Game Theoretic Modeling and Social Networks

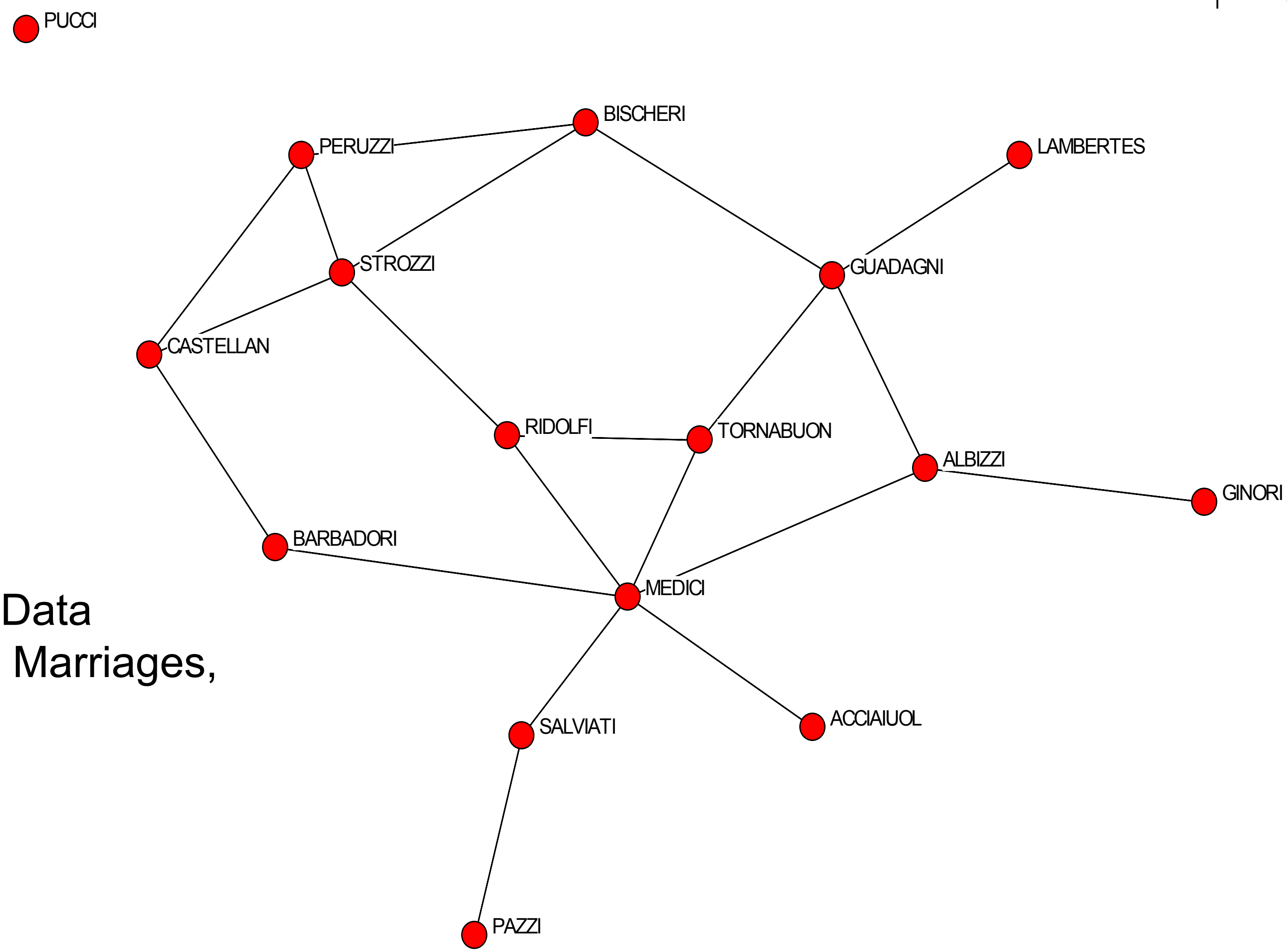
Matthew O. Jackson

Nemmers Conference

Modeling Social Networks: Where we are and where to go

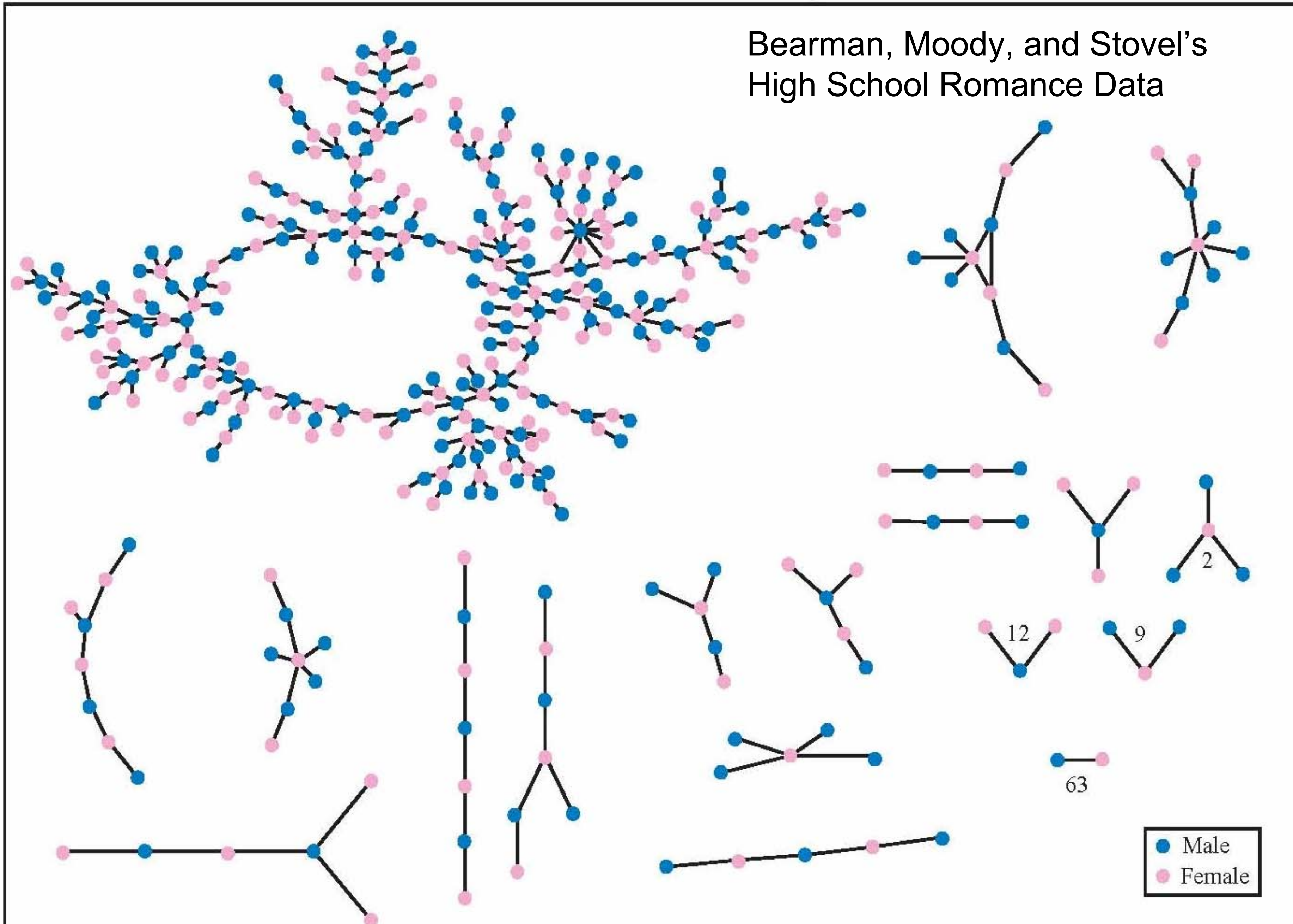
- Some empirical background
- What are the interesting questions?
- Random graph models
  - a few representative examples
  - strengths and weaknesses
- Strategic/Game Theoretic models
  - a few representative examples
  - strengths and weaknesses
- Hybrids and the future

