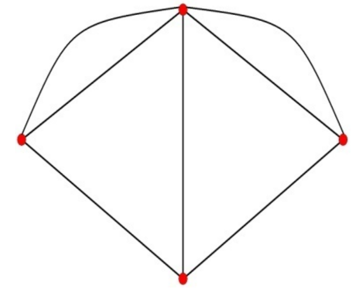
Griliches (1957): Hybrid Corn Diffusion



S-Shape, Spatial Pattern...

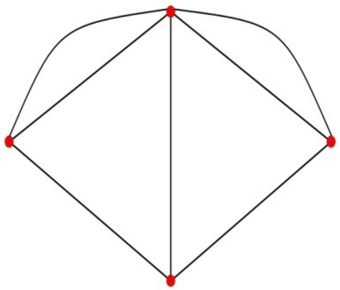


Diffusion of a product/ technology



- Complementarities in choices/compatibility
- Awareness – hear about through friends/ acquaintances
- Learning – about value
- Fads/fashion
- Characteristics - just similar tastes to friends due to homophily…

Dissecting diffusion

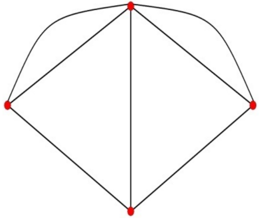


- Policy implications

  - Externalities can cause diffusion to be too slow, inefficient

  - What is driving diffusion? Should we/can we improve it?

Identification



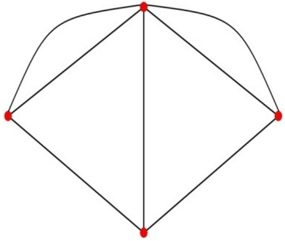
- Field/natural experiments (e.g., pseudo random injections in Indian data, identification – *but don't control networks*…)

- IV   (Just saw in Lecture 2)
  - exploiting network position (Bramoulle, Djebbari, and Fortin, - *does not address endogenous networks/unobservables*…)
  - things that affect network, but not behavior (Acemoglu, Garcia-Jimeno, and Robinson - rare …)

- **Structural modeling of behavior (e.g., diffusion model…)**

- **Model network formation**...

Application: Structural Modeling

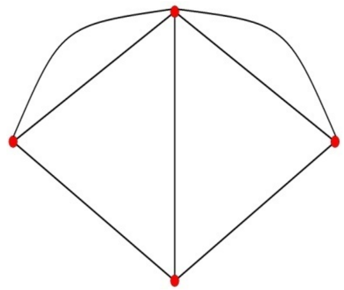


- Use networks in richer way than just mapping peers

- Model diffusion and use it to identify behavior:

  Track paths of information diffusion

Micro - Individual Behavior and Peer Effects:

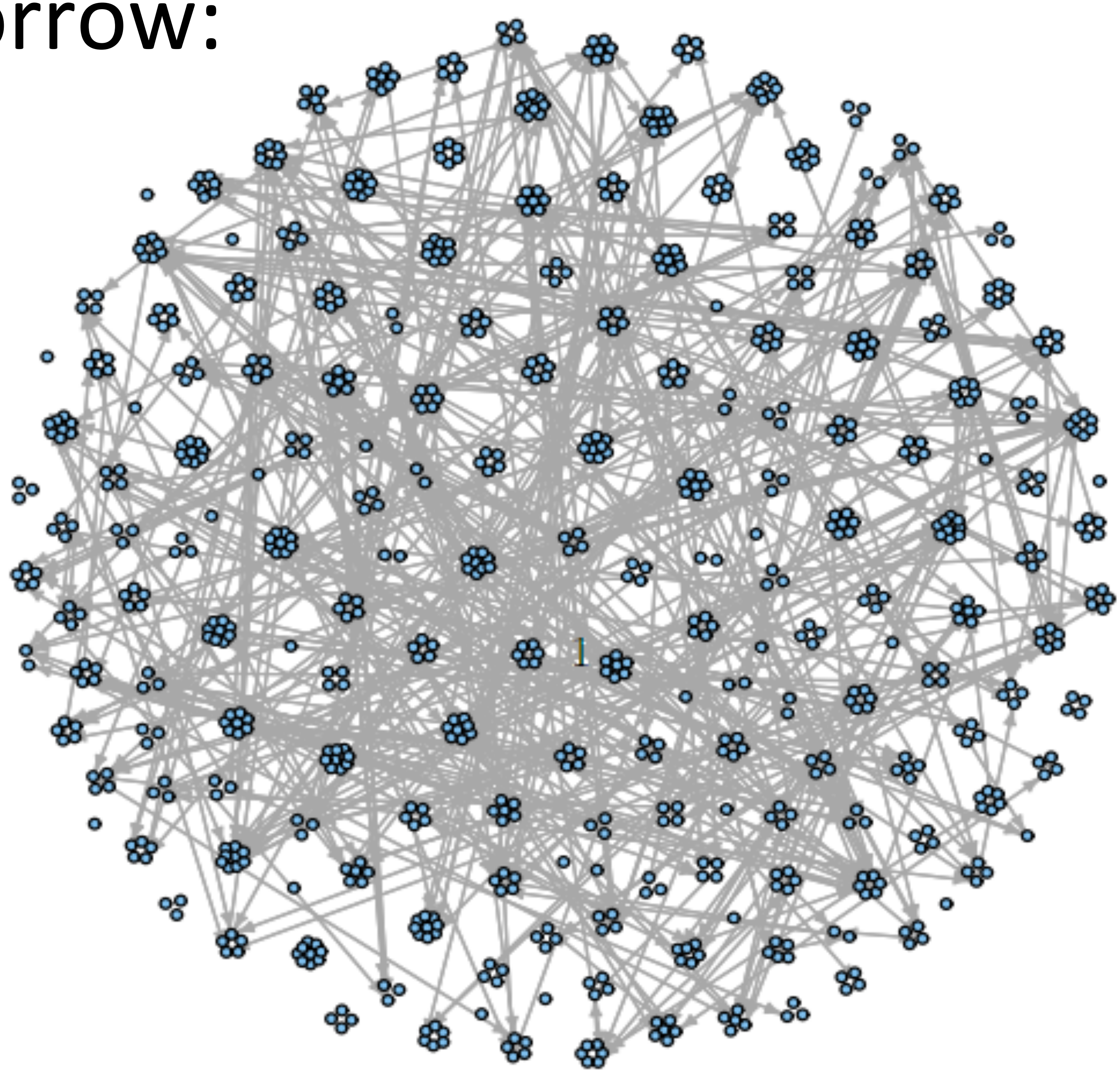


- ***Disentangling Peer effects:***

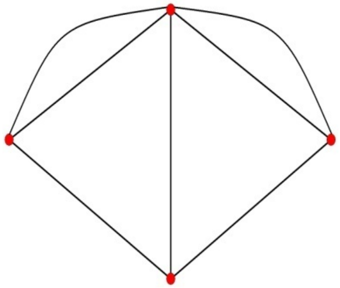
  - **Basic information diffusion:** about a product – being aware of new product

  - **Peer influence/Endorsement/Game on Network:** even if aware, more neighbors taking action leads to higher (or lower) action -- endorsement (learning), peer pressure, complementarities...

