Networks and Economic Behavior

Matthew O. Jackson

October 1, 2008
To Appear: *Annual Review of Economics, Vol. 1, pp. 489-513, 2009*

Abstract: Recent analyses of social networks, both empirical and theoretical, are discussed, with a focus on how social networks influence economic behavior, as well as how social networks form. Some challenges of such research are discussed as are some of the important considerations for the literature going forward.

Keywords: Networks, Social Networks, Economic Networks

Department of Economics,
Stanford University, Stanford, CA 94305-6072, USA
(jacksonm@stanford.edu)
http://www.stanford.edu/~jacksonm
and
the Santa Fe Institute.

Financial support under NSF grant SES-0647867 is gratefully acknowledged.

1. Introduction

Social networking web sites such as Facebook, MySpace and LinkedIn have made social networks more prominent than ever. Even though these and the many other forms of networked communications that have emerged with increased computerization provide a wealth of rich data for analysis, social networks are not new to economic interaction, nor are they new to researchers. These prominent examples simply make more broadly evident the importance of research on networked interactions and the opportunities for such scientific inquiry.

To some extent, the increased attention to social structure by economists parallels the growth of behavioral economics. Some of the interest in behavioral economics stems from the realization that psychological factors and context can be important determinants of decision making and ultimately of economic behavior. Similarly, the interest in social networks and the interaction patterns underlying economic activity stems from the realization that social context is an important determinant of economic behavior. The realization that social factors are critical to understanding a great deal of economic behavior is not new. The social embeddedness of economic activity was evident in sociological analyses of a wide variety of economic behavior even before Granovetter's important (1985) article popularized the concept of the embeddedness of economic activity. For instance, social structure is central in the study of the role of word of mouth communication in purchasing decisions by Katz and Lazarsfeld (1955), as well as in Myers and Schultz's (1951) study of the role of networks in the spread of job information, and in the study of exchange networks by Cook and Emerson (1978), and also in a string of analyses of social capital emerging with writings of Bourdieu (1972), Loury (1977), and Coleman (1988), among others. The recent growth of attention to the social context of economic activity comes in part from the maturing of the neoclassical models to the point where it is clear that they beg for additional context in order to explain a variety of observed phenomena, such as some patterns of wages and employment discussed below. It also comes in part from the development of the toolbox of researchers in economics to include a variety of methods that are well-suited to analyzing social interaction.

In line with this perspective, my discussion here focuses on (1) how examining the social context in which economic activity is embedded can enhance our understanding of economic behavior, both empirically and theoretically, and (2) which methods have been used in network analysis, how economic perspectives complement other perspectives, and what some of the basic hurdles are in conducting network analyses. In keeping with the format of the Annual Review, I do not attempt to survey the vast and growing literature on social and economic networks, but rather I provide a critical view of the state of the literature with an eye towards issues just mentioned. Given the space limitation, I will also keep the discussion at a non-technical level and so readers wanting a broader view or a more detailed look at specific subjects are referred to Jackson (2008).[1]

2. Empirical Analysis of the Impact of Networks on Economic Activity

As alluded to above, from an economist's perspective, there are two primary reasons for analyzing social networks.[2] One reason is that the methodology that has evolved in economic research is also very useful in modeling and analyzing social interactions. For example, new insights are obtained from bringing game theoretic reasoning to study network formation as well as the interactions between networked agents, as discussed in more detail below. A second reason that an economist should be interested in network analysis is that many economic interactions are embedded in networks of relationships and the structure of the network plays an important role in governing the outcome. For instance, many labor markets are decentralized with substantial flows of information about jobs being communicated via personal connections. This affects patterns of wages and employment as well as education decisions and social mobility. Understanding the impact of the network in determining information flows is essential to understanding some aspects of labor markets. These two reasons sit on different footings. The first is a reason why some economists should dedicate time to the study of social networks. The second is a reason why all economists should pay attention to the analysis of social networks. While social network analysis by researchers trained in economics and related fields will undoubtedly continue for the first reason alone, the continued growth of network analysis within the field of economics will ultimately rest on the showing that incorporating social context significantly enhances our understanding of an array of economic activities. There are already a number of important examples of economic applications where it is clear that social networks play a central role, as I now discuss.

---

[1] Jackson (2008) synthesizes the analyses of networks from sociology, economics, statistical physics, mathematics, and computer science. There are also various texts that focus on specific literatures. Wasserman & Faust (1992) present the tools of social network analysis stemming from the sociology literature. Bollobas (2001) surveys random graph theory. Vega-Redondo (2007) focuses on the analysis of complex networks based on some of the random graph techniques from mathematics, statistical physics and computer science. Goyal (2008) provides a look at some of the recent analyses of networks from the economics literature. There are also many helpful collected volumes that include some analyses of networks, such as Dutta & Jackson (2003), Demange & Wooders (2005), Newman, Barabasi & Watts (2006), Rauch (2007), and The Handbook of Social Economics (forthcoming). There are some popular texts such as Watts (2003) and Barabasi (2004), as well as a history of thought of the sociology literature by Freeman (2006) .

[2] Clearly, social scientists should be interested in social networks beyond the draw on an economist's toolbox and beyond network implications in economic applications, as there is also an interest in the pure science of social interactions.

2.1 The Relevance and Economic Implications of Networks

Despite the fact that we idealize markets as centralized and effectively anonymous institutions, many, if not most, markets function in a decentralized fashion, involving networks of bilateral interactions. Many markets involve networks, not only in terms of who transacts with whom, but also in the transmission of information about potential transactions. For example, one of the most extensively studied interfaces between social structure and markets is that of labor markets because of the large role that social networks play in disseminating information about both job openings and candidates. For example, Myers and Shultz (1951) interviewed textile workers in a New England mill town and asked how they had heard about their jobs. Myers and Shultz found that 62 percent had found out about and applied to their first job through a social contact, in contrast with only 23 percent who applied by direct application, and the remaining 15 percent who found their job through an agency, ads, or other means. There have been many studies since then that have examined the role of networks in communicating job information in different professions (e.g., Rees and Shultz (1970)), geographic areas and countries (e.g., Pellizzari (2004)), as well as comparatively across ethnicities, gender, and other dimensions (e.g., Corcoran, Datcher, and Duncan (1980)).[3] An influential study by Granovetter (1973), based on interviews in Amherst Massachusetts, noted not only the importance of social networks in obtaining information about jobs, but also demonstrated an important role for what he termed ``weak ties." That is, distinguishing social relationships that are ``strong" in terms of frequency or intensity of interaction from those that are more casual or infrequent and deemed ``weak," Granovetter found that a significant percentage (more than one fourth) of the job information obtained through social channels was obtained through weak ties.[4]

The fact that social networks are important in transmitting information about job openings and about potential employees will not come as a surprise, especially to anyone who has been on either side of a job market (even an academic job market). The impact of this fact ultimately depends on how it affects the functioning of job markets, and such questions have also been examined, for instance by analyzing how wage and employment patterns relate to social structure. For example, Bayer, Ross and Topa (2005) make use of census data to demonstrate higher correlation in employment among people living in the same block compared to correlations among those living on different blocks but still relatively nearby and with similar characteristics. They also find evidence suggesting that referrals can significantly influence wages and employment.

---

[3] See Ioannides and Datcher-Loury (2004) for more background.

[4] In quantifying the tie strength in the data, Granovetter defined tie strength based on the frequency of interaction between two individuals in the previous year. ``Strong" indicates interaction at least two interactions per week on average, ``medium" indicates less than two interactions per week but more than one per year, and ``weak" indicates on interaction per year or less. He analysis is based on a sample of 54 interviewees who found their most recent job through a social contact; finding that 16.7 percent had found their job through a strong tie, 55.7 percent through a medium tie, and 27.6 percent through a weak tie.

A challenge in such studies is that the social context is endogenous and so it can be difficult to be sure that the social context is responsible for the effect. In order to deal with this, beyond carefully controlling for all discernable characteristic of individuals as in the above study, another approach is to look for some mechanism that exogenously and randomly affects social interaction patterns and then seeing how that variation translates into variation in labor market outcomes. For instance, Laschever (2007) examines the random grouping of troops into military units in the United States World War I draft and finds that a ten percent increase in the average employment rate of a veteran's unit increases the veteran's employment rate by around three percent in expectation after correcting for other observables. Other examples of such techniques are Munshi (2003), who finds significant impacts from exogenous immigration patterns due to weather events, and Beaman (2007) who examines the random relocation of political refugees and sees significant differences in labor market outcomes based on the social setting that the refugees encounter.