Examples of Social and Economic Networks

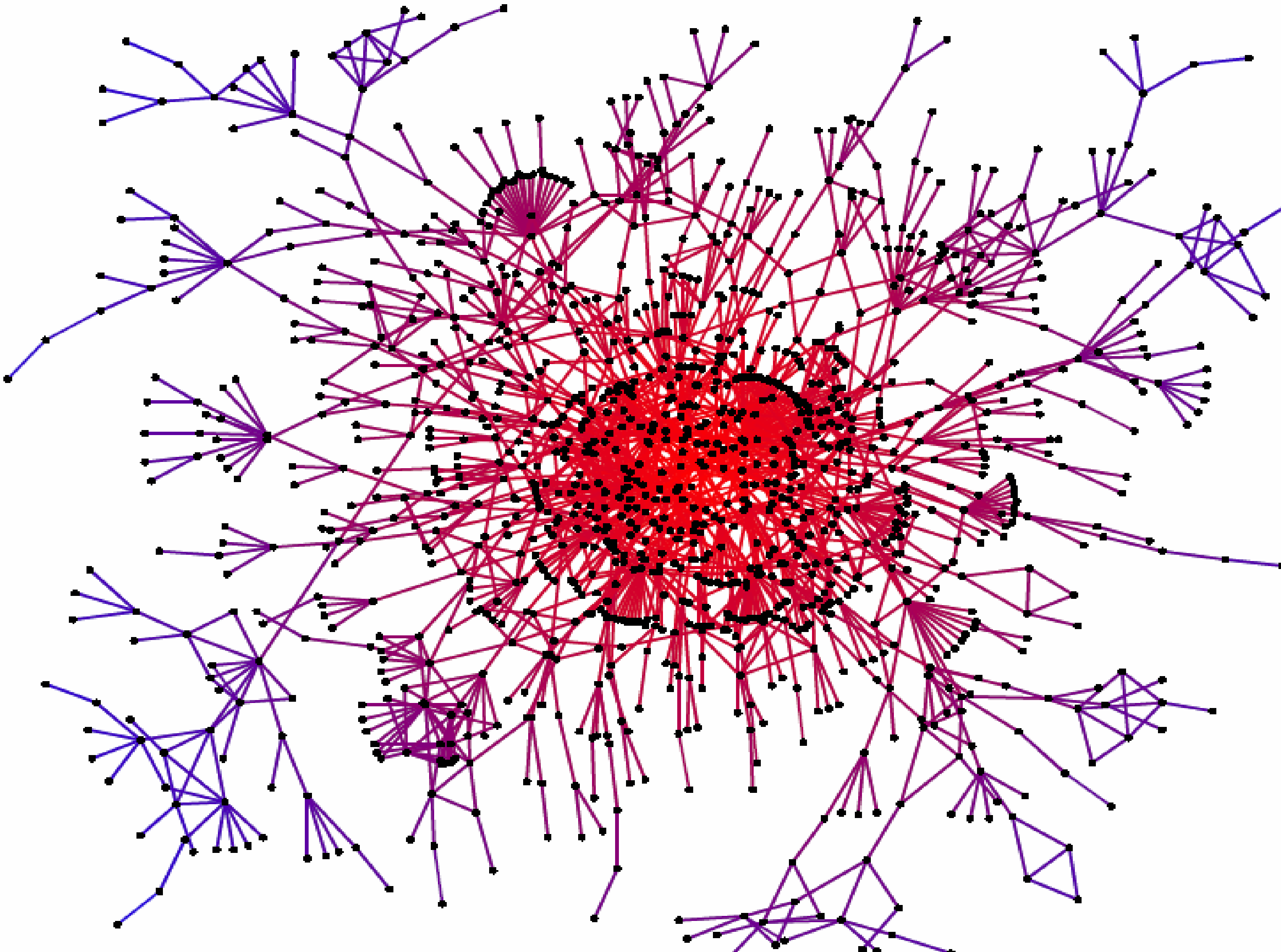


Padgett’s Data
Florentine Marriages, 1430’s

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



Adamic – Stanford homepage links (largest component)



What do we know?

- Networks are prevalent
  - Job contact networks, crime, trade, politics, ...
- Network position and structure matters
  - rich sociology literature
  - Padgett example – Medicis not the wealthiest nor the strongest politically, but the most central
- ``Social'' Networks have special characteristics
  - small worlds, degree distributions...

Networks in Labor Markets

- Myers and Shultz (1951)- textile workers:
  - 62% first job from contact
  - 23% by direct application
  - 15% by agency, ads, etc.
- Rees and Shultz (1970) – Chicago market:
  - Typist 37.3%
  - Accountant 23.5%
  - Material handler 73.8%
  - Janitor 65.5%, Electrician 57.4%…
- Granovetter (1974), Corcoran et al. (1980), Topa (2001), Ioannides and Loury (2004) ...

Other Settings

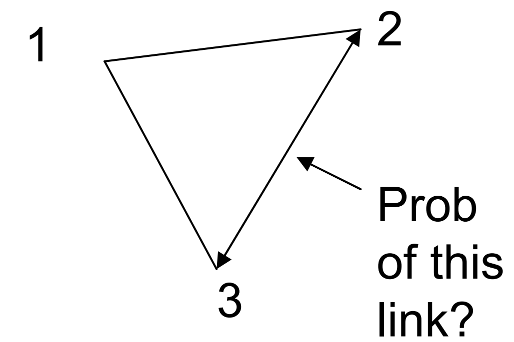
- Networks and social interactions in crime:
  - Reiss (1980, 1988) - 2/3 of criminals commit crimes with others
  - Glaeser, Sacerdote and Scheinkman (1996) - social interaction important in petty crime, among youths, and in areas with less intact households
- Networks and Markets
  - Uzzi (1996) - relation specific knowledge critical in garment industry
  - Weisbuch, Kirman, Herreiner (2000) – repeated interactions in Marseille fish markets
- Social Insurance
  - Fafchamps and Lund (2000) – risk-sharing in rural Phillipines
  - De Weerdt (200
- Sociology literature – interlocking directorates, aids transmission, language, ...

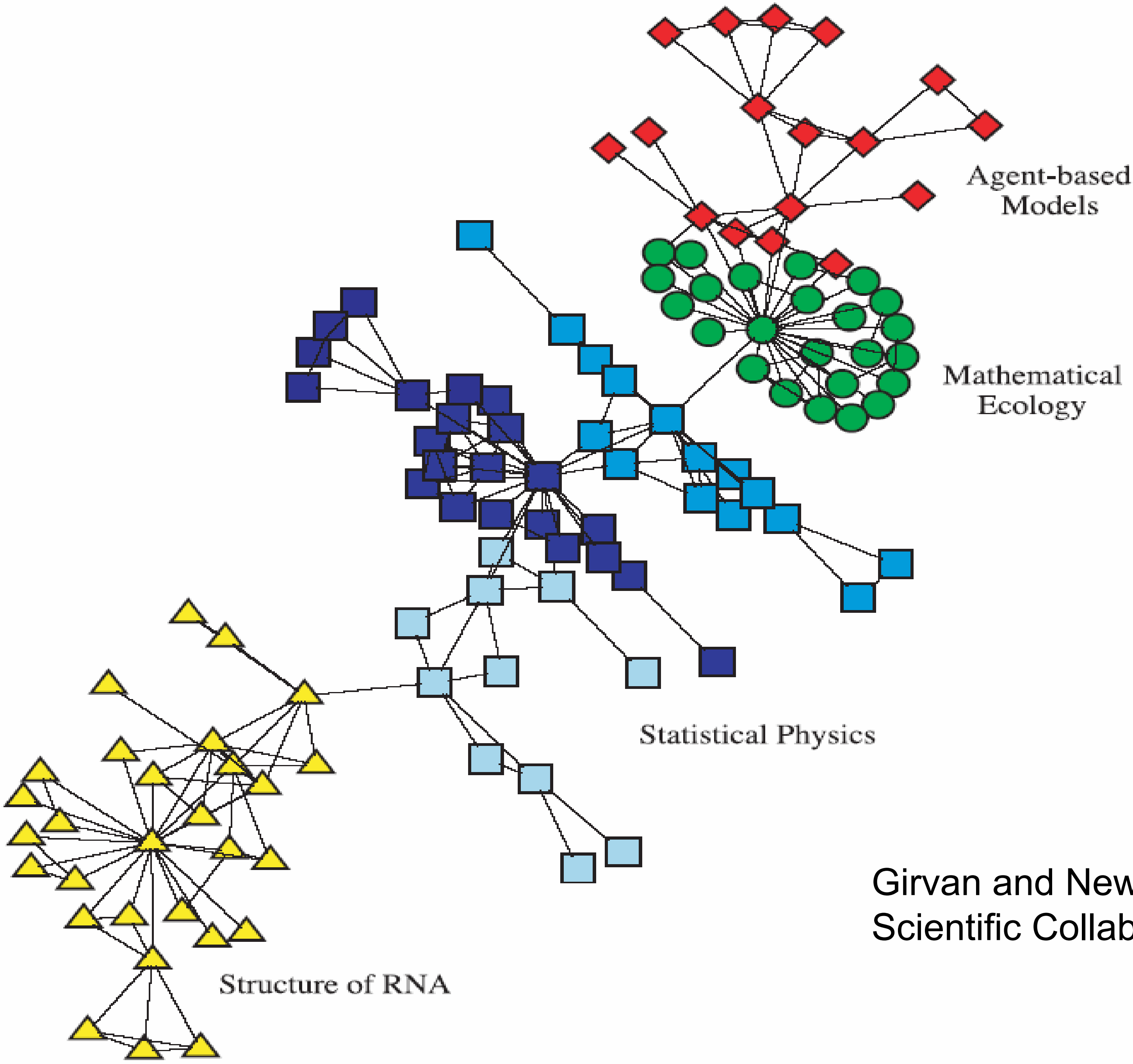
Stylized Facts: Small diameter

- Milgram (1967) letter experiments
  - median 5 for the 25% that made it
- Actors in same movie (Kevin Bacon Oracle)
  - Watts and Strogatz (1998) – mean 3.7
- Co-Authorship studies
  - Grossman (1999) Math mean 7.6, max 27,
  - Newman (2001) Physics mean 5.9, max 20
  - Goyal et al (2004) Economics mean 9.5, max 29
- WWW
  - Adamic, Pitkow (1999) – mean 3.1 (85.4% possible of 50M pages)