Borrow:



Start with Standard Peer-effects analysis:



Let p_i be i's choice of whether to participate

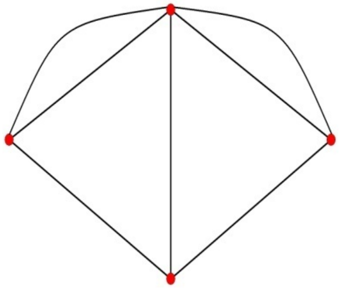
- $\text{Log}(p_i/(1-p_i))$

  $= \; b_0$

  $+ \; b_{char}$ characteristics$_i$

  $+ \; b_{Peer}$ frac$_i$ friends participating

Start with Standard Peer-effects analysis:



Let p_i be i's choice of whether to participate

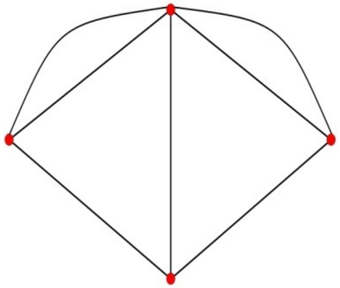
- $\text{Log}(p_i/(1-p_i))$

  $= \quad b_0$

  $+ \, b_{char}$ characteristics$_i$

  $+ \, \mathbf{2.5}^{***}$ frac$_i$ friends participating

Start with Standard Peer-effects analysis:



Let p_i be i's choice of whether to participate

- $\text{Log}(p_i/(1-p_i))$

  $= \; b_0$

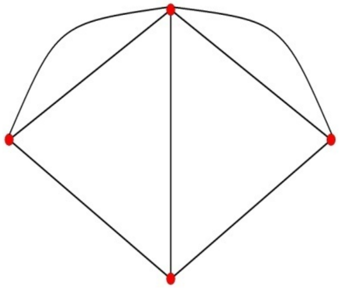
  $+ \; b_{char}$ characteristics$_i$

  $+$ **2.5***** frac$_i$ friends participating

frac 0 to 1 increases $p_i/(1-p_i)$ by factor 12.2,

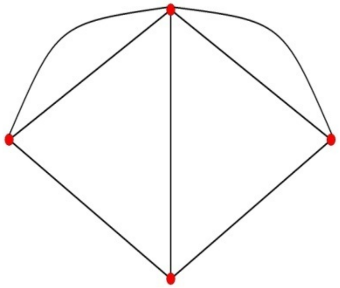
frac .1 to .3 increases $p_i/(1-p_i)$ by factor 1.65,

Modeling diffusion:



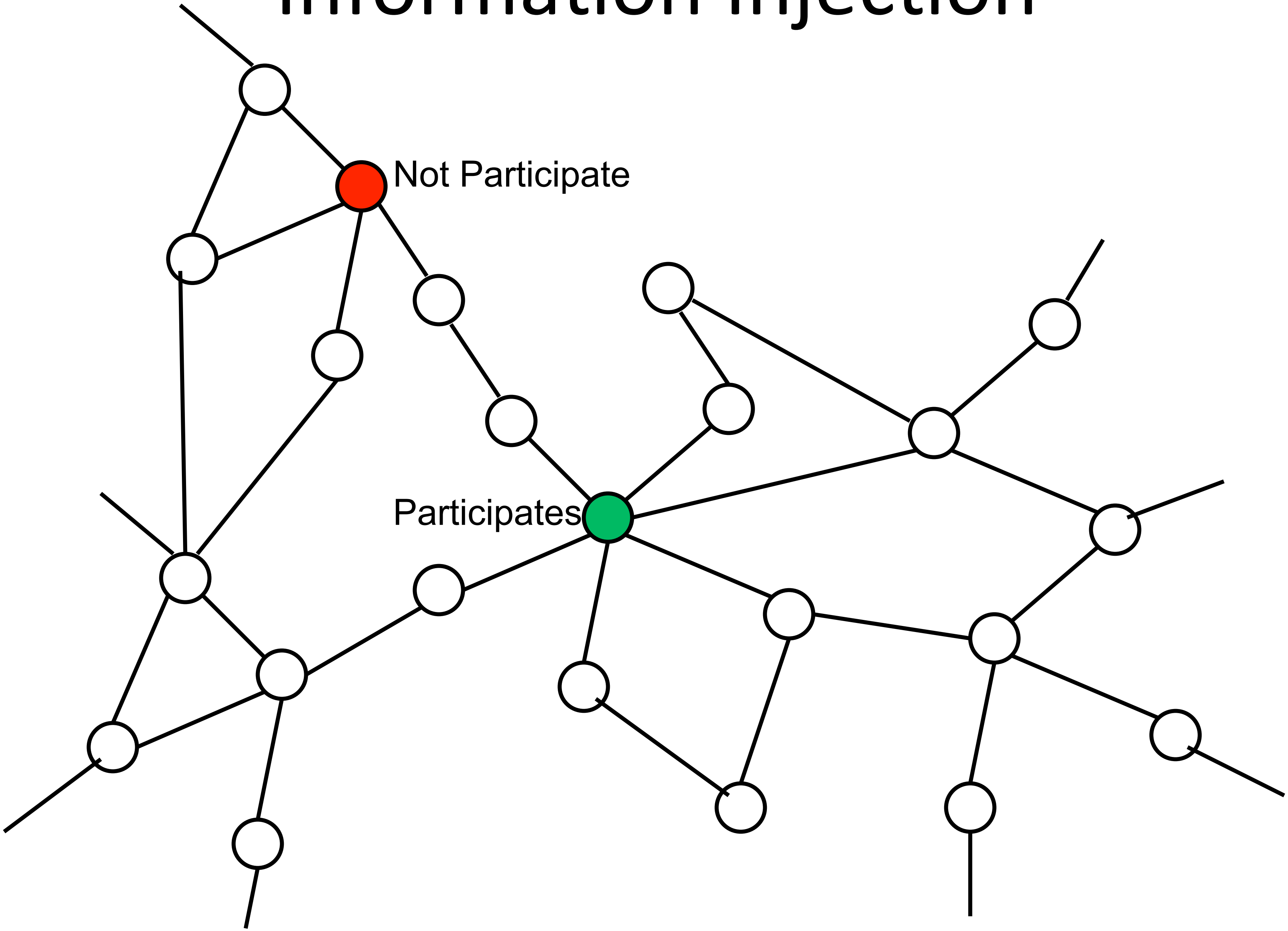
- We know the set of initially informed nodes
- Informed nodes (repeatedly) pass information randomly to their neighbors over discrete times
- Once informed (just once), nodes choose to participate depending on their characteristics and their neighbors' choices