These studies provide important evidence that social context influences wage and employment outcomes, and also provide some insight into the direction and magnitude of such effects. Beyond these empirical studies, there are also some applied theoretical analyses that bring job contact networks into models of employment and wages and show that this incorporation can provide new insight into a number of well-documented patterns in wages and employment. In particular, Calvó-Armengol and Jackson (2004, 2007, 2009) examine how the explicit transfer of job information through a social network impacts employment and wage patterns. Their model is such that if a worker is unemployed, then he or she can either receive information about a job opening either directly or via one of his or her employed friends. Increasing the number of employed friends that a worker has leads to increased employment prospects the worker. This also extends to wages, as hearing about more jobs (whether a worker is already employed or unemployed) leads to better matches and to more offers of employment and leads to higher wages. Thus, a worker's employment prospects and expected wages increase with the employment status of his or her acquaintances, and so such a network-based model of job information exhibits positive correlation in the employment and wages of connected workers. It also exhibits positive correlation of indirectly connected agents' employment and wages, and is such that the correlation between workers' status decays with their social distance.[5] Beyond such correlations, the model also exhibits duration dependence: the longer a worker is unemployed, the lower the probability that the worker will become employed in the next period. This is due to the fact that the longer a worker is unemployed the more likely it is that a worker's friends are also unemployed, which leads to a lower expectation that the worker will find a job in the future. These sorts of correlation patterns and time series can also translate into sustained differences in wages across different groups, such as by ethnicity, gender, or age, to the extent that

[5] Although this is intuitive on one level, there is a confounding effect that a friend of a friend is a potential competitor for information about a given job. That is suppose i and k are both linked to j and j hears about a job. Conditional on k changing from being unemployed to being employed, there is a lower probability that i makes the same change, as they are competitors for the job information from j. This short-term effect is of a second order compared to the fact that k helps keep j employed over time, which in turn helps keep i employed.

friendships are based on such factors. As the portion of workers of a given ethnicity who are unemployed is increased, the prospects for their acquaintances to hear about job information declines. So, if workers tend to have a disproportionate share of their social ties within their own group based on race, age, gender, or other attribute, then we will see employment and wage outcomes be correlated within groups even after adjusting for any other relevant characteristics of the workers. This, in turn, has implications for agents' decisions to become educated or to make any investments that affect their employability. If an agent expects that his or her friends are unlikely to become educated, then this lowers the agent's future job prospects and can lead the agent to under-invest in education and other sorts of human capital. Thus, as we consider the passing of job information explicitly through a social network, we can find network-based poverty traps and also decreased social mobility.

Another setting where social networks have been shown to influence economic decisions is in criminal behavior. For example, Reiss (1988) found that two thirds of criminals commit crimes with other criminals. Since such studies face problems of endogeneity, it can also be helpful to look at a model in estimating the effects. In this direction, Glaeser, Sacerdote and Scheinkman (1996) estimated a simple model of social influence, where criminal activity increases with neighbors' criminal activity, and found that petty crime and the tendency of youths to participate in crime were significantly influenced by their peers. Recently, richer models have been developed that allow for more complex network structures than the simple lattices in the Glaeser, Sacerdote, and Scheinkman model. Ballester, Calvó-Armengol, and Zenou (2006) develop such a model, where there are local complementarities in activities and global substitution effects. That is, the benefits from engaging in criminal activity increases with the criminal activity of one's friends, due to learning effects as well as production synergies from committing crimes together, but benefits decrease with overall competition from economy-wide criminal activity. Using a linear-quadratic payoff specification of how the payoff from criminal (or other) activity depends on the activity of one's neighbors, Ballester, Calvó-Armengol, and Zenou show an elegant and intuitive connection between the level of activity of a given agent and how central that agent is in the network (as defined via an eigenvector-based definition of network centrality due to Bonacich (1972, 1987)). The intuition is that a central agent is connected to other agents who are well-connected and so forth. Better connections lead to higher complementarities with neighbors and so higher marginal incentives to engage in crime, and this feeds back through the network in a way that is proportional to the centrality of the agents. This also allows Ballester, Calvó-Armengol, and Zenou to identify the ``key player'', that is the agent who if removed would lead to the largest change in the criminal activity. Although highly stylized, such a model allows tractable comparative statics to be derived that relate economic behavior directly and intuitively to network structure.

The examples of labor markets and criminal activity are two settings where there is empirical work relating social networks to behavior, and also where models are emerging that help provide additional predictions relating network structure to economic decisions and outcomes. There are many other such settings, that I will not detail here, but just to mention a few, they include studies of risk-sharing among networks of individuals (e.g., Fafchamps and Lund (2004), De Weerdt (2004), De Weerdt and Dercon (2006), Bloch, Genicot and Ray (2005), and Bramoullé and Kranton (2005)); networks of research and development, patent, and other joint ventures among firms (see the survey by Bloch (2004)); as well as more every-day activities such as how smoking and obesity are affected by friends' behaviors (e.g., Christakis and Fowler (2007, 2008)).

The applications above are ones where we see social networks influencing economic behavior. There are also empirical observations about social networks that economic models can help illuminate. An example is ``homophily,'' which is the tendency of nodes to be attached to other nodes that have similar characteristics. The background on this subject is rich beginning with Katz and Lazarsfeld (1954), and including important work by Blau (1977) and Marsden (1987, 1988), among others (see McPherson Smith-Lovin and Cook (2001) for a survey), and strong tendencies of individuals to associate with others with similar attributes have been widely documented, whether across age, race, gender, profession, religion or other dimensions. To get some impression of this, consider the following networks which are networks of friendships from the ``Add Health'' data set.[6] The network below consists of 624 nodes, which are the students in a US high school. A link indicates that at least one of the two students claimed the other as a friend in an interview. The figures are drawn using an algorithm that places linked nodes as close together as possible, while maintaining some overall average distance between nodes. Thus, groups of nodes with higher densities of connections are grouped together. Figure 1 is coded by the (self-reported) race of the students.

Figure 1 here

Figure 1 exhibits homophily in that we can see that nodes tend to be grouped by race. To see the phenomenon more closely, consider Table 1 below. We see that whites comprise 55 percent of the student body and yet 75 percent of their friendships (as an average per capita) are with other whites. If there was no homophily this percentage should be closer to 55. The difference in the percentage of own-race friendships from the expected number with no homophily is statistically significant (with a p-value near 0). Similarly, the Asian students also exhibit significantly higher percentages of friendships with other Asians than would be expected with no homophily: they comprise only 32 percent of the population but have 65 percent of their friendships with other Asians. Hispanics and Blacks comprise much smaller portions of the population of this high school and their percentages of own-type friendship do not exhibit homophily.

[6] Add Health is a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (addhealth@unc.edu). I thank James Moody for making available the data organized in Pajek files from which I derived these figures.

Table 1: Friendships by race in an Add Health high school

Race	% of the Population	% of friendships with own race
White	55	75
Asian	32	65
Hispanic	6	5
Black	1	1
Other/Unknown	6	-

We can also examine the same high school with respect to grade (i.e., year in school) rather than race. There we see even stronger patterns of inbreeding, as pictured in Figure 2.

Figure 2 here

The strong homophily patterns by grade are not so surprising, given that students will interact much more frequently with other students in the same grade.

This is an example of a setting where modeling network formation explicitly can help shed light on the sources and patterns of homophily. Currarini, Jackson and Pin (2009, 2010) document two empirical observations in the Add Health data related to homophily. First, they note that if a given ethnic group comprises a larger fraction of a high school's population, then the agents in that group tend to form more friendships on average: a group that comprises nearly 100 percent of a school forms over 8 friendships per capita, while a groups that comprises close to 0 percent of a school forms less than 5 friendships per capita. Second, Currarini, Jackson and Pin show that the extent to which a group inbreeds is nonlinear in the group's size (as a fraction of their school). Groups that comprise a middle-sized fraction of a school exhibit the highest level of inbreeding, even when normalizing by their relative size. With these observations about homophily in hand, Currarini, Jackson and Pin then examine a economic-style model of network formation, where there are two main influences on an individual's mixture of friends. Individuals meet potential friends via a random matching process and the decisions affecting which friendships are formed are based on an agent's preferences for various possible combinations of types of friends. Within that model, a bias in the meeting process, so that one is more likely to meet own types, leads to the pattern where middle-sized groups end up the most biased towards their own type in their friendships; but such a meeting bias does not lead to differences in numbers of friendships formed. In contrast, a same-type bias in preferences leads to an increase in the number of friendships formed as a group's size grows, but a preference bias does not lead to the right inbreeding patterns as a function of group size. Thus, within the model each of the two biases is needed to explain one of the empirical observations and neither can explain both of the empirical observations. This is certainly not the first (or last) word on what might lead to homophily patterns (again, see McPherson, Smith-Lovin and Cook (2001) for more background) and the full explanation probably involves a combination and can also differ across applications. Nevertheless such an analysis shows how an economic approach based on a simple matching model and choices by agents can help provide a new angle on a network phenomenon.