High Clustering Coefficients - distinguishes ``social'' networks

- Watts and Strogatz (1998)
  - .79 for movie acting
- Newman (2001) co-authorship
  - .496 CS, .43 physics, .15 math, .07 biomed
- Adamic (1999)
  - .11 for web links (versus .0002 for random graph of same size and avg degree)





Girvan and Newman's
Scientific Collaboration Data

Distribution of links per node: Power Laws

- Plot of log(frequency) versus log(degree) is ``approximately'' linear in upper tail

- prob(degree) = c degree^{-a}
  - log[prob(degree)] = log[c] – a log[degree]

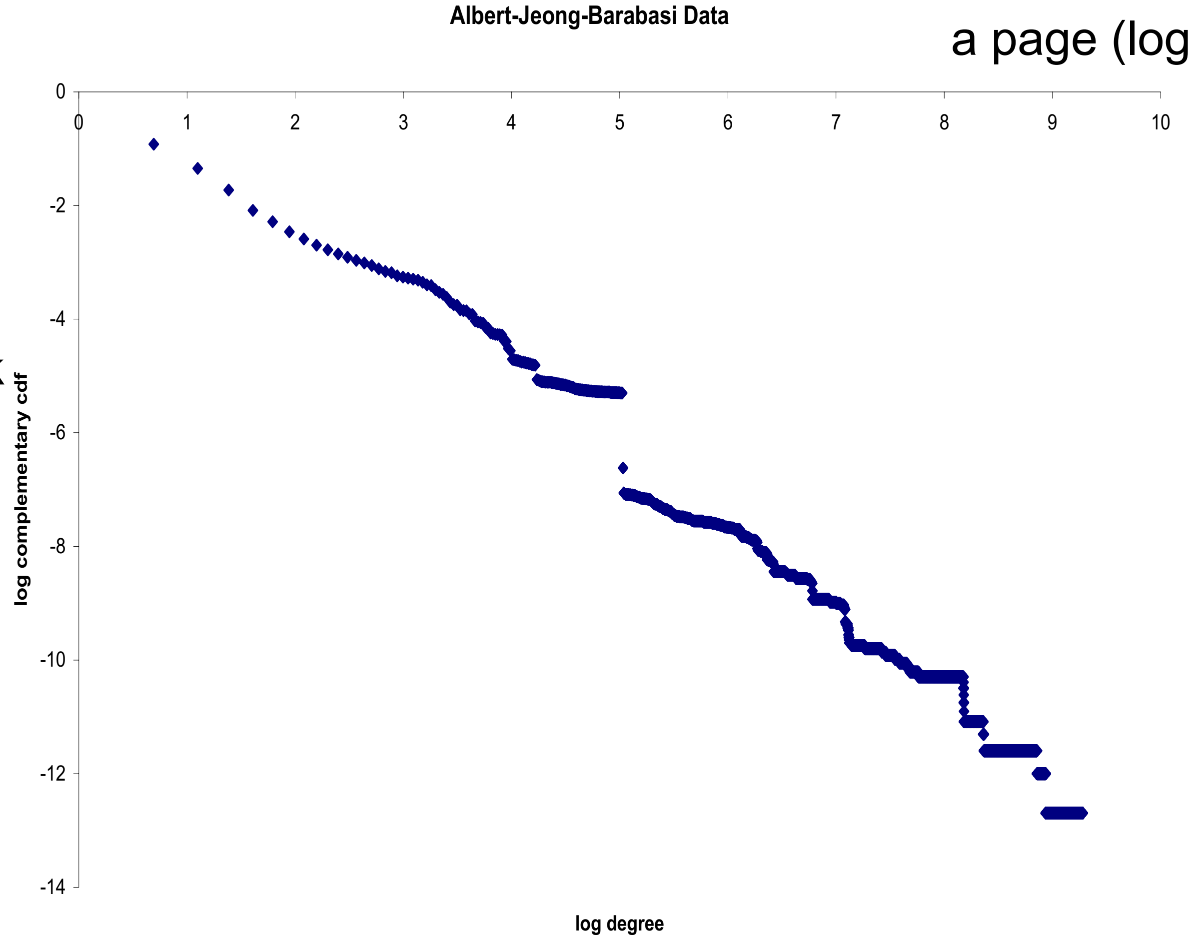
- Fat tails compared to random network

- Related to other settings: Pareto (1896), Yule (1925), Zipf (1949), Simon (1955),

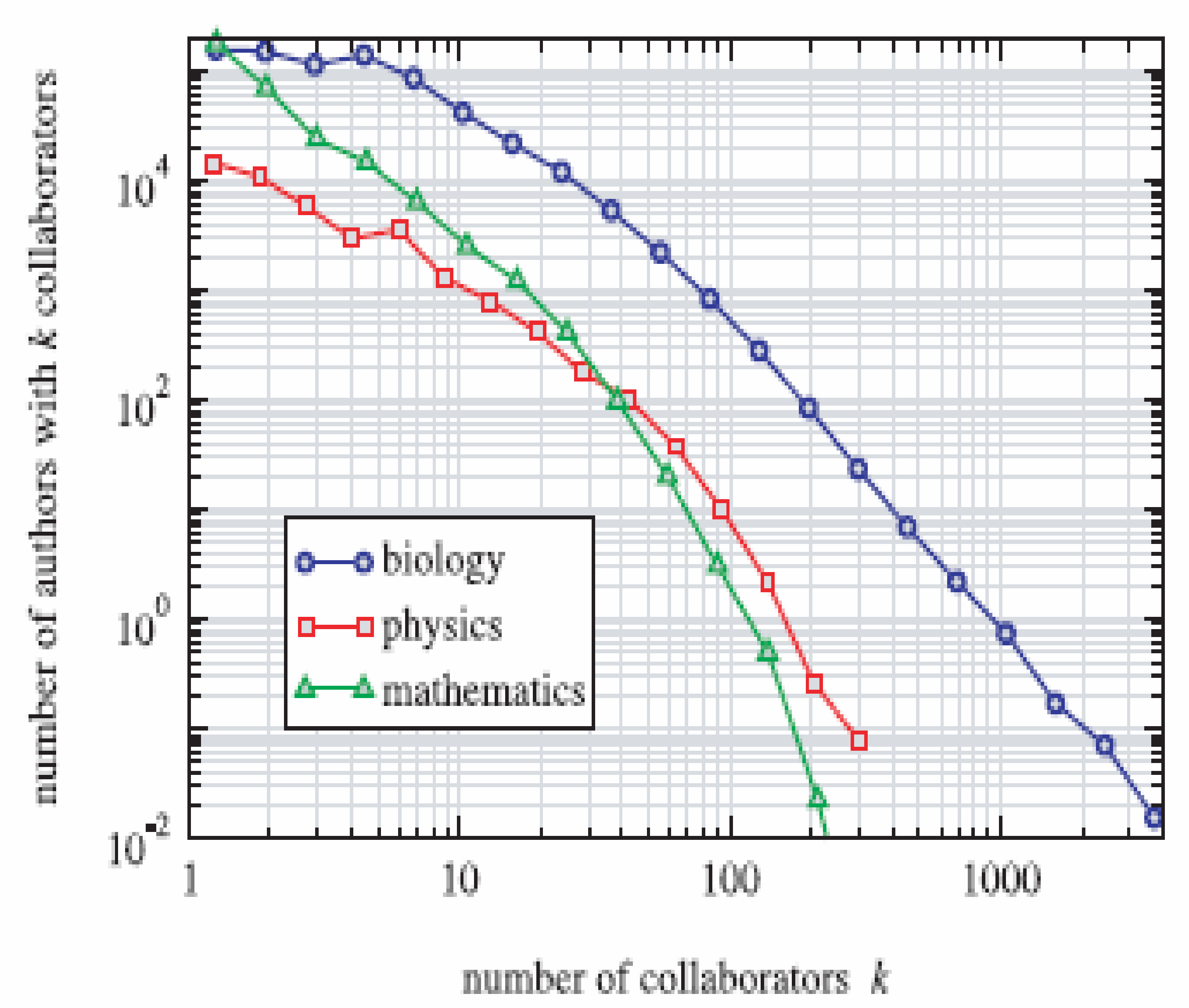
Degree – ND www Albert, Jeong, Barabasi (1999)

number of links to a page (log scale)

fraction of pages with more than k links (log)



