Networks in the Understanding of Economic Behaviors[†]

Matthew O. Jackson

The explosive changes in our abilities to communicate over distances—spurred by the evolution of communication technologies coupled with increased capabilities of the internet and social media—have made social networks very salient. Thus, it is perhaps more than coincidental that there has been a growth in network-related research that has accompanied the dramatic changes in the wiring of the world.[1]

Although the increased salience of networks may have awakened people to their importance, there is another more fundamental force that is driving the recent growth in the study of networks in economics. The main impetus is that, as economists endeavor to build better models of human behavior, they cannot ignore that humans are fundamentally a social species with interaction patterns that shape their behaviors. People's opinions, which products they buy, whether they invest in education, become criminals, and so forth, are all influenced by friends and acquaintances. Ultimately, the full network of relationships—how dense it is, whether some groups are segregated, who sits in central positions—affects how information spreads

---

[1] Note that this is not the first revolution in communication that has rewired human interaction, as the world has shrunk many times before: with the advent of letter writing, the telegraph, trains, the telephone, air travel, and others. Evidence from a study of the spread of the Black Death in 14th-century Europe suggests that the average social distance between individuals was much larger centuries ago than today (Marvel, Martin, Doering, Lusseau, and Newman 2013).

■ *Matthew O. Jackson is the William D. Eberle Professor of Economics at Stanford University, Stanford, California. He is also a Senior Fellow of the Canadian Institute for Advanced Research, Toronto, Ontario, Canada, and an external faculty member of the Santa Fe Institute, Santa Fe, New Mexico. His email address is jacksonm@stanford.edu.*

and how people behave. This impetus for the growth in network studies has been unleashed by the increased availability of data, which when coupled with increased computing power, allows us to analyze networks in economic settings in ways not previously possible.

In this paper, I discuss and illustrate this driving force, and describe some of the ways in which networks are helping economists to model and understand behavior. I begin with an example that demonstrates the sorts of things that researchers can miss if they do not account for network patterns of interaction. Next, to provide a broader perspective on the different ways in which networks provide insights into economic behaviors, I discuss a taxonomy of network properties and how they impact behaviors. Each of these properties and its impact is illustrated via applications.[2] Finally, I discuss an important frontier of networks research—developing tractable models of network formation—which is essential in addressing issues of endogeneity when estimating network effects on behavior.

Why Networks?

Do economists really need to map out the network of interactions in order to understand economic phenomena? Can't we ignore the patterns of interactions or proxy networks via other means such as peer groups or geography? Although we can circumvent network data in some situations, there are many settings in which network data provide essential insights into economic behaviors that are not available via other means. To make this point, I discuss a recent analysis in which network data play a vital role in disentangling potential explanations for behavior. In this example, households are the nodes of the network, and links in the network concern things like the exchange of favors, kinship, the sharing of advice, and so forth. There are also many other sorts of networks, including contractual relationships among firms, alliances between countries, overnight lending among banks, and others in which nodes are organizations and relationships are more formal but would illustrate similar points.

In Banerjee, Chandrasekhar, Duflo, and Jackson (2013), we used network data to analyze the diffusion of microfinance in a set of rural Indian villages. A bank entered 43 villages that had relatively limited access to formal loans and began offering microfinance loans (roughly $200 per unsecured loan over 50 weeks, with a limit of one loan at a time per household). Having sufficient participation in such a loan program is critical to making it viable, and so understanding what drives the participation in the program also becomes essential; especially since participation varied across villages, ranging from 7 percent to 44 percent of eligible households.

---

[2] There are large literatures analyzing social interactions that span several disciplines—with particular attention from sociology, economics, computer science, statistics, and statistical physics, and increasing attention from anthropology and political science. I do not attempt to survey the literatures here. Background can be found in Wasserman and Faust (1994); Granovetter (2005); Jackson (2005, 2008, 2011); Demange and Wooders (2005); Vega-Redondo (2007); Goyal (2007); Newman (2010); Benhabib, Bisin, and Jackson (2011); Jackson and Zenou (2014); Jackson, Rogers, and Zenou (2014).

In such poor villages, information is mainly spread via word-of-mouth. Thus, the bank entered a village and contacted a few individuals—shopkeepers, teachers, and self-help group leaders—to tell them that the bank would be offering loans and to spread the news to other members of the village. To be able to analyze how information spread, we surveyed the villagers before the bank entered the villages and mapped out networks of twelve types of relationships: who borrows kerosene and rice from whom, whom a given villager would borrow a small sum of money from, who seeks medical help from whom, who gets advice from whom, who visits whose house for dinner, and so forth.