2.2 The Challenges of Endogeneity and Correlated Unobservables

Before moving on to discuss some of the theoretical modeling of networked behavior and network formation, let me discuss a few of the challenges that make empirical work on networked interactions difficult. As an example, suppose that we wish to determine whether a person's decision to buy a certain product or adopt a new technology is influenced by his or her friends and acquaintances. Even with detailed data on the network of social interactions and also on the behavior in question, this can be difficult to sort out. The difficulty is related to the homophily mentioned above. People associate with others who have similar characteristics. Some of those characteristics might not be observed by the researcher. If we see that individuals are more likely to adopt a technology conditional on their neighbors adopting it after conditioning on all of the factors that we have observed, we cannot be certain of whether there really is a social interaction which affects the decision to adopt the technology, or whether there is still some hidden characteristic which is correlated across the friendships and is responsible for the adoption decision. As a simple illustration, suppose that we are considering adoption of a new textbook among university professors. Suppose also that the publisher has advertized the text via mailings to some professional associations but not to others. If links among professors occur with a higher frequency between those in the same professional association than between professors who do not have professional associations in common (a form of homophily), we might mistakenly attribute a correlation of decisions among linked professors to be due to the social link rather than to the unobserved advertizing patterns. In this example, this effect could be detected by keeping track of professors' professional association affiliations. Even if the social interaction is a primary driver of behavior, convincingly establishing this involves ruling out other drivers of behavior, many of which we might not observe or directly control for.

An example of a study that is confronted with this issue is Uzzi's (1996) influential research on the garment industry in New York City in 1991. One of the things that Uzzi examines is how the rate of bankruptcy differs across firms, and in particular how it relates to the interaction patterns of firms. Uzzi keeps track of the extent to which a firm does repeat business with other firms or to which it interacts with many different firms in more one-time transactions. He develops an index which looks like a sort of Hirfindahl index for each firm: the squared fraction of the business that a firm does with each potential partner is summed across potential partners. A firm that does all of its business with one other firm has a score of 1, while a firm that spreads its business evenly among four different firms will have a score of 4/16=1/4. Uzzi then regresses whether or not a firm survived the year (125 out of the 496 firms went bankrupt during the year) on this measure, and some other variables. He finds that an increase in a firm's concentration index leads to a significant increase in its expected survival rate. Uzzi argues that this is reflective of how more solid social connections can enhance business and that longer-term, repeated, and embedded relationships can help overcome some frictions in contracting and can thus lead to increased profitability. He also bolsters this with data from interviews with various people in the industry. While these arguments are quite reasonable and provide insight into how strong social relationships can overcome an inability to write or enforce complete contracts; it is still hard to draw causal conclusions from the empirical data. There are many hypotheses for why firms might differ in their patterns of repeat business with other firms, such as (even slight) differences in specialization, and some of these could be related to bankruptcy probabilities.

This problem of unobserved correlated characteristics is clearly not special to the empirical analysis of social effects, as it is a challenge in analyses of many economic variables. Nonetheless, it is particularly acute in social settings because of the strong homophily patterns. This makes it very difficult to prove the obvious. That is, even in settings where we might very reasonably expect social interaction to be a primary driver of behavior, it can be difficult to convincingly establish this. There are various approaches to dealing with such things, such as having appropriate exogenously generated variation in the independent variables, such as in the studies by Munshi (2003), Beaman (2007), and Laschever (2007) mentioned above. One can also use instrumental variables approaches, or else take advantage of timing. A nice example of using timing is a study by Conley and Udry (2004) who conduct a careful analysis of the timing of changes in the use of fertilizer among pineapple growers and their neighbors in a social network to show how pineapple growers' fertilizer use is significantly driven by their observations of their neighbors' experiences.

A second issue that is particularly acute in network analysis is endogeneity. Do people adjust their behavior in response to that of their friends, or do they choose their friends based on behavior? Observing a correlation between social proximity and behavior does not imply any causation. Just as an illustration of endogeneity issues, note that another possible explanation for the Uzzi (1996) data discussed above is that it could be that firms can sense when another firm is weak and near bankruptcy. If firms are then unwilling to invest in repeated relationships with such weak firms, we would see firms near bankruptcy endogenously forced to have low indices and lots of smaller one-time transactions. Even though this might not be the right explanation for the data, the fact that patterns of relationships are chosen and possibly affected by factors that lead to bankruptcy makes interpreting the correlation difficult. This is true in many, if not most, applications where one tries to estimate how social structure impacts some outcome. Social structure is generally endogenous and could be influenced by the dependent variable or by some of the other factors that drive the dependent variable. Sorting out causation in such settings requires careful attention to timing, or else some powerful instruments or other clever approaches. Indeed, studies that have looked at time series of the co-evolution of behavior and social network structure, such as that by Kandel (1977), find that people adjust their friendships based on the behavior of their friends, and also adjust their behavior in response to that of their friends.

Beyond these twin challenges of unobserved correlates and endogeneity in working with social structure as an explanatory variable, there are also challenges dealing with specification and identification. For example, Manski (1993) points out identification problems that arise with quite natural specifications of peer influence (e.g., where an individual's behavior is influenced by the mean of his or her peers' behaviors in a linear fashion). Such issues can partly be overcome with more complete observation of the friendship patterns in a society, so that a given individual's peers can be directly observed and need not be inferred from the individual's own characteristics (what Manski refers to as the ``reflection problem''). But identification adds another layer of difficulty in empirical analysis of social interactions. All in all, analysis of social effects present a healthy set of challenges that provide a rich agenda both in further empirical studies and in developing new methods of analysis.

3. The Theory of Networks

Much of the advance made by economists in the study of networks has come in developing theory about how networks form as well as how networks influence behavior. These theoretical advances make use of economic modeling techniques and help provide new insight into the structure and implications of social networks. Let me discuss each of these in turn.

3.1 Network Formation

The welfare implications of any interaction are central to an economic analysis, but not always so central to other disciplines. In part, this stems from the utility maximization/revealed preference perspective that is a foundation of the modern economic paradigm. Although this paradigm is sometimes constraining, it also provides for a powerful welfare analysis and a deep understanding of externalities. The link between economic insight and studying network formation comes from the fact that externalities play a prominent role in many network settings. For example, how well my friends are connected is important in determining what job information I have access to, what I learn from them, and more generally how I benefit from all sorts of interaction with them. As maintaining relationships involves some discretion, bringing a strategic perspective to network formation has provided several insights.

A useful illustration of these points is a simple model that presents a template of network formation by self-interested individuals and includes explicit benefits from maintaining links to well-connected individuals. This is the ``connections model'' introduced by Jackson and Wolinsky (1996). In that model the payoff to agent i in a network g is

$$u_i(g) = \sum_{j \neq i} \delta^{\ell(ij)} - d_i\, c,$$

where $\ell(ij)$ is the number of links in the shortest path between i and j in g (setting $\ell(ij)=\infty$ if there is no path between i and j), d_i is i's degree (the number of links that i maintains in g), and c is a parameter representing the cost of a link. So, individuals get benefits from having links both through the direct connections that links provide as well as the indirect connections links provide. Having a friend is worth δ, each friend of a friend is worth δ^2, and a friend of a friend of a friend is worth δ^3, and so forth, while each direct friend also costs some amount c.

The strategic formation of a network is modeled via a simple stability/equilibrium concept that incorporates the idea that mutual consent is needed to form a relationship,[7] but an individual can unilaterally sever a relationship. A network is pairwise stable if[8]

- whenever a link between two individuals is absent from the network then it cannot be that both individuals would benefit from adding the link (with at least one benefiting strictly), and
- whenever a link between two individuals is present in the network then it cannot be that either individual would strictly benefit from deleting that link.

The connections model is fairly easy to analyze with this tool in hand. With low enough costs, so that the cost of adding a link is less than the marginal gain from converting a friend of a friend into a friend, then all links will form and only the complete network will be pairwise stable. If costs are prohibitively high, then no links will form. In an intermediate range of costs there are a variety of network structures that are pairwise stable, depending on the particulars of the number of agents, the linking cost c and the decay parameter δ. The interesting aspect is to contrast these pairwise stable networks with the network that maximizes the total societal welfare. If costs are extremely low or high, then the total utility maximizing network is the unique pairwise stable network. When costs are intermediate, then the pairwise stable networks might not include any of the total utility maximizing networks, or might include other networks. A key to understanding this is that the unique total utility maximizing network architecture for intermediate cost ranges is a star: a network where one agent is linked to all others, and the others are only linked to this center agent. This network involves the minimum number of possible links and connects all agents at a distance of at most two from each other. In fact, this is the total utility maximizing network in a much wider variety of models (see Jackson (2008) for more discussion). However, it is easy to see that stars will often not be pairwise stable, even when they maximize total utility. The key is that the peripheral agents in the star benefit from all of the indirect connections that the center agent provides them, and yet the center agent bears most of the cost. For some cost ranges (e.g., when c exceeds δ) the only pairwise stable networks will be such that all non-isolated agents have at least two links as otherwise no agent would be willing to maintain a link with them.