Co-Authorship Data, Newman and Grossman



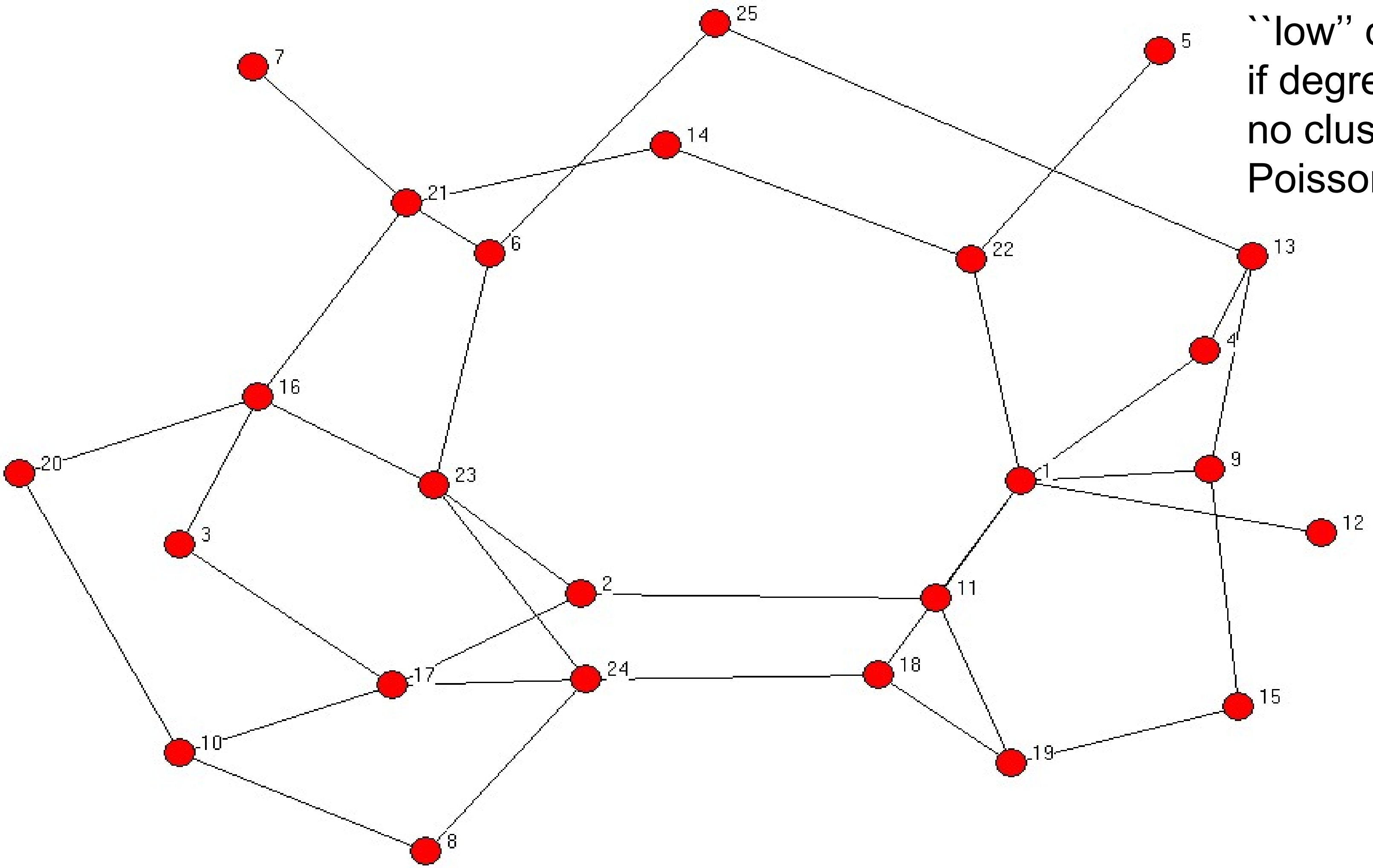
Three Key Questions:

- How does network structure affect interaction and behavior?

- Which networks form?
  - Game theoretic reasoning
  - dynamic random models

- When do efficient networks form?
  - Intervention - design incentives?

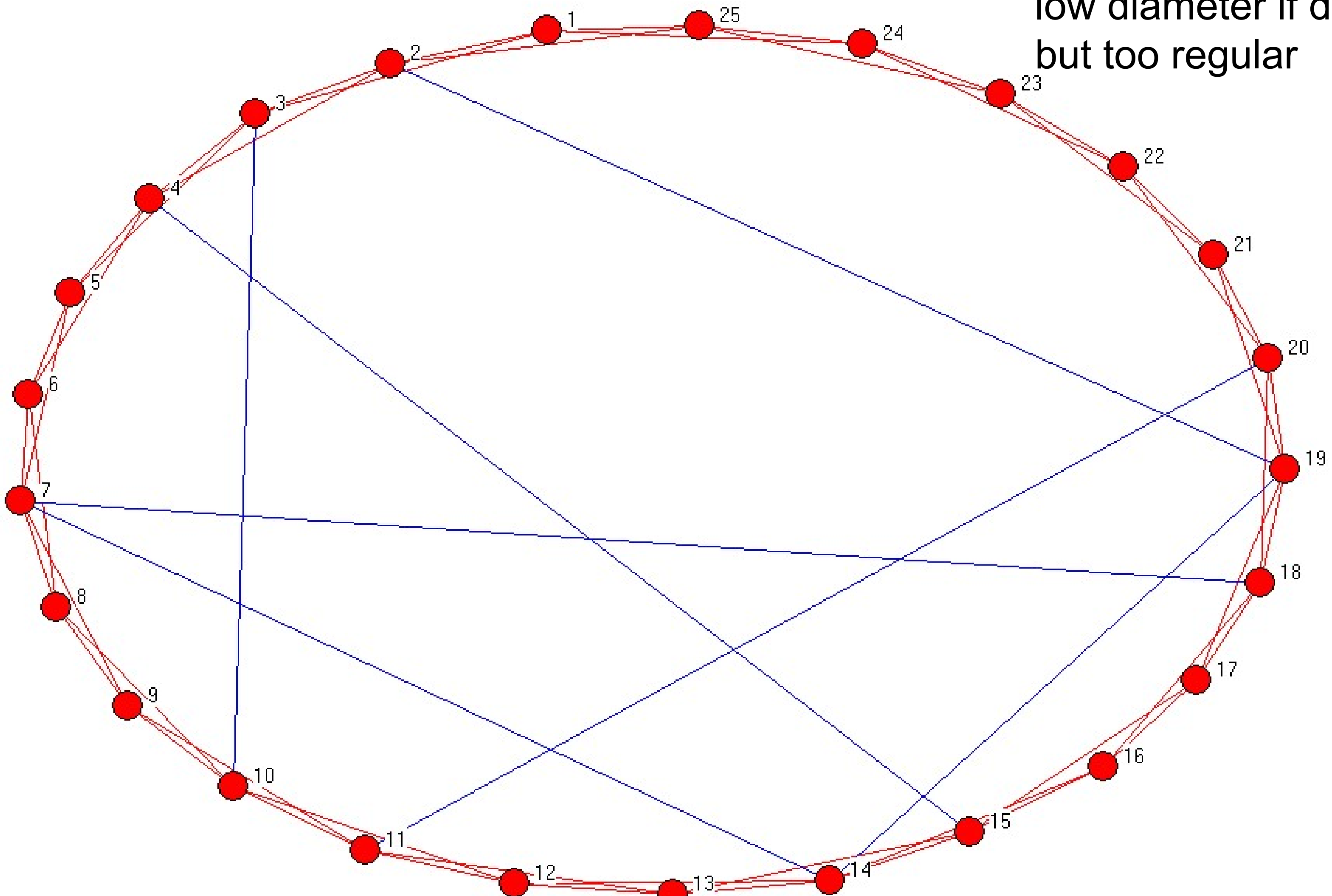
Random Graphs: Bernoulli (Erdos and Renyi (1960))