Modeling behavior/information diffusion:

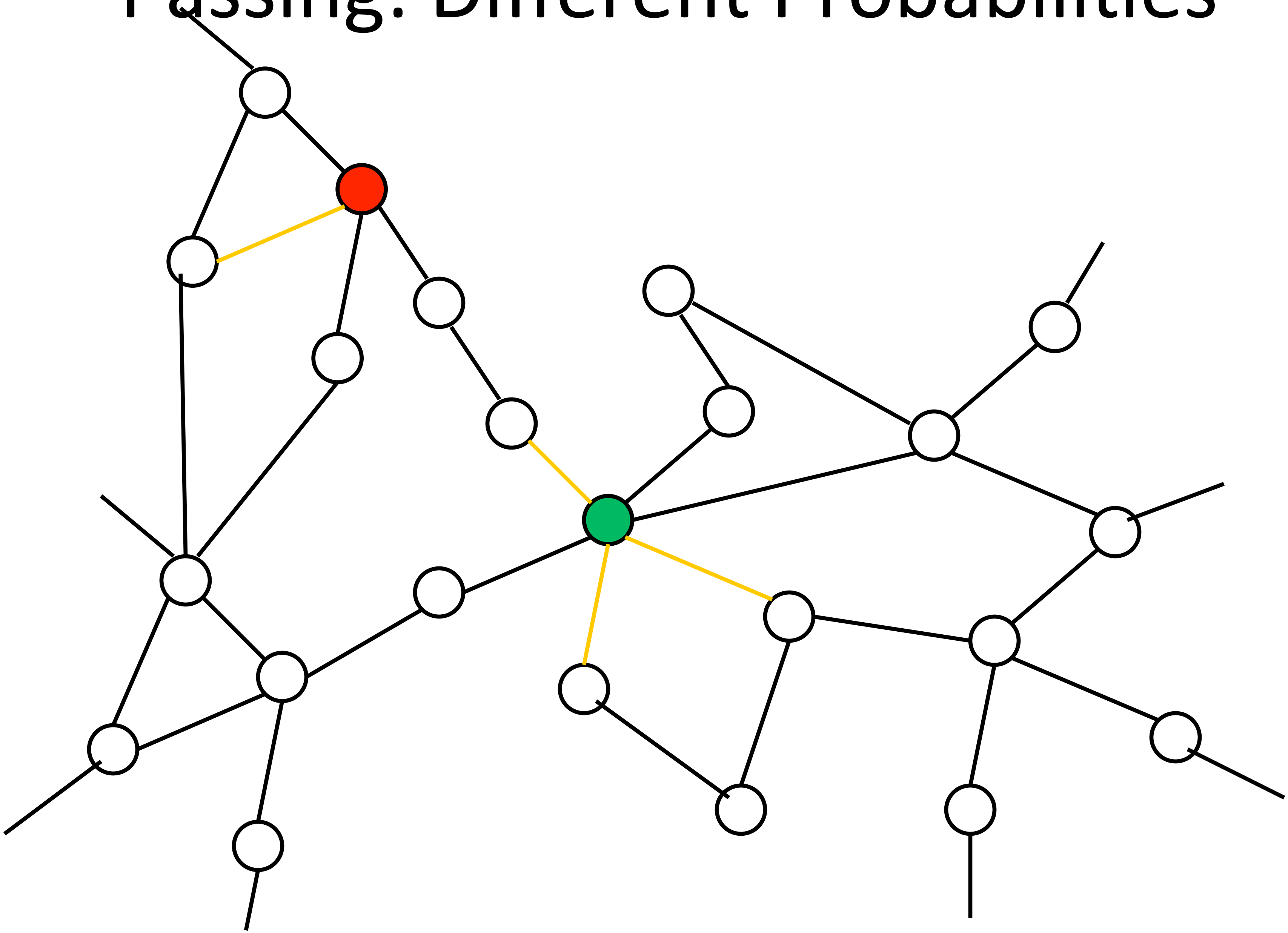


- Probability of passing to a given individual:
  - q^N if did Not participate
  - q^P if did Participate