---

[7] There are some applications where links can be formed unilaterally, such as when one author cites another. Directed networks provide some variations on the above results (e.g., see Bala and Goyal (2000) and Dutta and Jackson (2000)). More generally, there are many issues regarding how to model strategic network formation. For more discussion, see Jackson (2004, 2008).

[8] More formally, let $g+ij$ denote the network formed when the link ij is added to the network g and g-ij denote the network formed when the link ij is deleted from the network g, where a network is represented as a list of all the pairs of nodes that are linked. A network g is pairwise stable if: (i) if there is an i and ij not in g such that $u_i(g+ij)>u_i(g)$ then $u_j(g+ij)< u_j(g)$; and (ii) for all ij in g and i $u_i(g)\geq u_i(g-ij)$. For more detailed and formal definitions, see Jackson (2008).

The above analysis of the network formation is a starting point, and there are many other considerations that are also important to include. For example, it could be that the agents can offer payments to each other as a function of the network that forms (e.g., see Currarini and Morelli (2000), Bloch and Jackson (2006)) or it could be that there is some general bargaining procedure that determines how the total benefits of the network are divided (e.g., see Myerson (1977), Jackson and Wolinsky (1996), Slikker and van den Nouweland (2001) and Jackson (2005)). Interestingly, even with complete information and a very wide class of possible ways in which the total utility in a network can be reallocated, it can still be that no total utility maximizing network is pairwise stable, as shown by Jackson and Wolinsky (1996). In most economic settings inefficiencies arising from externalities can be rectified, in the absence of any frictions, through proper transfers; an idea which is known as the Coase (1960) Theorem. The multilateral and combinatorial nature of the externalities in the network case can sometimes preclude such efficiency, even without other bargaining frictions.

While game theoretic models of network formation provide novel insight into the patterns that might emerge and into the tension between individual incentives to maintain relationships and overall welfare, there are challenges in using game theory in working with data. Two such challenges are that there can be multiple equilibria even in relatively simple settings, and that such models can be difficult to solve when introducing natural sorts of heterogeneity among players. Nonetheless, such models can still be pushed quite far in these directions. For example, Carayol, Roux, and Yildizoglu (2006, 2008) use genetic algorithms and Monte Carlo simulations to solve large versions of the connections model where linking costs are based on a geographic distance and agents are located at different locations. This introduces some heterogeneity into the model that allows it to begin to exhibit the variation in connectedness and some of the spatial patterns that are observed in applications.

Beyond strategic network formation models there are also a variety of random network formation models that have been very useful in providing some basic insights regarding network structure and well as in statistical analysis. I briefly discuss some of these to give a feeling for the differences between the approaches and how they complement each other. The seminal papers in this strand of the literature are from the random-graphs literature in mathematics, including classics by Erdos and Renyi (1959, 1960, 1961). The canonical model that Erdos and Renyi studied (among others) is one where a network is formed by having each link form independently with a probability p. There are many things that are known about such networks, beginning with important early theorems proven by Erdos and Renyi. For instance, when looking at large networks (theorems are often about asymptotic properties as the number of nodes n grows), if the average degree, p(n-1) below one, then the network will generally consist of many small and separate components. Once the average degree exceeds one, then the network starts to coalesce and a ``giant component'' starts to emerge. Once the average degree exceeds log(n) then the network almost surely consists of just one component that includes all of the nodes.

Although this basic random graph model is too simple to match reality, some of its features are quite robust and offer some insight into observed network patterns. For example, in this random network model if we hold the average degree d does not grow too quickly as the number of nodes increase, but is above one, then as the number of nodes grows the average distance between the nodes in the giant component is proportional to log(n)/log(d). This offers some insight into Milgram's (1967) classic experimental study of the ``small worlds'' phenomenon, and its many follow ups.[9] This relatively short social distance between nodes in a uniformly random network is relatively easily understood, as with a constant degree, the network effectively looks like a tree. Just as a thought experiment, if we think about a tree network where every node has degree d, if we start at some node and follow all paths out t links, then we will have reached $d+d(d-1) + \ldots + d(d-1)^{t-1}$ nodes, which is on the order of d^t nodes. To reach all nodes, or to have d^t be on the order of n, we need t to be roughly log(n)/log(d). So, a node can reach any other node in its component via a path that is of an order of no more than log(n) links. Thus, the average distance in a randomly generated network is much smaller than the number of nodes in the network, just as empirically observed.

When we push a bit further, we find characteristics of networks that are not well matched by this basic model of uniformly random link formation. For example, such a model exhibits vanishing ``clustering'' unless the density of links becomes extremely high.[10] In contrast, many observed social networks exhibit substantial clustering. For instance, various networks of co-authorships among scientists show clustering on the order of hundreds or thousands of times what would arise if links were formed uniformly at random. Watts and Strogatz (1998) show a model that has some features of a regular lattice, and other features of random link formation can exhibit the characteristics of both. In particular, they show that beginning with a lattice structure among nodes, so that the starting network has high clustering, one needs only randomly change a small fraction of the links in order to dramatically reduce the average and maximum distance between nodes. Thus, starting with a highly structured network, adding a small amount of randomness can lead to a network that exhibits two common features of observed social networks: high clustering and low average path lengths between pairs of nodes.

Strategic models of network formation lead to very complementary explanations of these same phenomena of short average path length and high clustering. Specifically, versions of the connections model with geographic costs can also lead to small-world network

---

[9] Milgram's (1967) experiment involved picking some subjects in Kansas and Nebraska and asking them to direct letters to other people in Massachusetts, who were unknown to the people in Kansas and Nebraska other than in terms of a name and profession. The key was that the subjects could not send the letters to someone whom they did not know, but instead had to send the letter to someone they knew and ask that person to pass the letter along. Thus the letters had to follow the paths of social network. A surprising fraction (almost a quarter) of the letters reached their destination and did it in a median of 5 relays.

[10] Clustering refers to the tendency of linking to be a transitive relationship. That is, a measure of the clustering at a given node is to examine what fraction of pairs of that node's neighbors are connected to each other.

characteristics. That is, suppose that people have low costs of forming relationships to people who are close to them spatially, and higher costs of forming relationships to people who are far away from them. ``Close spatially'' could refer to being nearby geographically, or it could also mean having similar characteristics like age, profession, education, or other attributes. Having low costs of forming nearby links leads to dense networks on a local level, with substantial clustering: local friendships are likely to be transitive so that my friends are likely to be linked to each other. High costs of forming distant links means that there will be fewer such links. However, if there were no links between people who are spatially distant from each other, then such links would become extremely valuable. Long-distance links shorten path length to many indirect connections and so can be very valuable. Thus, as long as there are not too many long-distance links they can be very attractive and so the network will have many local links and high clustering among nodes that are close to each other, and then some longer distance links that ensure that average path lengths do not grow to be too large. Various forms of the spatial connections models are examined by Johnson and Gilles (2000), Carayol and Roux (2003), and Jackson and Rogers (2005).

Price (1965) and Albert, Jeong, and Barabasi (2000) note other interesting features of some networks that differ from a network where links are formed uniformly at random. They examine the frequency distribution of degrees across nodes, known as the degree distribution. If links are formed uniformly at random, then for large networks this distribution is roughly a Poisson distribution. However, some observed degree distributions have ``fat tails,'' in that they have a relatively high frequency of nodes with very high and very low degrees, and a relatively lower frequency of nodes with intermediate degrees. In particular, the observed degree distributions in the Price (1965) and Albert, Jeong, and Barabasi (2000) studies are closer to that of a power distribution where relative frequency of nodes of degree d is proportional to $d^{-\gamma}$ for some γ. Price (1976) and Barabasi and Albert (2001) offer random network formation models that yield such fat tailed distributions. The basic idea is that nodes are born over time and form new links as they come into the system. In particular, the probability with which they attach to a given node is proportional to the number of links that the node already has. This is termed ``preferential attachment'' by Barabasi and Albert, and leads nodes with more links to accumulate even more new links than nodes with fewer links, a sort of ``rich get richer'' phenomenon. This leads to a power degree distribution.