``low'' diameter
if degree is high,
no clustering,
Poisson degree

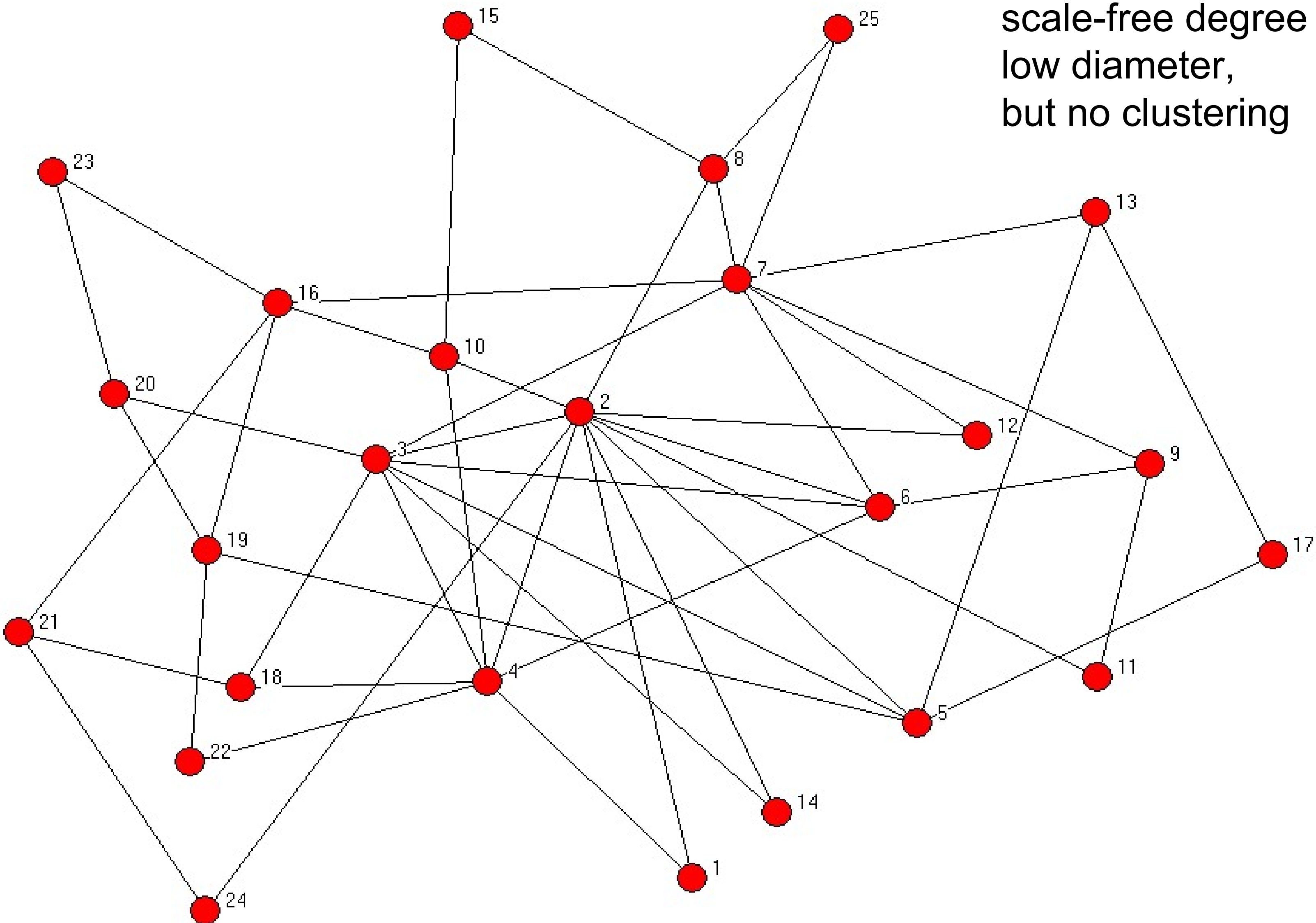
Rewired lattice (Watts and Strogatz (1999))

high clustering
low diameter if degree is high
but too regular



Preferential Attachment (Barabasi and Albert (2001))

scale-free degree distribution
low diameter,
but no clustering



Advantages of Random Graph Models

- Generate large networks with well identified properties
- Mimic real networks (at least in some characteristics)
- Tie a specific property to a specific process

What's Missing From Random Graph Models?

- The ``Why''?
  - Why this process? (lattice, preferential attach...)
- Implications of network structure: economic and social context or relevance?
  - welfare and how can it be improved...
- Careful Empirical Analysis
  - ``Scale-Free'' may not be
  - No fitting of models to data (models aren't rich enough to fit across applications)

Economic/Game Theoretic Models

- Welfare analysis – agents get utility from networks
  - $u_i(g)$
  - Efficient Networks: argmax $\sum u_i(g)$

- Decision making agents form links and/or choose actions

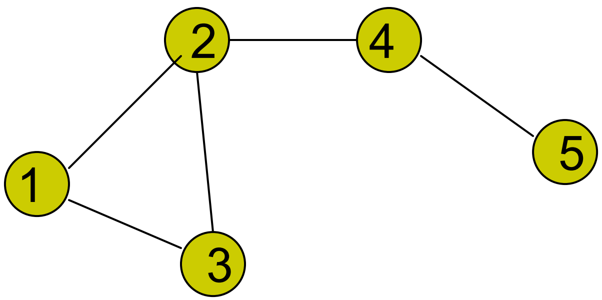
Example: Connections Model

Jackson and Wolinsky (1996):

- benefit from a friend is δ
- benefit from a friend of a friend is δ^2,...
- cost of a link is c

$u_2 = 3\delta + \delta^2 - 3c$

$u_1 = 2\delta + \delta^2 + \delta^3 - 2c$



$u_5 = \delta + \delta^2 + 2\delta^3 - c$

- Pairwise Stable networks
  - $u_i(g) \geq u_i(g\text{-}ij)$ for each i and ij in g
  - $u_i(g+ij) \geq u_i(g)$ implies $u_j(g+ij) \geq u_j(g)$ for each ij not in g

Efficient Networks

- low cost: $c < \delta - \delta^2$
  - complete network is efficient
- medium cost: $\delta - \delta^2 < c < \delta + (n-2)\delta^2/2$
  - star network is efficient
    - minimal number of links to connect
    - connection at length 2 is more valuable than at 1 ($\delta - c < \delta^2$)
- high cost: $\delta + (n-2)\delta^2/2 < c$
  - empty network is efficient

Pairwise Stable Networks:

- low cost: $c< \delta-\delta^2$
  - complete network is pairwise stable (and efficient)

- medium/low cost: $\delta-\delta^2 < c < \delta$
  - star network is pairwise stable (and efficient)
  - others are also pairwise stable

- medium/high cost: $\delta< c < \delta+(n-2)\delta^2/2$
  - star network is not pairwise stable (no loose ends)
  - nonempty pairwise stable networks are over-connected and may include too few agents

- high cost: $\delta+(n-2)\delta^2/2 < c$
  - empty network is pairwise stable (and efficient)

Some Settings stable=efficient

Buyer-Seller Networks: Kranton-Minehart (2002):

- Sellers each with one identical object
- Buyers each desire one object, private valuation
- buyers choose to link to sellers at a cost
- sellers hold simultaneous ascending auctions

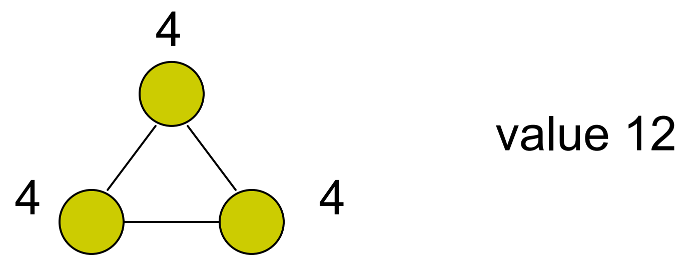
Example: values iid U[0,1], 1 seller

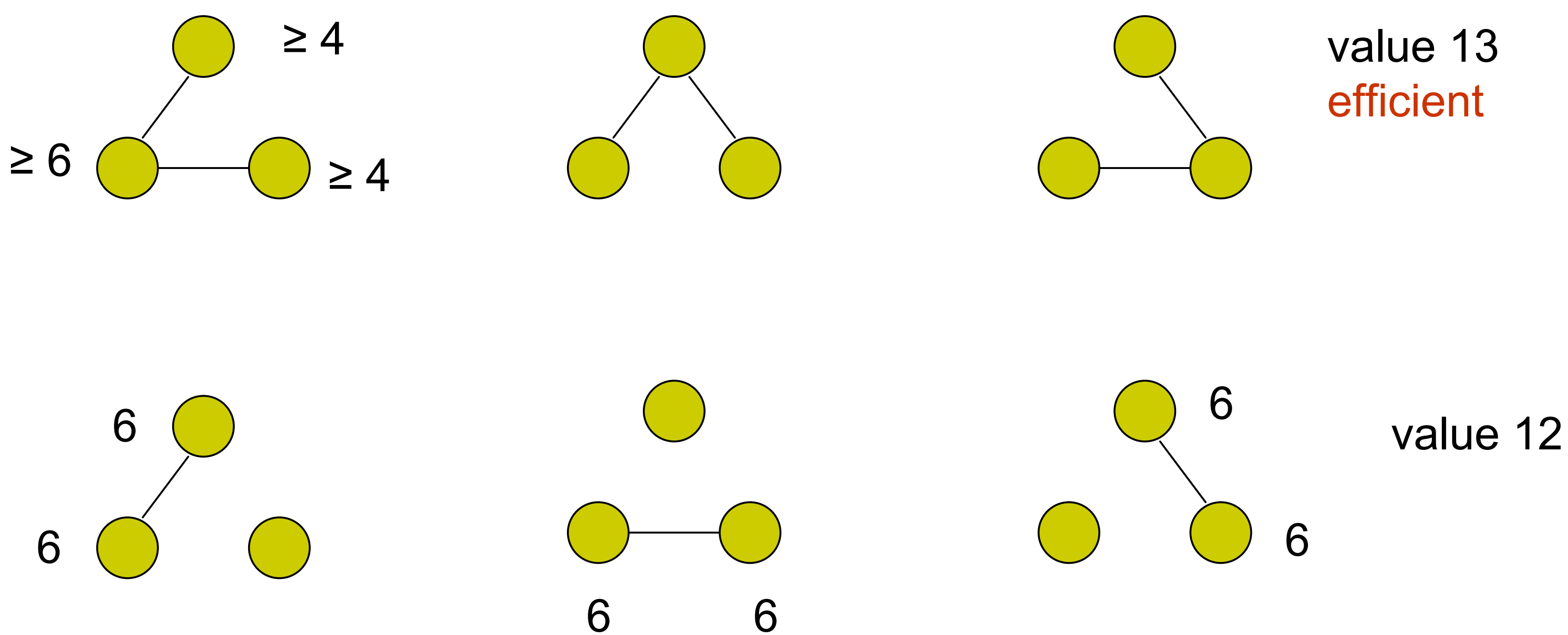
	Each buyer's expected utility	Seller's expected utility	Total social value
n buyers	1/[n(n+1)]	(n-1)/(n+1)	n/(n+1)
n+1 buyers	1/[(n+1)(n+2)]	n/(n+2)	(n+1)/(n+2)
change	-2/[n(n+1)(n+2)]	2/[(n+1)(n+2)]	1/[(n+1)(n+2)]