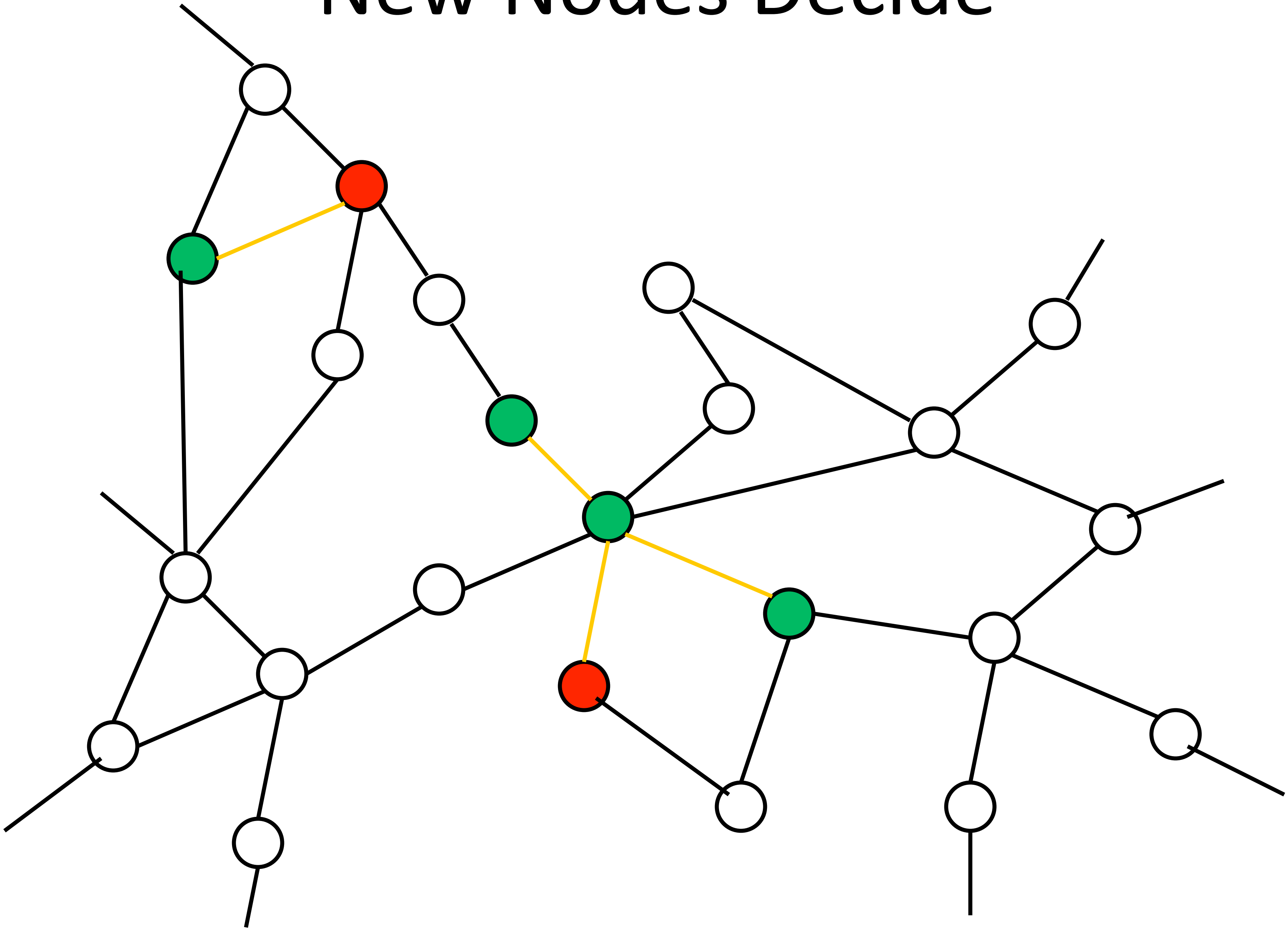
Information Injection



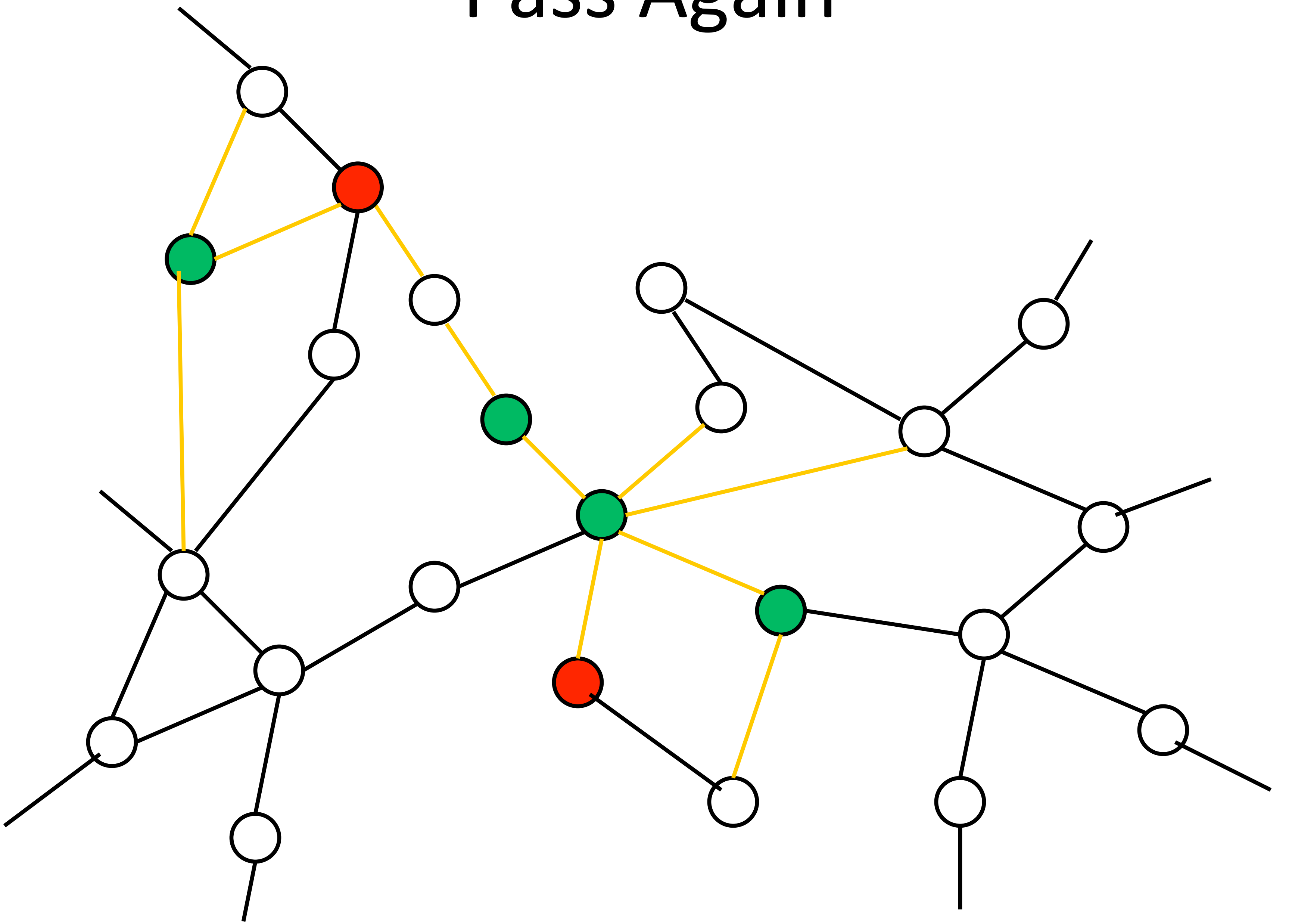
Passing: Different Probabilities



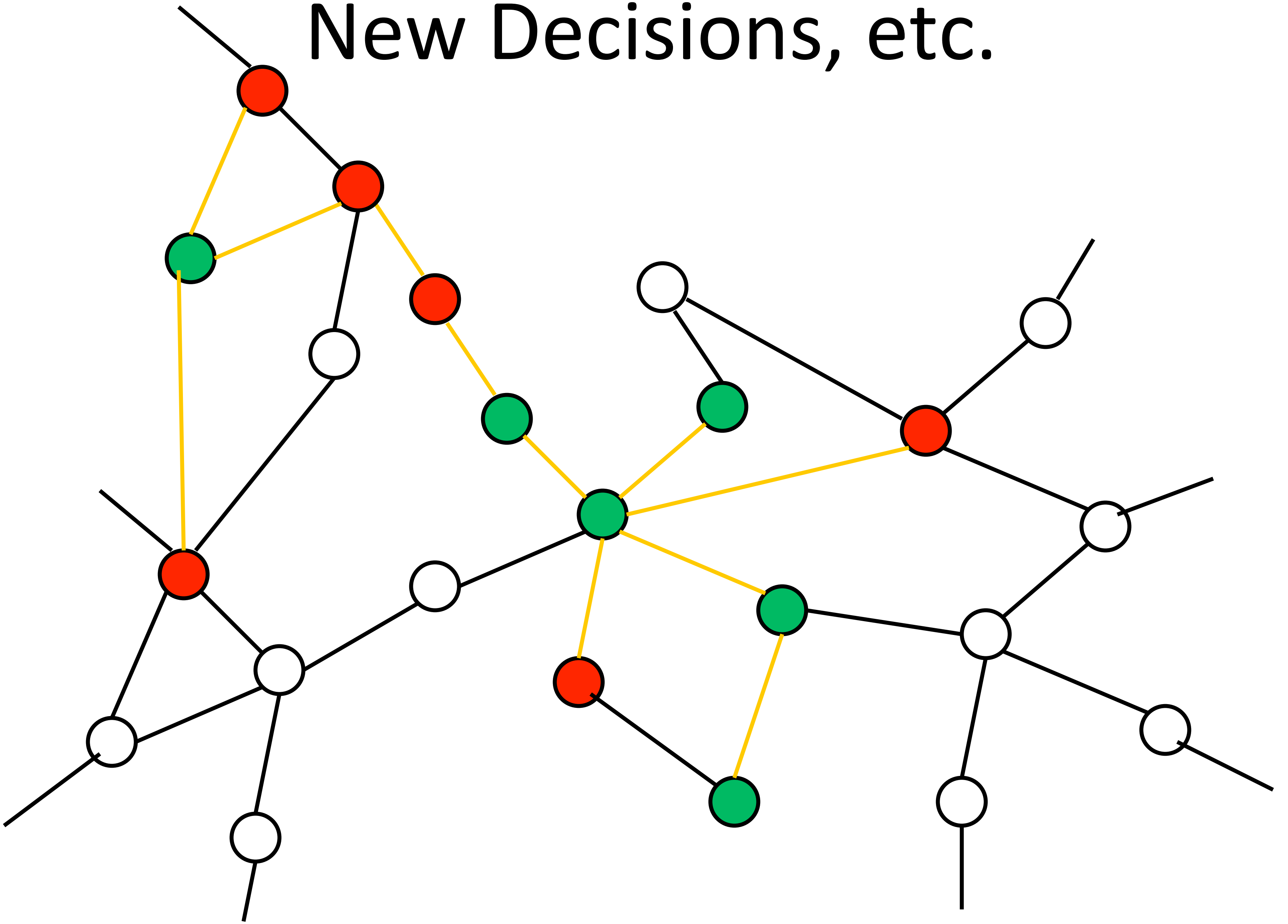
New Nodes Decide



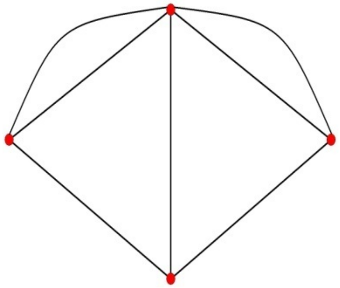
Pass Again



New Decisions, etc.



Choice Decision