There are also models of random network formation that are hybrids, combining some uniformly random attachment with some other sorts of attachment, and that result in degree distributions that lie somewhere between that of uniformly randomly generated links and the fully fat tails of the preferential attachment process. These models are useful, especially in empirical analyses, since many observed degree distributions do not lie at either extreme, and so one can then estimate to what extent links are formed uniformly at random and to what extent the formation process is driven by the existing network structure. Such models can also exhibit the significant clustering and low average path lengths discussed above, as well as other features of observed networks (e.g., see Pennock et al (2002) and Jackson and Rogers (2005)).

There are far too many other aspects of network formation modeling to discuss in the limited space here.[11] But the above discussions and models provide a feeling for some of the approaches and issues that have been examined. One of the important remaining gaps in the literature is in somehow bridging between the process-based random graph formation models that are good at answering ``how'', and the strategic-based formation models that are good at answering ``why''.[12] These modeling techniques are quite complementary and developing models that combine some random opportunities for link formation coupled with some discretion in which relationships are actually chosen, could be quite useful resulting in models with the heterogeneity that is needed to fit data but also understanding some of the forces shaping network formation and allowing for a welfare analysis. The Currarini, Jackson and Pin (2009, 2010) model discussed above is one example that combines discretion with random meetings, but a richer paradigmatic approach is still missing.

3.2 Networks and Behavior

Another growing area of the modeling of networked interaction examines how network structures impact behavior. This embodies a series of questions about both how overall average behavior of a society is affected by the structure of its social network, as well as how individual behavior is affected by position in the network. Theoretical work on this topic is essential to the goal of understanding how network structure impacts economic outcomes.

In thinking about how network structure impacts behavior, it is useful to roughly partition settings into two categories. In one situation, communication, contagion or learning occurs through the network. Here network structure is primarily involved in transmission and determining flows of some information or behavior. One example of an application that falls into this category is learning about the value of some consumer product via word of mouth, and another is the contagion of a computer virus that is carried in email attachments. In a second situation, agents are making choices and their payoffs from those choices depend on the choices of their neighbors. Here, the network structure again is important, but it is not because of the flow or transmission of some information, but instead because it affects the patterns of interactions and thus the patterns of externalities that impact decisions. An example of this is deciding on which of several technologies to adopt, such as choosing a software package when compatibility of the software with acquaintances' software matters. Of course, there are situations that involve some aspects of both pure transmission and of local external effects in decisions. In the job contact networks discussed above both roles of networks were present at once:

---

[11] For instance, there is a whole other class of models that come out of the statistics and sociology literatures that were developed to work directly with data. A prominent class, known as p* models, is discussed by Wasserman and Faust (1998). These can allow for rich interactions in the probabilities that different combinations of links form. There are also classes of models that emerge from the random graphs literature that are useful in deriving large graph properties while incorporating varied degree sequences. Again, see Jackson (2008) for an overview.

[12] See Jackson (2005) for more discussion of this point.

information about jobs flowed through the network and also the benefits from investing in education by individuals depended on their neighbors' decisions.

The difference between these two roles of networks in affecting behavior is important because the tools needed to analyze one are different from the other. The transmission, diffusion, and contagion sort of role of networks is to some extent mechanical: it is process based and much of the mathematics of various dynamic systems (e.g., Markov chains, percolation theory, …) can be brought to bear on the analysis. The situation where individual decisions have external effects and the choices of neighbors are complements or substitutes requires some form of either equilibrium or agent-based modeling techniques to handle the strategic interactions. Such settings can still involve dynamic systems, but the external effects of decisions require some sort of analysis aimed at strategic interaction.

Both types of analysis face the hurdle that incorporating complex social networks can present roadblocks to tractability and so there is a delicate balance in keeping a model rich enough to study interesting interactions and dynamics, and yet simple enough to work with in the face of the daunting combinatorics that emerge in networked settings. Let me discuss a few of the areas that are active areas of research with regards to how social structure impacts behavior and some of the hurdles they face.

A starting point for understanding contagion and diffusion through networks is to examine the transmission of disease through a social network. In the simplest cases, this just involves understanding the network structure directly. For example, if a disease were completely virulent so that anyone linked to an infected individual were to become infected, then one could trace the course of a disease simply by examining the structure of the components of the network. If a network were path-connected then the result of such a disease would be catastrophic. Of course, most transmission is not so virulent. Nonetheless, starting from this simple case, one can build up a model to allow transmission to be probabilistic across links, for instance if people only have some chance of interacting or of transmitting a disease conditional on being linked. One can also add random times for which nodes are infected. In some contexts, like a common cold, individuals recover from being infected but then can later be re-infected, while in other contexts an infected individual recovers and cannot be re-infected and can no longer infect others. Some of the analysis of such models can be conducted using random graph models, where the extent of the infection is studied simply by altering the random graph model to adjust the probability that links and/or nodes are present. Models of such contagion are relatively tractable and provide results regarding how network structure affects both whether or not an infection can gain a toe hold as well as the extent to which it eventually spreads (e.g., see Pastor-Satorras and Vespignani (2000, 2001), Jackson and Rogers (2007), and Lopez-Pintado (2008)). There are some simple but important intuitions that emerge from these analyses, mainly regarding how the distribution of the degrees of nodes in a network affects contagion and eventual epidemic size: Very highly connected nodes are more susceptible to infection, holding all else constant, than less connected nodes, simply because they have more interactions and are more likely to meet infected nodes. This leads networks with fatter-tailed degree distributions to be more susceptible to the initial spreading of a disease. That is, if we hold the total number of links in a network constant, but rearrange them so that some nodes are more highly connected and others are less connected, effectively instituting a mean-preserving spread in the degree distribution, then we end up with a network that looks increasingly ``hub and spoke" like. The highly connected agents, or hubs, serve as conduits for infection and help infections to spread more easily. These agents are easily infected and also contact many other agents. On the other hand, the extent to which an infection eventually spreads can actually be lower in a network with fatter tails. The presence of hub agents makes it easier to get past the initial infection threshold, but there are also more agents with very low degrees, and such agents are relatively more difficult to infect.

Models of diffusion of a disease are nice examples of situations where one can fairly cleanly make predictions about outcomes working from some simple characteristics of network structure, at least in stylized random network models that capture some of the basic features of actual social networks. Note that such models also embody the spread of some sorts of information, rumors, or adoption of some technologies. To the extent that such behaviors are simply dependent on contact with other ``infected'' (i.e., informed agents), much of the analysis extends directly. The main departure from the above modeling comes once we enhance it to cover learning, when learning involves processing diverse information from multiple sources.

Modeling learning in networked settings has employed two basic approaches. One is a Bayesian approach, where agents update their beliefs based either on communication or observation of other agents' actions over time. This approach provides a nice benchmark for what happens with ``full rationality.'' Another approach is more mechanical where agents repeatedly process the information from their neighbors according to fixed rules.

To get some impression of the Bayesian approach, consider an example where each day agents are faced with a choice of two different actions say A and B, which have stochastic payoffs and the actions lead to similar average payoffs to the different agents, but such that which action has the higher payoff is initially unknown. For instance, suppose that the agents are fisherman and the A and B represent two different types of baits that can be used to try to catch fish in a certain area. The agents learn about which action is better through their own experience and also from observing what actions other fishermen take and how much fish the other fishermen bring home each day. Let who observes whom be described by a network. So each day a fisherman chooses either bait A or B and then goes out and fishes. When he or she comes home, she knows how many fish she caught, and also sees what bait her neighbors used as well as how many fish they caught. The Bayesian inference problem becomes complicated very quickly. If I see other fishermen change baits, it could in part be due to their experience, but it could also be due to what they have seen happen to their neighbors, or even because they have seen their neighbors change actions, and so forth. Even if I know the structure of the entire network, the updating problem quickly becomes intractable.[13] Nonetheless, there are

[13] See Gale and Kariv (2003) and Choi, Gale, and Kariv (2005, 2007) for more discussion and some experimental investigations of the extent to which agents are fully rational learners in networked settings.

still some important things that can be deduced. One is that agents in the same components of the network will, almost surely, converge to eventually getting the same long run payoff. If not, then there would be two neighboring agents somewhere in the network who would be getting different long-run payoffs. The agent getting the lower payoff should eventually realize this and change behaviors. This result is pointed out by Bala and Goyal (1998). Note that this does not mean that all agents eventually learn which is the long-run expected payoff maximizing action, but instead that they will all settle down to getting the same payoff. Additional conditions are needed in terms of diversity of initial beliefs, the observation patterns, and/or the network structure in order to get convergence to the correct action (e.g., see Bala and Goyal (2001) as well as Acemoglu, Dahleh, Lobel, and Ozdaglar (2008)).