Transfers cannot always help

anonymity: same transfers
to identical players

balance: no transfers
outside of component





Rich literature on such issues

- loosen anonymity (Dutta-Mutuswami (1997))
- directed networks (Bala-Goyal (2000), Dutta-Jackson (2000),...)
- bargaining when forming links (Currarini-Morelli(2000), Slikker-van den Nouweland (2000), Mutuswami-Winter(2002), Bloch-Jackson (2004))
- dynamic models (Aumann-Myerson (1988), Watts (2001), Jackson-Watts (2002ab), Goyal-Vega-Redondo (2004), Feri (2004), Lopez-Pintado (2004),...)
- farsighted models (Page-Wooders-Kamat (2003), Dutta-Ghosal-Ray (2003), Deroian (2003),...)
- allocating value (Myerson (1977), Meessen (1988), Borm-Owen-Tijs (1992), van den Nouweland (1993), Qin (1996), Jackson-Wolinsky (1996), Slikker (2000), Jackson (2005)...)
- modeling stability (Dutta-Mutuswami (1997), Jackson-van den Nouweland (2000), Gilles-Sarangi (2003ab), Calvo-Armengol and Ikilic (2004),...)
- experiments (Callander-Plott (2001), Corbae-Duffy (2001), Pantz-Zeigelmeyer (2003), Charness-Corominas-Bosch-Frechette (2001), Falk-Kosfeld (2003), ...)

Models of Networks in Context

- crime networks (Glaeser-Sacerdote-Scheinkman (1996), Ballester, Calvo, Zenou (2003),...)
- markets (Kirman (1997), Tesfatsion (1997), Weisbach-Kirman-Herreiner (2000), Kranton-Minehart (2002), Corominas-Bosch (2005), Wang-Watts (2002), Galeotti (2005),Kakade et al (2005)...)
- labor networks (Boorman (1975), Montgomery (1991, 1994), Calvo (2000), Arrow-Borzekowski (2002), Calvo-Jackson (2004,2005), Cahuc-Fontaine (2004), Currie...)
- insurance (Fafchamps-Lund (2000), DeWeerdt (2002), Bloch-Genicot-Ray (2004),...
- IO (Bloch (2001), Goyal-Moraga (2001), Goyal-Joshi (2001), Belleflamme-Bloch (2002),Billard-Bravard (2002), ...)
- international trade (Casella-Rauch (2001), Furusawa-Konishi (2003),
- public goods (Bramoulle-Kranton (2004)
- airlines (Starr-Stinchcombe (1992), Hendricks-Piccione-Tan (1995))
- network externalities in goods (Katz-Shapiro (1985), Economides (1989, 1991) , Sharkey (1991)...)
- organization structure (Radner (), Radner-van Zandt (), Demange (2004)...)
- learning (Bala-Goyal (1998), Morris (2000), DeMarzo-Vayanos-Zweibel (2003), Gale-Kariv (2003), Choi-Gale-Kariv (2004),...)

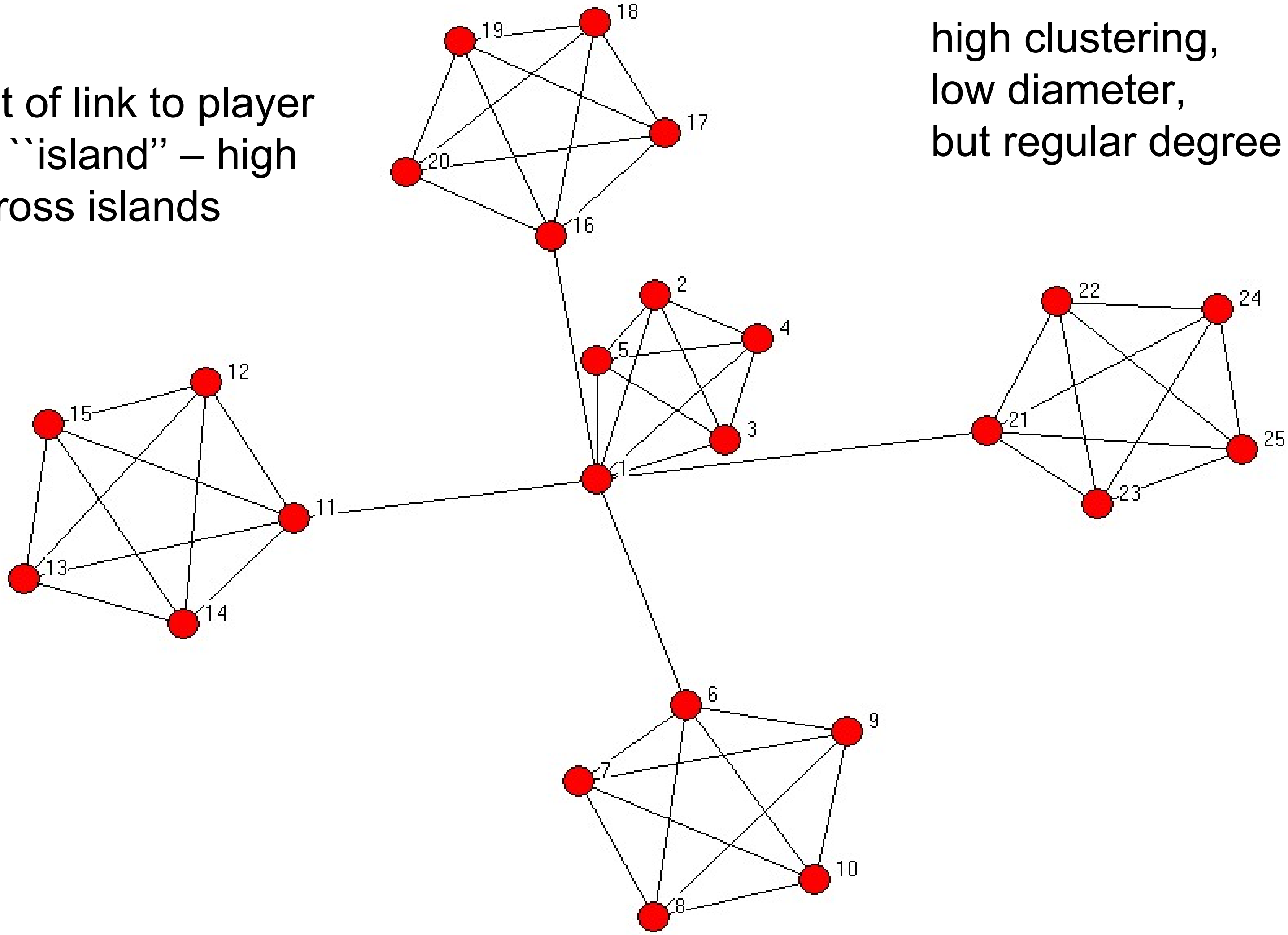
Can economic models match observables?

- Small worlds related to costs/benefits
  - low costs to local links – high clustering
  - high value to distant connections – low diameter

Geographic Connections (Johnson-Gilles (2000), Carayol-Roux (2003), Galeotti-Goyal-Kamphorst (2004), Jackson-Rogers (2004))

low cost of link to player on own ``island'' – high cost across islands

high clustering,
low diameter,
but regular degree



Advantages of an economic approach

- Payoffs allow for a welfare analysis
  - Identify tradeoffs – incentives versus efficiency
- Tie the nature of externalities to network formation...
- Put network structures in context
- Account for (and *explain*) some observables

What's missing from Game theoretic models?