- Now *conditional upon being informed*:

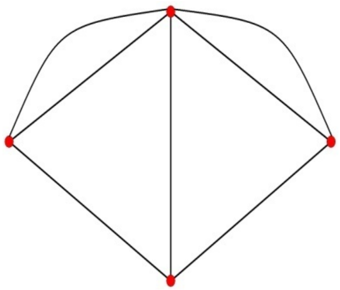
- $\mathrm{Log}(p_i/(1-p_i))$

  $= \quad b_0$

  $+ \; b_{char}$ characteristics$_i$

  $+ \; b_{Peer}$ frac$_i$ informing friends participating

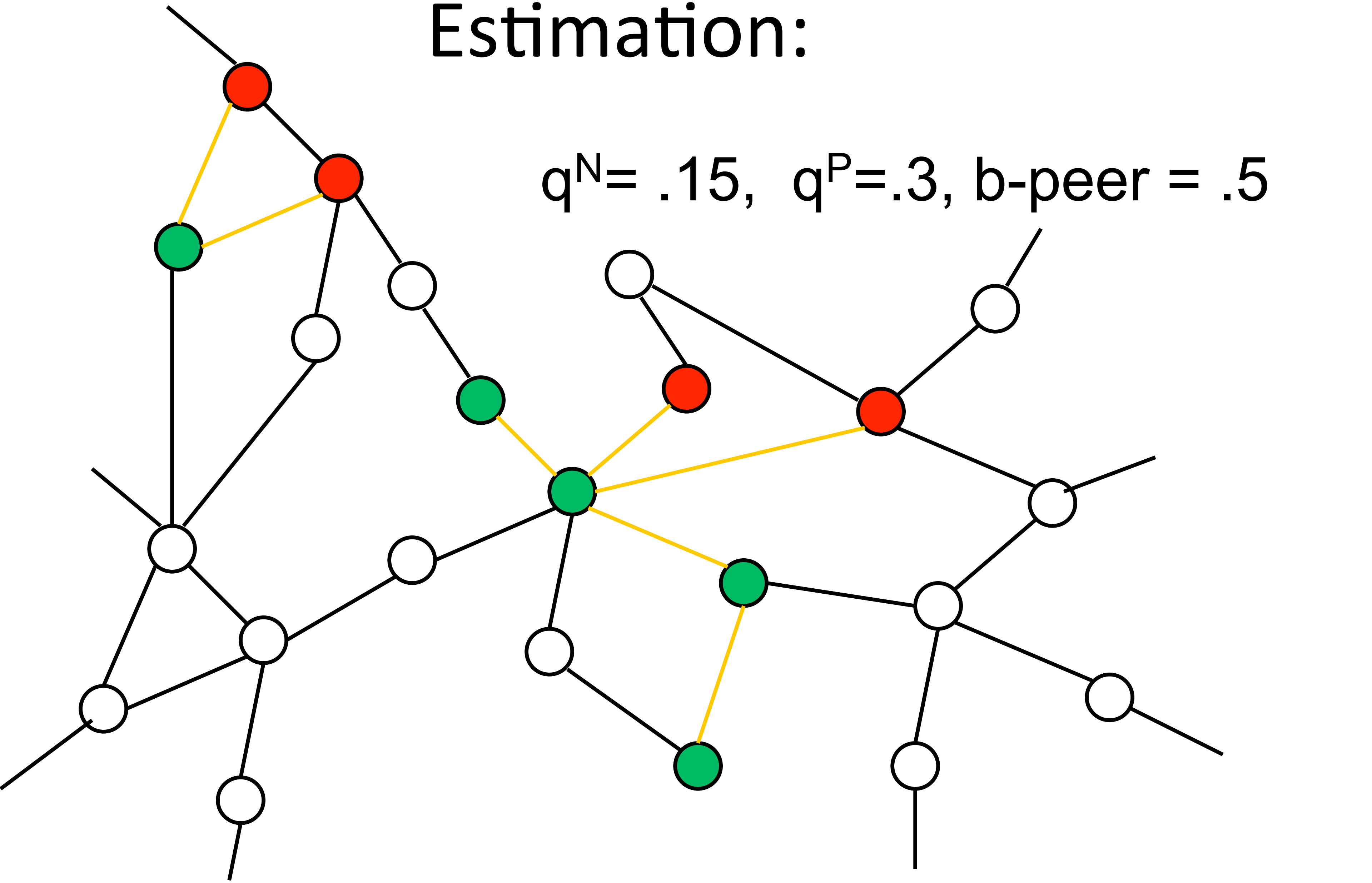
Estimation technique:



- Estimate b_0, b_{char}, from initially informed (saves on computation size of grid)
- q^N, q^P, b_{peer} - For each choice of parameters, simulate on the actual networks of the villages for time period proportional to number of trimesters in data for village (3 to 8 times)
- Choose parameters to best match simulated participation rates and various moments to observed moments (SMM)

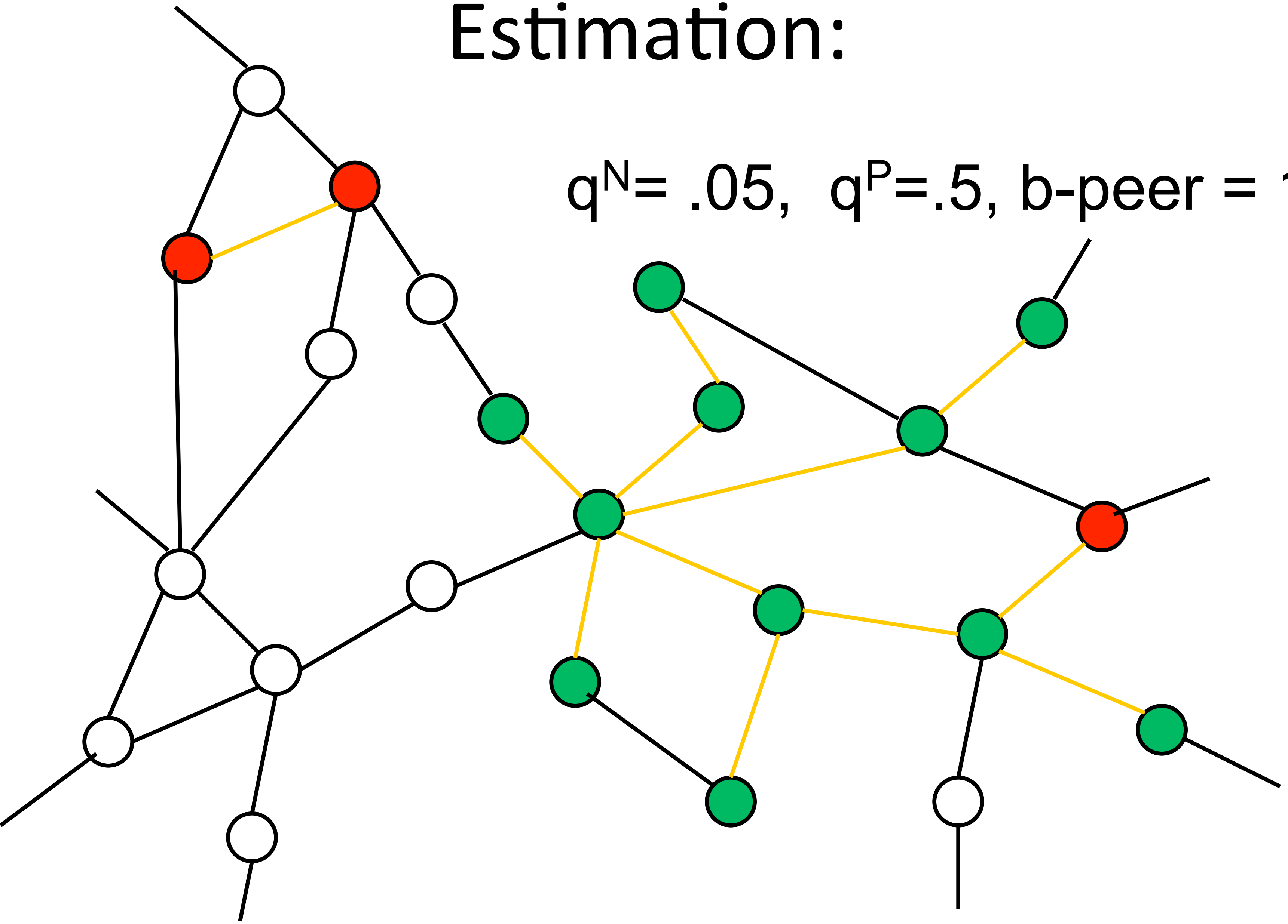
Estimation:

$q^N = .15$, $q^P = .3$, b-peer = .5

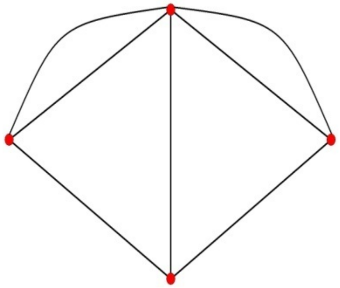


Estimation:

$q^N = .05$, $q^P = .5$, b-peer = 1



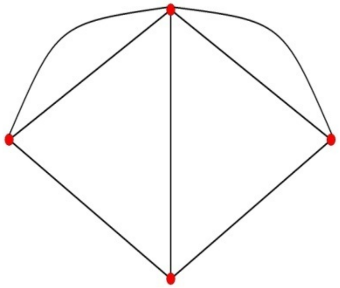
Estimated parameters:



- Information significant, peer/endorse effect not

	qN	qP	b-peer	qN – qP
Estimates:	0.05***	0.55***	-0.20	-0.50***
	[0.01]	[0.13]	[0.16]	[0.13]

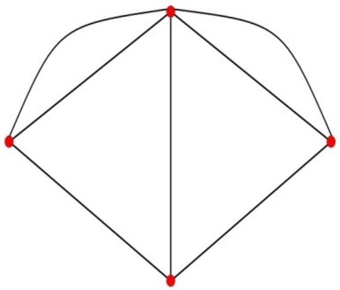
Estimated parameters:



- Information significant, peer/endorse effect not

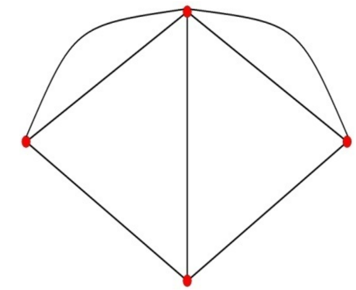
	qN	qP	b-peer	qN – qP
With Diffusion	0.05***	0.55***	-0.20	-0.50***
	[0.01]	[0.13]	[0.16]	[0.13]
just peer:			2.5***	

Results from Fitting Model of Diffusion in this case:



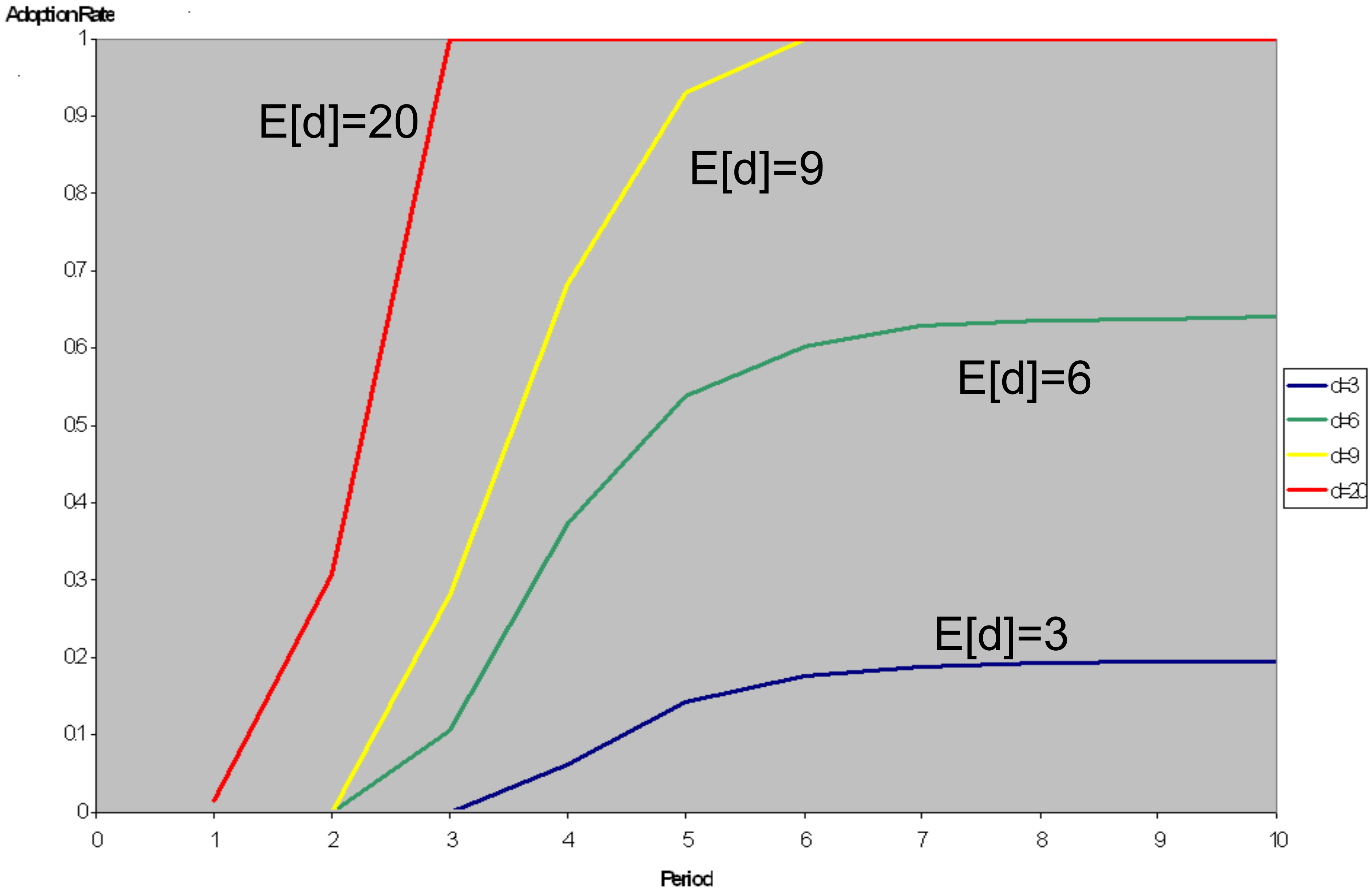
- Significant information passing parameters
- Insignificant, limited Peer Effects
- Information passing depends on whether participate: more likely if participate
- Nonparticipants play a substantial role (1/3 of total)

Broader Messages:



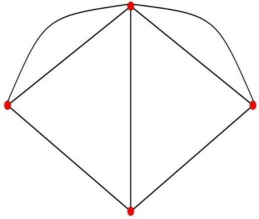
- Simple network models can help *estimate and dissect* peer effects and diffusion processes: policy consequences

- Network structures have consequences for behavior:
  - Tractable and intuitive ways to quantify despite complexity of networks



fraction adopting over time, $P(d) = ad^{-2}$,
Simulated diffusion process, threshold of neighbors

Approaches



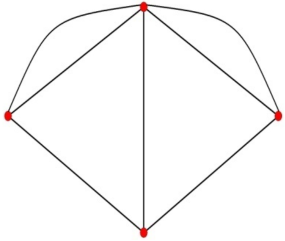
- Field/natural experiments (e.g., pseudo random injections in Indian data, identification – *but don't control networks*…)

- IV (Just saw in Lecture 2)
  - exploiting network position (Bramoulle, Djebbari, and Fortin, - *does not address endogenous networks/unobservables*…)
  - things that affect network, but not behavior (Acemoglu, Garcia-Jimeno, and Robinson - rare …)

- **Structural modeling of behavior (e.g., diffusion model…)**

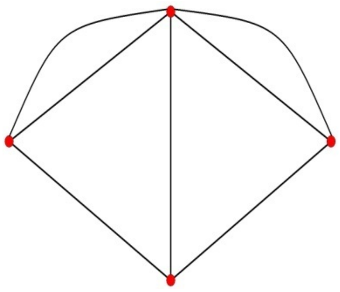
- **Model network formation**…

Network Formation



- Main challenges driving current literature
  - multiplicity
  - integrating formation with behavior: unobservables
  - link dependencies!!

Questions



- Always lurking: correlated unobservables

- Peoples' behaviors correlate with network position because of homophily

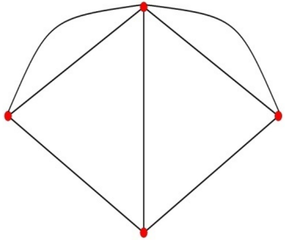
Example

- Goldsmith-Pinkham and Imbens (2013)

$$Y_i = b_0 + b_1 X_i + b_2 Y_{(i)}^{peer} + b_3 X_{(i)}^{peer} + b_4 Z_i + e_i$$

Z_i *unobserved* characteristics

Example

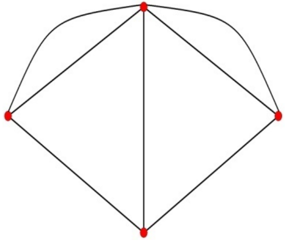


- Utility from friendship based on homophily:

$$U_i\,(j) = a_0 + a_1 \mid X_i - X_j \mid + a_2 \mid Z_i - Z_j \mid + \ldots + e_{ij}$$