The other different extreme in terms of modeling approach is well represented by a model by DeGroot (1974). It that model agents repeatedly communicate with each other, and repeatedly update by taking a weighted average of the opinions of themselves and their neighbors. The repetition of communication allows information to diffuse throughout the network. The process is not Bayesian in that the weights the agents use to average the signals that they get in a given period are not optimally adjusted over time, and as such this provides a boundedly rational benchmark. One advantage of the model is that the repeated weighted averaging of signals very tractable as it involves repeatedly multiplying an initial vector of beliefs by a weighted updating matrix. This is a simple linear algebra problem, and the long run beliefs, convergence properties, the relative influence of the agents, and a host of other things can be explicitly calculated for any given network (e.g., see DeMarzo, Vayanos and Zwiebel (2003), and Golub and Jackson (2008, 2010, 2011)). One can even examine things like how homophily influences the speed of learning. It is also not clear which of these extreme sorts of models of updating better matches reality. The Bayesian updating becomes too intractable for agents to undertake, and yet at the other extreme repeated myopic updating is perhaps too simplistic. But combined, the models complement each other well.

As mentioned above, one needs different techniques when examining networked interactions where agents are making decisions and their decisions are influenced by other agents. This is another important area of research because of the large number of applications. Many choices that we make on a daily basis, such as which phone plan to use, whether or not to smoke, whether to take up a sport, how much education to pursue, and so forth, are dependent on the actions of our friends, family, and other acquaintances. Modeling this coherently presents some substantial challenges because of the combinatorial nature of the problem. To get a feeling for this, consider a simple but natural variation of a network-based game considered by Bramoullé and Kranton (2007). Suppose that each of individual in a network decides whether or not to buy a tool. If agent i does not have the tool but one of i's neighbors does, then i can borrow the tool. Agent i cannot borrow the tool from the friend of a friend. If none of i's friends has the tool then it is worthwhile for i to buy it. If at least one of i's friends has the tool then it is strictly better not to buy the tool but simply borrow it instead (and there are no congestion problems with respect to borrowing tools). There are many equilibrium configurations to this game and they depend on the network structure. Pure strategy equilibria are configurations such that at least one person in each agent's neighborhood buys a tool, and no two linked agents both buy tools. There are generally many such configurations for any network. Moreover, slight changes in the network structure can lead to dramatic changes in the equilibrium configurations. Even so, there are still some properties of the equilibria of such games that can be established, and some comparative statics can be obtained in terms of how the equilibrium structure changes as the game changes. Interestingly, in some cases, if agents have to make their choices before they are sure of who their neighbors will be, then the problem can actually simplify dramatically and a variety of equilibrium properties and comparative statics can be more directly obtained (e.g., see Galeotti et al (2010)).

Again, the above discussion only scratches the surface of some of the things that have been investigated. Just to name a few applications, there is a growing literature on bilateral trade between buyers and sellers in networked markets and how terms of trade are influenced by, and influence, network structure. There are also analyses of risk-sharing, favor trading, advertising in networks, and also some modeling of the co-evolution of network structure and behavior.

3. Concluding Remarks

As should be clear by now, the literature on networks is rapidly growing. It is an exciting area because of its multi-disciplinary nature, and it is difficult to think of other areas of research that so naturally draw from, and apply to, as many disciplines. As should also be clear, there are many challenges that still lie ahead and a lot of wide open problems. Most notably, whether or not networks become an essential part of the economic paradigm will depend on the extent to which there are multiple settings where the network structure is a primary driver of economic behavior. There is a growing body of empirical research which suggests that social effects are substantial and that incorporating social context into economic studies will become increasingly necessary. We have also seen that both the empirical research and theoretical research face their own challenges in the area, related to issues of endogeneity and to the inherently complex and combinatorial nature of networked settings. These challenges, and the myriad of interesting network applications to explore, should keep researchers busy for some time.

References

D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar 2008 Bayesian learning in social networks, Mimeo, M.I.T.

Albert, R., H. Jeong, and A.L. Barabasi (1999) "Diameter of the World Wide Web," *Nature* 401:130–131.

Aumann, R. and Myerson, R. 1988 Endogenous Formation of Links Between Players and Coalitions: An Application of the Shapley Value, In: Roth, A. ed. *The Shapley Value*, Cambridge University Press, 175-191.

Bala, V., and S. Goyal (1998) “Learning from Neighbors,” *Review of Economic Studies* 65: 595–621.

Bala, V. and Goyal, S. 2000 A non-cooperative model of network formation, *Econometrica*, 68: 1181.1230.

Bala, V. and Goyal, S. (2001) “Conformism and Diversity under Social Learning,” *Economic Theory*
17:101–120.

Ballester, C., A. Calvó-Armengol, and Y. Zenou 2006 Who’s Who in Networks: Wanted the Key Player, *Econometrica*, 74(5): 1403-1417.

Barabási A. and R. Albert 1999, Emergence of scaling in random networks, *Science*, 286: 509-512.

Bayer, P.J., S.L. Ross, and G. Topa (2005) “Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes,” Economic Growth Center Discussion Paper 927, New Haven, Conn., Yale University.

Beaman, L. (2007) “Refugee Resettlement: The Role of Social Networks and Job Information Flows in the Labor Market,” New Haven: Conn., Yale University, manuscript.

Bloch, F. 2004 Group and Network Formation in Industrial Organization, Chapter 11 in Group Formation in *Economics; Networks, Clubs and Coalitions* , edited by Gabrielle Demange and Myrna Wooders, Cambridge University Press:
Cambridge U.K.

F. Bloch, G. Genicot, and D. Ray 2005 Social Networks and Informal Insurance, mimeo: GREQAM, Georgetown University and NYU.

Bloch, F. and M.O. Jackson 2007 The Formation of Networks with Transfers Among Players, *Journal of Economic Theory*,

Bonacich, P. 1972, Factoring and weighting approaches to status scores and clique identification, *Journal of Mathematical Sociology*, 2: 113.120.

Bonacich, P. (1987) “Power and Centrality: A Family of Measures,” *American Journal of Sociology* 92:1170–1182.

Boorman, S. 1975 A Combinatorial Optimization Model for Transmission of Job Information through Contact Networks, *Bell Journal of Economics*, 6: 216.
249.

Bramoullé, Y. and R. Kranton 2007 "A Model of Public Goods: Experimentation and Social Learning,"
Journal of Economic Theory 135:478–494.

Bramoullé, Y., and R. Kranton (2005) "Risk-Sharing Networks," *Journal of Economic Behavior and Organization* 64(3–4):275–294.

Burt, R. 1992 *Structural Holes: The Social Structure of Competition*, Harvard University Press: Cambridge, MA.

Calvó-Armengol, A. 2004 Job Contact Networks, *Journal of Economic Theory,* 115: 191-206.

Calvó-Armengol, A. and M.O. Jackson 2004 The Effects of Social Networks on Employment and Inequality,. *American Economic Review*, 94(3): 426-454.

Calvó-Armengol, A. and M.O. Jackson 2007 "Networks in Labor Markets:Wage and Employment Dynamics and Inequality," *Journal of Economic Theory* 132(1): 27–46.

Calvó-Armengol, A. and M.O. Jackson 2009 Like Father, Like Son: Labor Market Networks and Social Mobility, *American Economic Journal: Microeconomics,* Vol. 1, No. 1, 124-150.

Carayol, N. and P. Roux 2003 Collective Innovation in a Model of Network Formation with Preferential Meeting, mimeo: Université Louis Pasteur and Université de Toulouse I.

Carayol, N., P. Roux and M. Yildizoglu 2006 In search of efficient network structures: the needle in the haystack, *Review of Economic Design*, 11(4): `434-`470

Carayol, N., P. Roux and M. Yildizoglu 2008 **Inefficiencies in a model of spatial networks formation with positive externalities** ***Journal of Economic Behavior & Organization*** 67(2): 495-511

Charness, G. and M. Corominas-Bosch, and G.R. Frechette 2005 Bargaining on Networks: An Experiment, *Journal of Economic Theory*.

Choi, S., D. Gale, and S. Kariv (2005) "Behavioral Aspects of Learning in Social Networks: An Experimental Study," in *Advances in Applied Microeconomics*, Vol. 13, Behavioral and Experimental Economics, J. Morgan, ed., Berkeley, Calif.: BE Press.

Choi, S., D. Gale, and S. Kariv (2007) "Social Learning in Networks: A Quantal Response Equilibrium
Analysis of Experimental Data," Berkeley, University of California at Berkeley, preprint.

Christakis, N.A., and J.H. Fowler (2007) "The Spread of Obesity in a Large Social Network over 32 Years," *New England Journal of Medicine* 357:370–379.