- Stark network structures emerge
  - need more heterogeneity
- over-emphasize choice versus chance determinants for *large* applications?
- more on network structure and outcomes

Hybrid Models Needed

- Build richer models with random/heterogeneity
- allow for welfare analysis
- take model to data and fit observed networks
- relate structure to outcomes

Example: can we learn about welfare from fitting networks? (w Rogers)

- Nodes are players
- Indexed by date of birth t={1,2,3,...}
- Find m_r other nodes at random
- Search their neighborhoods to find m_s more nodes
  - think of entering at a random web page and following its links
- Attach to a given node if net utility is positive
  - random utility or
  - increasing in node's degree

Degree Distribution

Expected increase in the in-degree of a node i

prob found at random

prob found through search

$$p\,(\,m_r/t \;+\; d_i\;[m_s/(t\,m)])$$

prob linked to given found

number of neighbors

prob my neighbor is entry point

m – average links/node,   r – ratio random/search

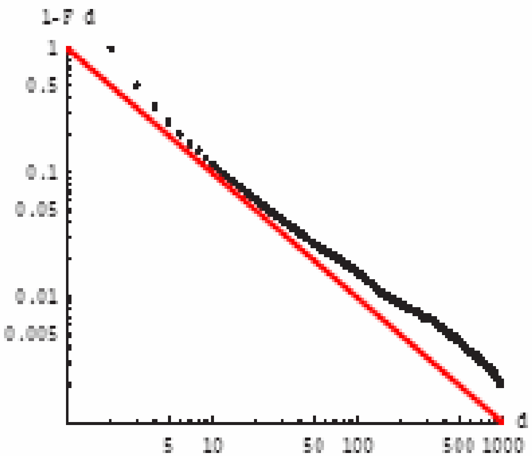
Proposition (Mean field)

The degree distribution of the mean field approximation to the process has a degree distribution having complementary cdf of

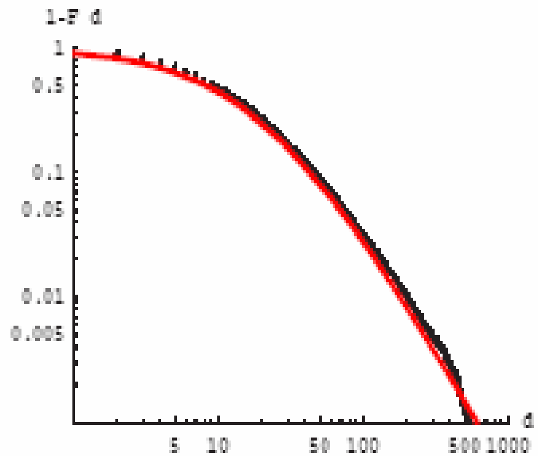
$$F(d) = 1 - (rm)^{1+r} (d + rm)^{-(1+r)}$$

Clustering is bounded away from 0 and decreasing in r

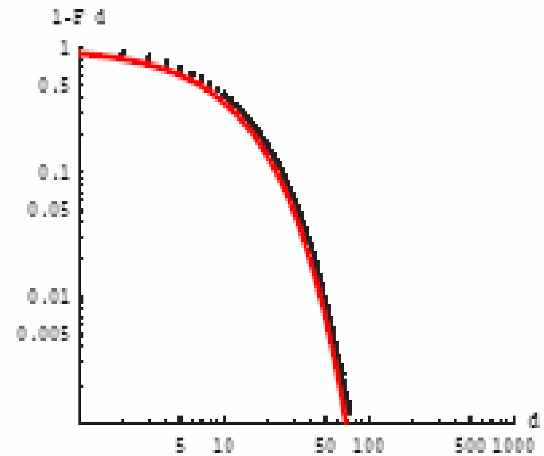
Varying the relative Random and Search probabilities



r=0



r=1

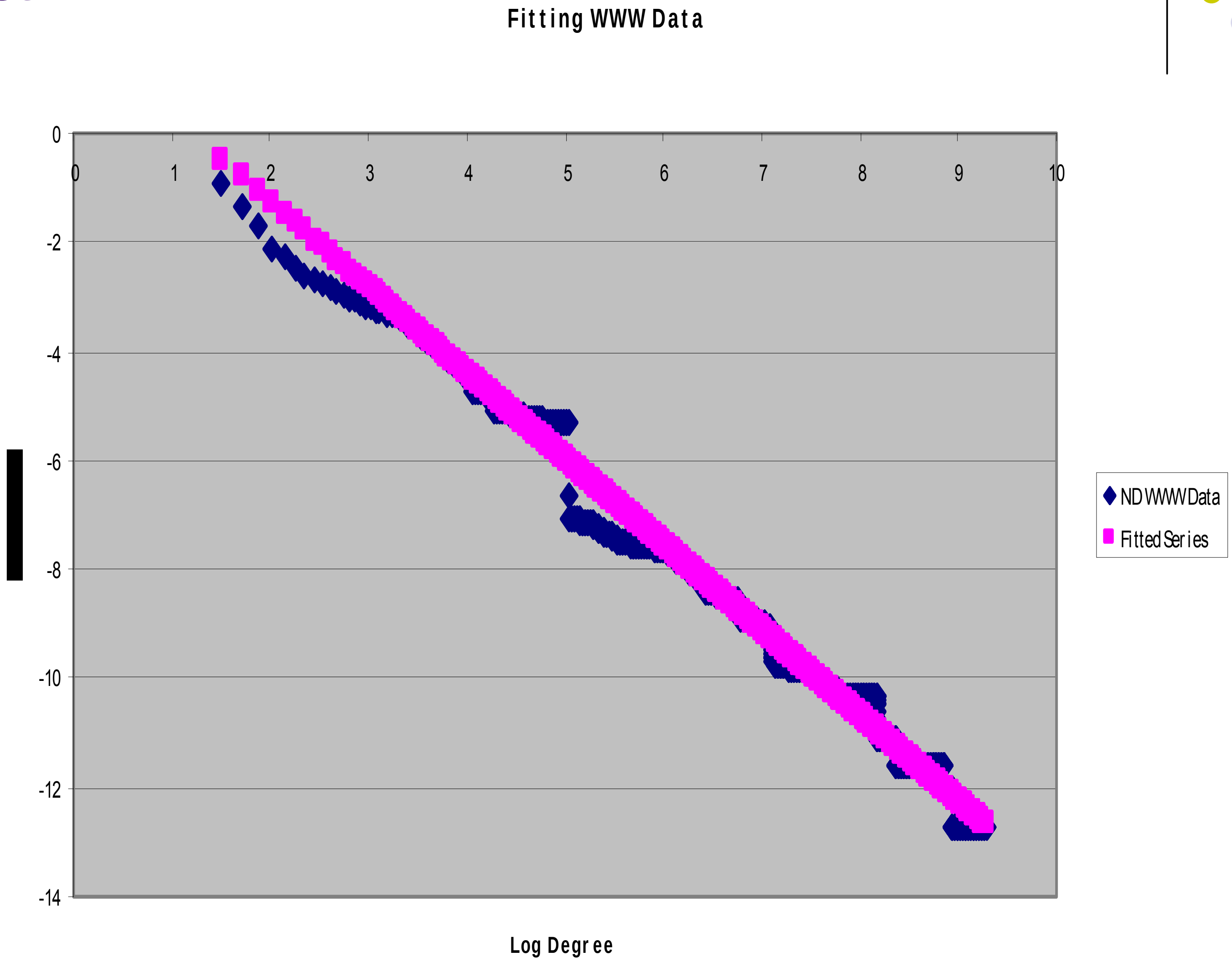


r= ∞

Fitting the Data

- fix our m by direct calculation from data
- estimate r by fitting the degree distribution
- examine implied clustering coefficients and compare to data
- simulate the model to get accurate estimates for diameter
- other characteristics?