(... = past network relationships if available, e.g., past friends in common, linked in past)

Estimate Unobservables



$$Y_i = b_0 + b_1 X_i + b_2 Y_{(i)}^{peer} + b_3 X_{(i)}^{peer} + b_4 Z_i + e_i$$

$$U_i(j) = a_0 + a_1 |\ X_i - X_j\ | + a_2\ |\ Z_i - Z_j\ | + \ldots + e_{ij}$$

Links logistic in $U_i(j)$, $U_j(i)$

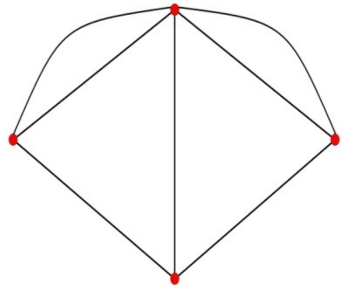
Estimate system (Bayesian, MLE)

Infer unobservable Z_i's :

ij connected with distant X_i's have similar Z_i's

ij unlinked with similar X_i's have differing Z_i's

Lesson:

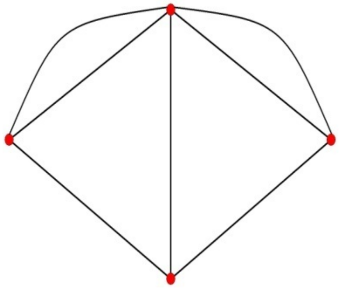


$$Y_i = b_0 + b_1X_i + b_2Y_{(i)}^{peer} + b_3X_{(i)}^{peer} + b_4Z_i + e_i$$

- Accounting for link formation can help infer unobservables

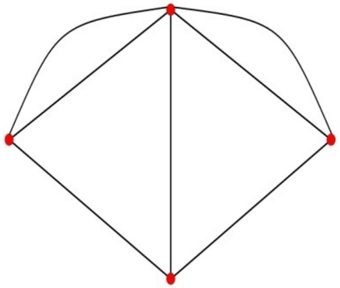
- Can help correct estimates of strategic interaction with friends/acquaintances

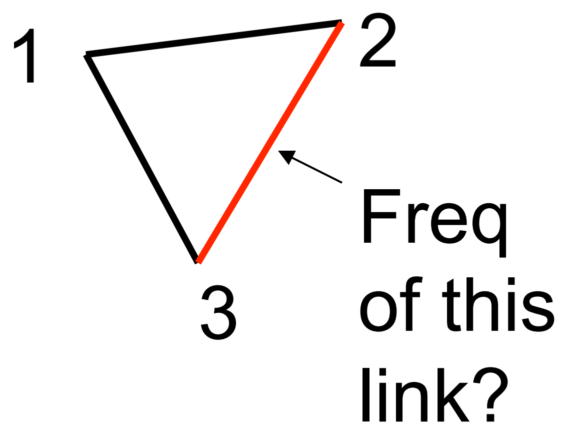
Link Dependencies



- Link formation is significantly correlated!
- Friends of friends
- Value to having closure (enforcement of incentives...)

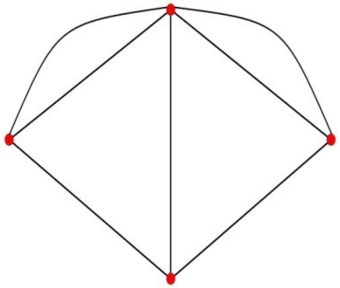
Link Dependencies - Clustering Coefficients:





- Prison friendships
  - **.31** (MacRae 60) vs .0134
- Co-authorships
  - **.15** math (Grossman 02) vs .00002,
  - **.09** biology (Newman 01) vs .00001,
  - **.19** econ (Goyal, van der Leij, Moraga 06) vs .00002,
- Florentine Marriage and Business dealings
  - **.46** on 15 central families vs .29...

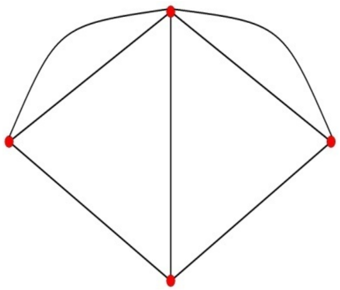
Challenges



- No longer talk about probabilities at link level

- But cannot calculate probabilities at network level: too many networks to do MLE/Bayesian calculations!!

Models of Network Formation with Dependencies

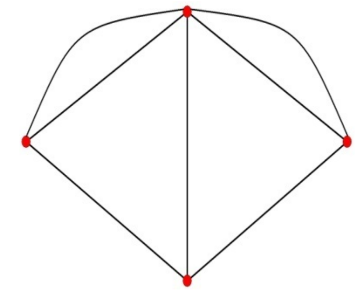


- Dynamic/Specific Models (Jackson-Wolinsky 96, Barabasi-Albert 99, Bala-Goyal 00, Jackson-Watts 00, Jackson-Rogers 07, Currarini-Jackson-Pin 09,10, Christakis et al. 10, Bramoulle et al. 12, Mele 12…)

- ERGMs (Frank-Strauss 86, Wasserman-Pattison 96, Snjiders 02, Handcock 03…) estimation problems!

- Subgraphs, probabilities of seeing specific configurations of links (Chandrasekhar-Jackson 13)

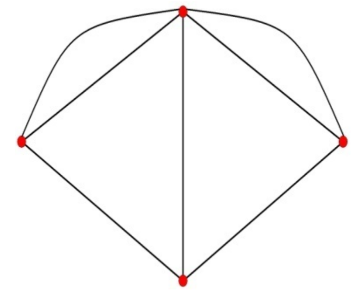
Broader Messages:



- Simple network models can help *estimate and dissect* peer effects and diffusion processes: policy consequences

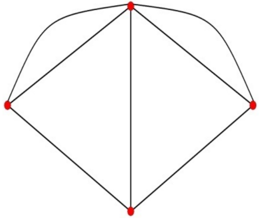
- Network structures have consequences for behavior:
    - Tractable and intuitive ways to quantify despite complexity of networks

Simplifying the Complexity



- Global patterns of networks
  - path lengths
  - degree distributions...
- Segregation Patterns: node types and homophily
- Local Patterns
  - Clustering
  - Support...
- Positions in networks
  - neighborhoods
  - Centrality, influence...

Identification



- Field/natural experiments (e.g., pseudo random injections in Indian data, identification – *but don't control networks*…)

- IV   (Just saw in Lecture 2)
  - exploiting network position (Bramoulle, Djebbari, and Fortin, - *does not address endogenous networks/unobservables*…)
  - things that affect network, but not behavior (Acemoglu, Garcia-Jimeno, and Robinson - rare …)

- Structural modeling of behavior (e.g., diffusion model…)

- Model network formation...