Christakis, N.A., and J.H. Fowler 2008 **The Collective Dynamics of Smoking in a Large Social Network** *New England Journal of Medicine* **358(21): 2249-2258.**

Coase, R.H. 1960 The Problem of Social Cost *The Journal of Law and Economics*, 3: 1-44.

Conley, T.G., and C.R. Udry (2004) "Learning About a New Technology: Pineapple in Ghana," Economic Growth Center Working Paper 817, New Haven, Conn., Yale University.

Corcoran, M., L. Datcher, and G. Duncan (1980) "Information and Influence Networks in Labor Markets," in *Five Thousand American Families,*volume 8, G. Duncan and J. Morgan, eds., Ann Arbor: University of Michigan.

Corominas-Bosch, M. 2004 On Two-Sided Network Markets, Journal of Economic Theory, 115: 35-77.

Currarini, S., M.O. Jackson, and P. Pin (2009) "An Economic Model of Friendship: Homophily, Minorities and Segregation," *Econometrica,* Vol. 77, No. 4, 1003–1045.

Currarini, S., M.O. Jackson, and P. Pin (2010) ``Identifying the roles of race-based choice and chance in high school friendship network formation,'' *the Proceedings of the National Academy of Sciences.* 107(11): 4857–486.

Currarini, S. and Morelli, M. 2000 .Network Formation with Sequential Demands, *Review of Economic Design*, 5: 229.250.

DeGroot, M.H. (1974) "Reaching a Consensus," *Journal of the American StatisticalAssociation* 69:118–121.

DeMarzo, P., D. Vayanos and J. Zwiebel 2003 Persuasion Bias, Social In.u-ence, and Unidimensional Opinions, *Quarterly Journal of Economics*, 1183: 909 - 968.

De Weerdt, J. (2004) "Risk-Sharing and Endogenous Network Formation," in *Insurance Against Poverty,* S. Dercon, ed., Oxford: Oxford University Press.

DeWeerdt, J., and S. Dercon (2006) "Risk-Sharing Networks and Insurance against Illness," *Journal of Development Economics* 81:337–356.

Dutta, B., S. Ghosal, and D. Ray 2005 Farsighted Network Formation, *Journal of Economic Theory*, 122: 143-164.

Dutta, B., and Mutuswami, S. 1997 Stable Networks, *Journal of Economic Theory*, 76: 322.344.

Economides, N. 1996 The Economics of Networks, *International Journal of Industrial Organization*, 16(4): 673-699.

Erdos, P., and A. Renyi (1959) “On Random Graphs,” *Publicationes Mathematicae Debrecen* 6:290–297.

Erdos, P., and A. Renyi (1960) “On the Evolution of Random Graphs,” *Publication of the Mathematical Institute of the Hungarian Academy of Sciences* 5:17–61.

Erdos, P., and A. Renyi (1961) “On the Strength of Connectedness of a Random Graph,” *Acta Mathamatica Academy of Sciences of Hungarica* 12:261–267.

Fafchamps, M., and S. Lund (2003) “Risk-Sharing Networks in Rural Philippines,” *Journal of Development Economics* 71:261–287.

Furusawa, T. and H. Konishi 2005 Free Trade Networks, *Japanese Economic Revie*w, 56: 144-164.

Gale, D., and S. Kariv (2003) “Bayesian Learning in Social Networks,” *Games and Economic Behavior* 45(2): 329–346.

Galeotti, A., S. Goyal, M.O. Jackson, F. Vega-Redondo, and L. Yariv 2010 ``Network Games,'' *Review of Economic Studies* Vol. 77: No. 1, pp 218-244.

Gilles, R.P. and S. Sarangi 2005 Stable Networks and Convex Payoffs, *Review of Economic Design*.

Glaeser, E., B. Sacerdote, and J. Scheinkman (1996) “Crime and Social Interactions,” *Quarterly Journal of Economics* 111:507–548.

Golub, B. and M.O. Jackson 2008 ``How Homophily affects the Speed of Learning and Best Response Dynamics" SSRN paper 1443787, http://papers.ssrn.com/abstract=1443787

Golub, B. and M.O. Jackson 2010 ``Naive Learning in Social Networks: Convergence, Influence, and Wise Crowds,'' *American Economic Journal: Microeconomics* 2(1): 112–49.

Golub, B. and M.O. Jackson 2011 ``Network Structure and the Speed of Learning: Measuring Homophily Based on its Consequences'' **forthcoming: Annals of Economics and Statistics:** http://papers.ssrn.com/abstract= 1784542

Granovetter, M. 1973 The Strength of Weak Ties, *American Journal of Sociology*, 78: 1360-1380.

Granovetter, M. [1974]1995 *Getting a Job: A Study of Contacts and Careers*, 2nd edition, University of Chicago Press, Chicago.

Granovetter, M. 1978 "Threshold Models of Collective Behavior," *American Journal of Sociology* 83(6): 1420–1443.

Granovetter, M. 1985 "Economic Action and Social Structure: The Problem of Embeddedness," *American Journal of Sociology* 91(3):481–510.

Herings, P.J.J., A. Mauleon and V. Vannetelbosch, 2004 Rationalizability for Social Environments, *Games and Economic Behavior* 49: 135-156.

Hojman, D. and A. Szeidl 2006 Endogenous Networks, Social Games and Evolution, *Games and Economic Behavior*, 551: 112-130.

Ioannides, Y.M. and L. Datcher-Loury 2004 Job Information Networks, Neighborhood Effects and Inequality, *Journal of Economic Literature*, 424: 1056-1093.

Jackson, M.O. 2003 The Stability and Efficiency of Economic and Social Networks, in *Advances in Economic Design*, edited by S. Koray and M. Ser-tel, Springer-Verlag: Heidelberg; reprinted in *Networks and Groups: Models of Strategic Formation*, edited by B. Dutta and M.O. Jackson, Springer-Verlag: Heidelberg.

Jackson, M.O. 2004 A Survey of Models of Network Formation: Stability and Efficiency, in *Group Formation in Economics; Networks, Clubs and Coalitions* , edited by Gabrielle Demange and Myrna Wooders, Cambridge University Press: Cambridge U.K.

Jackson, M.O. 2005 The Economics of Social Networks, Lecture prepared for the 9-thWorld Congress of the Econometric Society, August 2005, in *the Proceedings of the 9-thWorld Congress of the Econometric Society*, edited by Richard Blundell, Whitney Newey, and Torsten Persson , Cambridge University Press.

Jackson, M.O. 2005 "Allocation Rules for Network Games," *Games and Economic Behavior* 51: 128–154.

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

Jackson, M.O. and B.W. Rogers 2005 The Economics of Small Worlds, *Journal of the European Economic Association, Papers and Proceedings* 32(3): 617-627.

Jackson, M.O., and B.W. Rogers 2007 Meeting Strangers and Friends of Friends: How Random Are Social Networks? *American Economic Review,* 97(3):890–915

Jackson, M.O., and B.W. Rogers 2007b “Relating Network Structure to Diffusion Properties through Stochastic Dominance,” *B.E. Press Journal of Theoretical Economics* 7(1): 1–13.

Jackson, M.O. and van den Nouweland, A. 2005 Strongly Stable Networks, *Games and Economic Behavior*, 51: 420-444.

Jackson, M.O. and Watts, A. 2002 The Evolution of Social and Economic Networks, *Journal of Economic Theory*, 106(2): 265-295.

Jackson, M.O. and Wolinsky, A. 1996 A Strategic Model of Social and Economic Networks, *Journal of Economic Theory*, 71(1): 44-74.

Jackson, M.O. and L. Yariv (2007) The Diffusion of Behavior and Equilibrium Structure on Social Networks, *American Economic Review (Papers and Proceedings)*, Vol 97, No. 2, pp 92-98.

Johnson, C. and Gilles, R.P. 2000 Spatial Social Networks, *Review of Economic Design*, 5: 273-300.

Kakade, S.M., M. Kearns, L.E. Ortiz, 2004 Graphical Economics, *COLT* 23

Kakade, S.M., M. Kearns, L.E. Ortiz, R. Pemantle, and S. Suri 2004 Economic Properties of Social Networks, *Proceedings of Neural and Information Processing Systems, NIPS* MIT Press.

Katz, E., and P.F. Lazarsfeld (1955) *Personal Influence: The Part Played by People in the Flow of Mass Communication,* New York: Free Press.

Katz, L. 1953 A new status index derived from sociometric analysis, *Psychometrica* 18: 39-43.

Katz, M. and Shapiro, C. 1994 Systems Competition and Networks Effects, *Journal of Economic Perspectives*, 8: 93-115.

Kirman, A.P. 1983 Communication in Markets: A Suggested Approach, *Economics Letters*, 12:1-5.