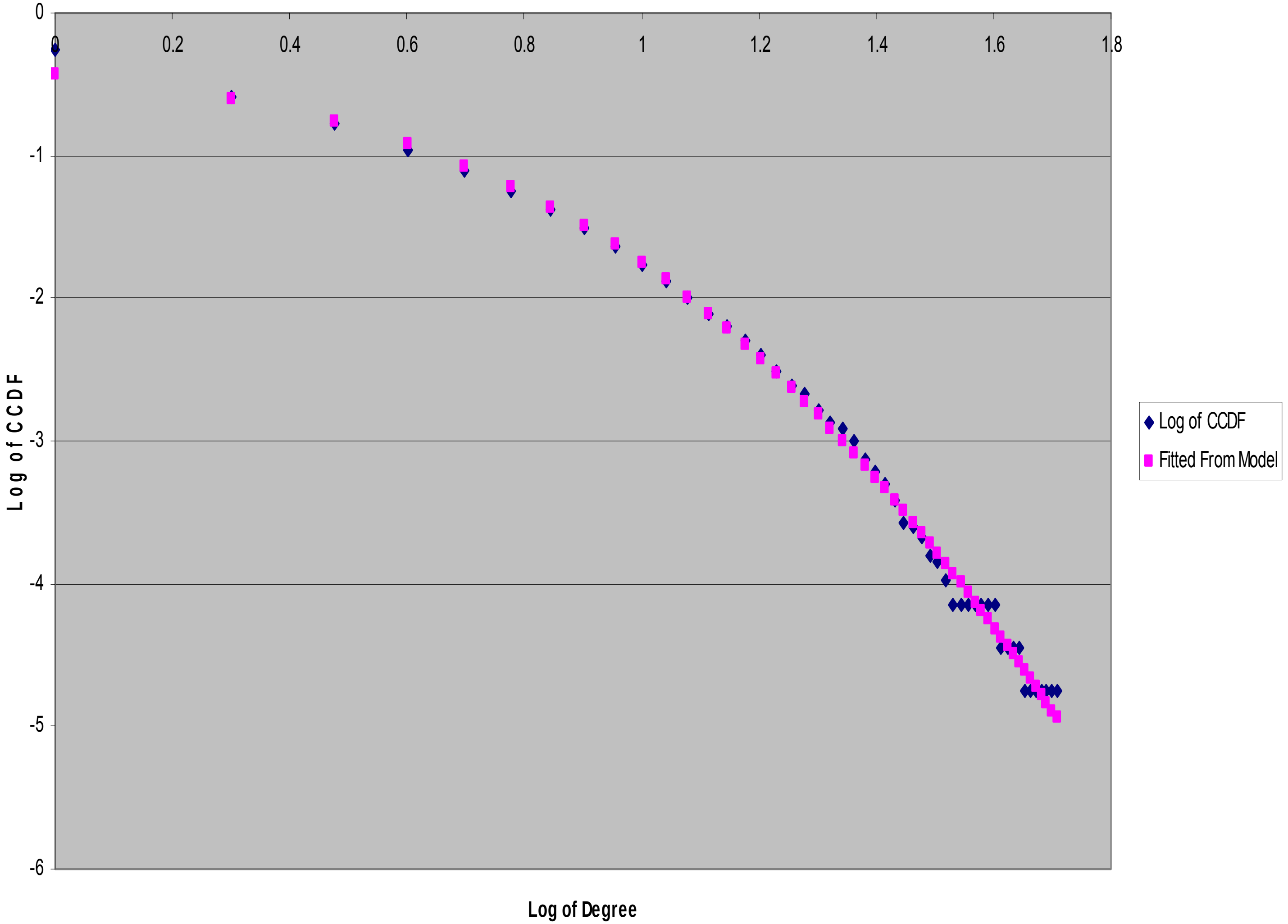
Comparison: fitting the www data



Other Characteristics

- m=5 on average in data
- our estimate for r = .5 (R^2 is .97)
- average clustering .11 (at p=1/3)
  - data .11 Adamic
- total clustering goes to 0
  - data?
- diameter: bracketed 16 to 32
  - data 20

Fitting the Model to Data: co-author data of Goyal et al



Comparisons:

- Random/Search:
  - WWW links: r=.5
  - Small World Citation: r=.62
  - Econ co-authors: r=3.5
  - Ham radio: r=5
  - Prison Friendships: r=590
  - High School Romances: r=1000

Relating Network structure to outcomes

- Diffusion of viruses, information, behavior...
  - Bailey (1975), Pastor-Satorras and Vespignani (2001), Lopez-Pintado (2003), ..., SIS models
- Model relates network to outcomes
  - Higher r degree distribution SOSD lower r
  - utility concave in degree implies efficiency ↑ r

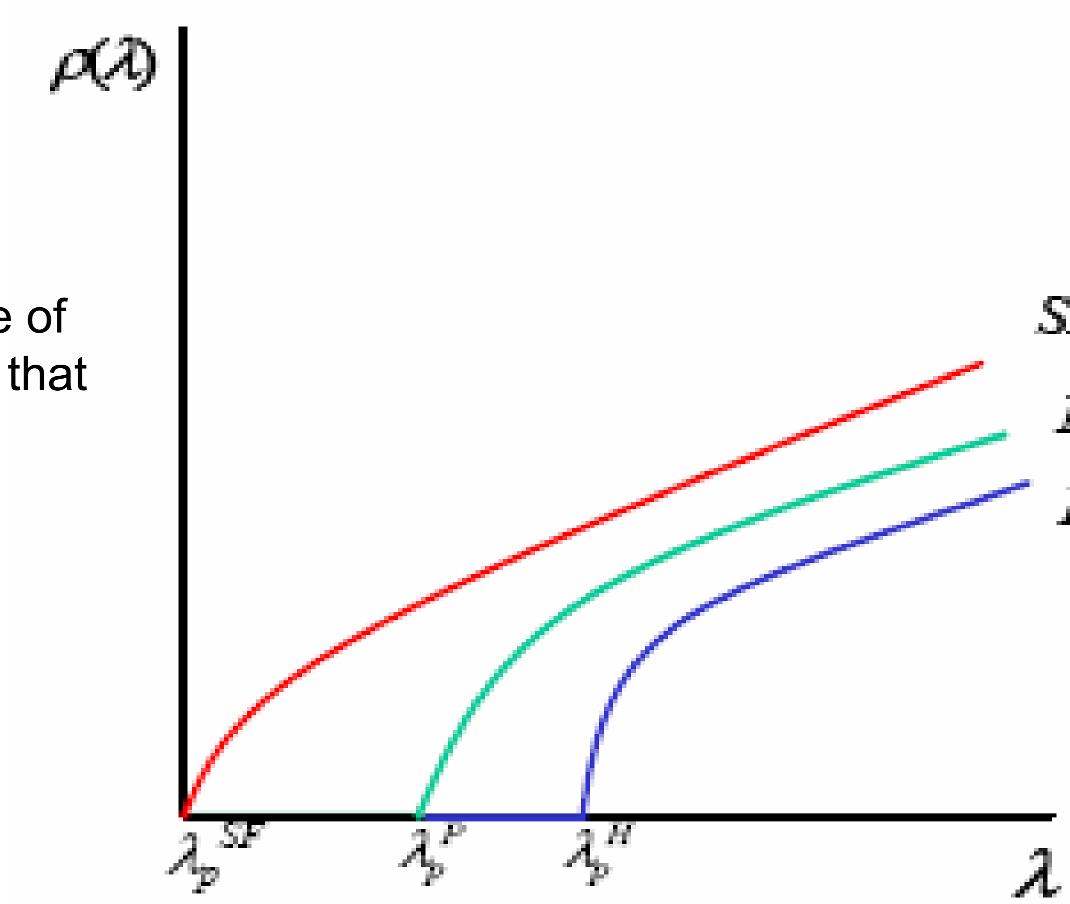
SIS Model (Bailey (1975))

- Nodes are infected or susceptible
- Probability that get infected is proportional to number of infected neighbors with rate ν
- get well randomly in any period at rate δ

Lopez-Pintado - infection rates

(Relates to lower r)

percentage of population that is infected



SF Scale Free

P Poisson (random)

H Homogeneous (regular)

infection rate/recovery rate

Infection rates related to Network structure

Proposition: For any r' > r there exist λ and λ' such that

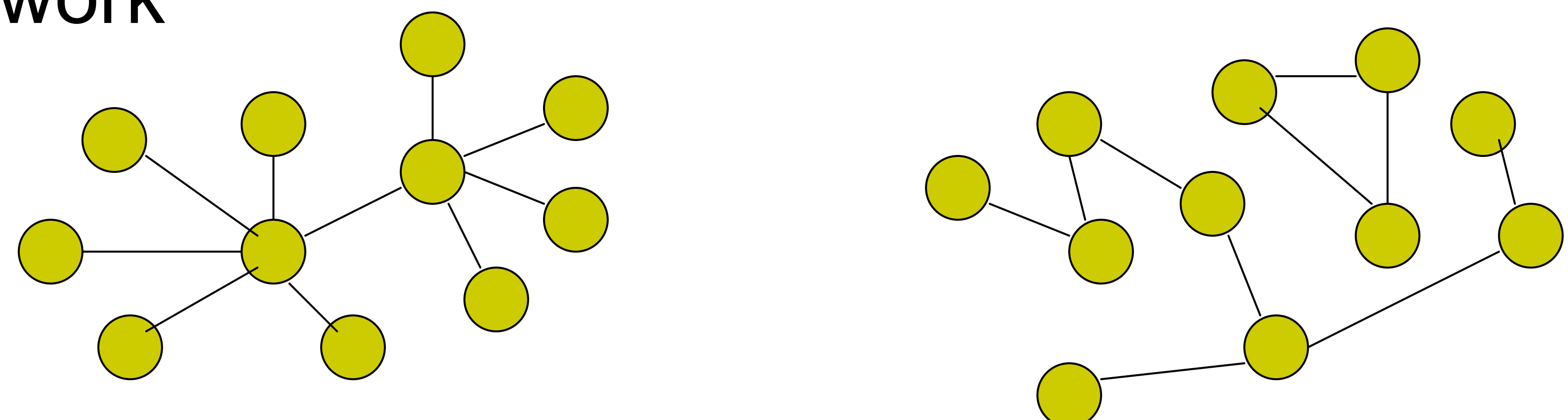
- If v/ δ<λ then the steady-state average infection rate is lower under r' than r.
- If v/ δ>λ' then the steady-state average infection rate is higher under r' than r.

Whither now?

- Bridging random/mechanical – economic/strategic
- Networks in Applications
  - Diffusion of information, technology– relate to network structure
  - Labor, mobility, voting, trade, collaboration, crime, www, ...
- Empirical/Experimental
  - case studies lack economic variables, tie networks to outcomes,
  - enrich modeling of social interactions from a structural perspective
- Furthering game theoretic modeling, and random modeling
- Foundations and Tools– centrality, power, allocation rules, community structures, ...

Connection to Information?

- Less random is more a like a ``hub and spoke'' network



- applications: infectious diseases, computer viruses, job information and employment, consumer behavior, social mobility...