Kirman, A., Oddou, C. and Weber, S. 1986 Stochastic Communication and Coalition Formation, *Econometrica*, 54: 129-138.

Kranton, R. and Minehart, D. 2001 A Theory of Buyer- Seller Networks, *American Economic Review*, 91 (3): 485-508.

Laschever, R. (2007) "The Doughboys Network: Social Interactions and Labor Market Outcomes of World War I Veterans," University of Illinois Working Paper, Champaign-Urbana.

Lopez-Pintado, D. 2008 Diffusion in Complex Social Networks," *Games and Economic Behavior* 62(2):573–590.

Marsden, P.V. (1987) "Core Discussion Networks of Americans," *American Sociological Review* 52:122–131.

Marsden, P.V. (1988) "Homogeneity in Confiding Relations," *Social Networks* 10:57–76.

Mauleon A. and V.J. Vannetelbosch 2004, Farsightedness and Cautiousness in Coalition Formation, *Theory and Decision*, 56: 291-324.

McPherson, M., L. Smith-Lovin, and J.M. Cook (2001) "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology* 27:415–444

Milgram, S. 1967, The small-world problem, *Psychology Today*, 2: 60-67.

Montgomery, J. 1991 Social Networks and Labor Market Outcomes, The *American Economic Review*, 81: 1408-1418.

Munshi, K. 2003 "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *Quarterly Journal of Economics* 118(2):549–597.

Mutuswami, S. and Winter, E. 2002 Subscription Mechanisms for Network Formation, *Journal of Economic Theory*, 106: 242-264.

Myers and G.P. Shultz 1951 *The Dynamics of a Labor Market*, Prentice-Hall: NY.

Myerson, R 1977 Graphs and Cooperation in Games, *Math. Operations Research*, 2: 225-229.

Page, F. and M. Wooders, and S. Kamat 2005 Networks and Farsighted Stability, *Journal of Economic Theory*, 1202: 257-269.

Pastor-Satorras, R., and A. Vespignani (2000) "Epidemic Spreading in Scale-Free Networks," *Physical Review Letters* 86:3200–3203.

Pastor-Satorras, R., and A. Vespignani (2001) "Epidemic Dynamics and Endemic States in Complex Networks," *Physical Review E* 63:066117.

Pellizzari, M. (2004) "Do Friends and Relatives Really Help in Getting a Job?" Centre of Economic Research Discussion Paper 623, London, London School of Economics.

Pennock, D.M., G.W. Flake, S. Lawrence, E.J. Glover, and C.L. Giles (2002) "Winners Don't Take All: Characterizing the Competition for Links on the Web," *Proceedings of the National Academy of Sciences, U.S.A.* 99(8):5207–5211.

Price, D.J.S. (1965) Networks of Scientific Papers, *Science* 149:510–515.

Price, D.J.S. 1976 A General Theory of Bibliometric and Other Cumulative Advantage Processes, *Journal American Society of Information Science* 27:292–306.

Rees, A. and G.P. Shultz 1970 *Workers in an Urban Labor Market*, University of Chicago Press: Chicago.

Reiss, A.J. (1980) "Understanding Changes in Crime Rates," in *Indicators of Crime and Criminal Justice: Quantitative Studies*, Washington, D.C.: Bureau of Justice Statistics.

Reiss, A.J. (1988) "Co-Offending and Criminal Careers," in *Crime and Justice: A Review of Research,* Vol. 10, M. Tonry, ed., Chicago: University of Chicago Press.

Rogers, B.W. 2006 A Strategic Theory of Interdependent Status, PhD dissertation: Caltech.

Sundararajan, A. 2005 Local Network Effects and Network Structure, mimeo: NYU Stern.

Slikker, M. and van den Nouweland, A. 2001 *Social and Economic Networks in Cooperative Game Theory*, Kluwer.

Tesfatsion, L. 1997 .A Trade Network Game with Endogenous Partner Selection,. in *Computational Approaches to Economic Problems*, H. Amman et al eds., Kluwer Academic Publishers, 249- 269.

Uzzi, B. 1996 The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect, *American Sociological Review*, 61:674–698.

Watts, D.J. 1999 *Small Worlds: The Dynamics of Networks between Order* *and Randomness*, Princeton University Press.

Watts, D.J. and S. Strogatz 1998, Collective dynamics of .small-world.net-works, *Nature*, 393: 440-442.

Figure 1

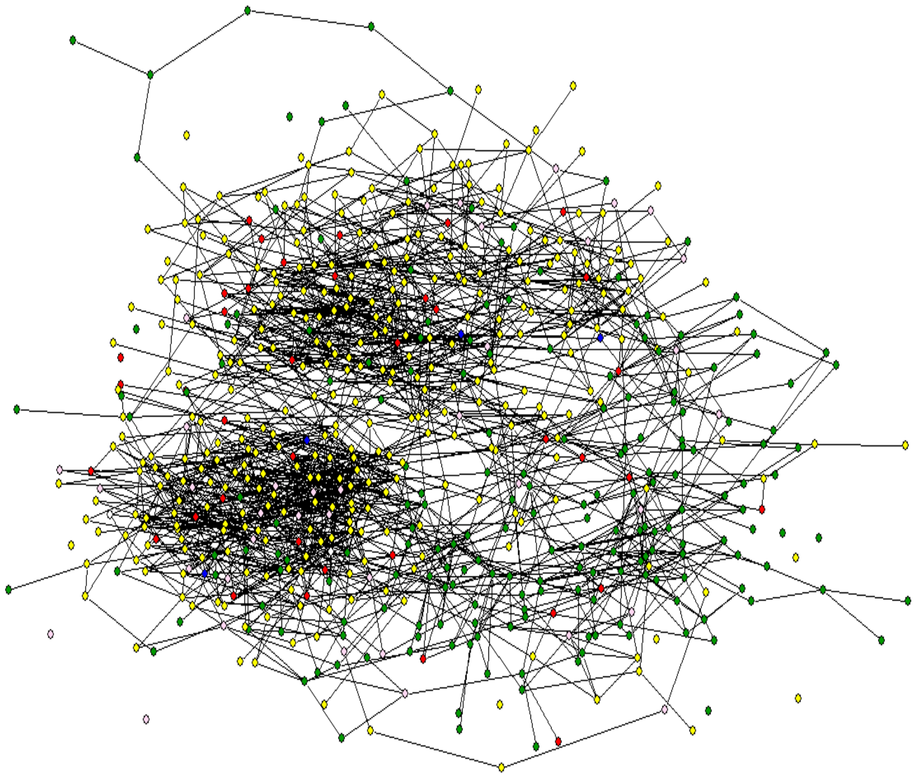

Figure 1: A network of friendships in a high school from the Add Health data set. Nodes are students and links indicate friendships. The color of the node indicates the student's race: Asian=green, Black=blue, Hispanic=red, White=yellow, Other=pink, Unknown=clear

Figure 2

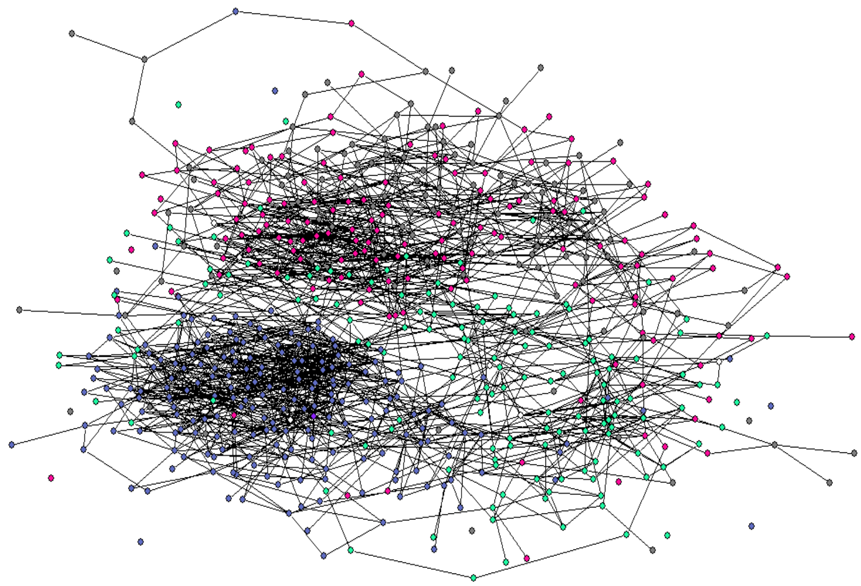

Figure 2: A network of friendships in a high school from the Add Health data set. Nodes are students and links indicate friendships. The color of the node indicates the student's grade: 9th grade = blue, 10th grade=green, 11th grade=pink, 12th grade = grey, unknown=clear