This network information was vital in understanding the large variation in participation across villages. Was it that basic information about loan availability was not reaching many households in some villages, or was it that there were strong complementarities and dependencies in decisions to participate across households? Answering these questions is essential in shaping policies to enhance participation. If it was simply that information was not spreading in the villages with low participation rates, then the microfinance organization should work to increase awareness. If, instead, information was spreading extensively but villagers' decisions to participate were heavily tied to the decisions of their friends and participation never got rolling, then that would suggest very different policies—for example, policies designed to educate and encourage participation among well-positioned households.

To understand why network information was essential in answering these questions, let us begin by largely ignoring the rich network information that we have and simply analyzing decisions of households through a canonical peer-effects regression. In particular, consider a standard logistic regression of the form

$$\log\left(\frac{p_i}{1 - p_i}\right) = \beta X_i + \lambda F_i,$$

where p_i is the probability that household i participates in the loan program, X_i is a vector of household characteristics (caste, wealth, profession, and others), and F_i is the fraction of household i's "friends" who participate in the program. This is a standard way of formulating a discrete choice problem with peer influence (and similar results hold for a probit formulation). Here we are using the network information only in determining who a given household's friends are.[3]

---

[3] Actually, even this basic network information is helpful in such direct peer effect analyses in overcoming the "reflection problem" identified by Manski (1993). If the behavior of one individual depends on the average behavior of that person's peers (in a linear way), then if the peers are all peers of each other it can be impossible to disentangle the peer effect from the effects of the characteristics. The resulting set of equations and unknowns has many solutions: everyone influences everyone else's behavior and so those can be scaled up or down in ways that make it impossible to identify the peer influence separately from the influence of other exogenous characteristics and factors. This can be overcome with rich enough network data about who is friends with whom, as then the influences are not completely circular and the system of equations is no longer singular (Bramoullé, Djebbari, and Fortin 2009).

If we run such a regression on these data, we find an estimate of λ above 2 with high significance (with a p-value below .01). How does one interpret this finding? It suggests strong "peer effects": a given household's decision is highly correlated with their friends' decisions. For instance, increasing the fraction of the households' friends' participation from 0 to 1 increases the odds ratio $\frac{p_i}{1 - p_i}$ by a factor of roughly ten: so a household with a likelihood of $p_i = .05$ of participating based on its characteristics and having no friends participating, ends up with an odds ratio of nearly .5, or a participation likelihood of roughly $p_i = .30$ if all of its friends participate.

However, this analysis does not sort out whether the correlation in behaviors is due to basic information spreading and awareness, or other complementarities in decisions. To distinguish between these two possibilities, we can use the network to track the spread of information, and then re-estimate the logistic decision above *after we have controlled for information spread*. Thus, the richer estimation is based on a model that has two parts: households who have heard about microfinance randomly tell some of their friends, and once a household has heard about microfinance, *then* they make a decision based on their demographics and their friends' decisions. The model involves three key parameters: 1) households who choose to participate in the microfinance program tell their friends with some probability q^P; 2) the households who choose not to participate tell their friends with some probability q^N; and, 3) a parameter λ that captures endorsement effects: having more friends participate might provide additional information concerning how appropriate a loan is for a given household or might provide other sorts of peer pressure or influence, resulting in additional dependencies in decisions across households.

This estimation takes advantage of data about the full network structure to identify the parameters. Given that there is some randomness in the passing of information, a household's position relative to the first-informed people in the village affects the household's likelihood of becoming informed, and so the fuller network structure allows us to identify the information-passing probabilities. With low probabilities of information passing, information dies quickly and does not spread far beyond the initially informed households. With high probabilities, most households become informed, except households that are relatively isolated. With intermediate probabilities, households who are closer to those initially informed are more likely to hear, as are households who might be more distant but who have many paths to the initially informed nodes. Using the patterns of eventual participation as a function of network position, one can thus estimate the passing probabilities. With such probabilities in hand, one can then calculate the likelihood that any given household is informed as a function of their position in the network, and the logistic participation decision can be run *conditional* upon a household being informed.

The best-fitting parameters that we found are $q^P = .55$ (the probability that a participating household tells its neighbors in the network in a given time period), $q^N = .05$ (the probability that a nonparticipating household tells its neighbors in the network in a given time period), and $\lambda = -.2$ (the peer influence or endorsement

parameter). The information passing probabilities are both significantly different from zero and from each other, and λ is not significantly different from zero. These estimates paint a richer and very different picture than the original significantly positive estimate of λ, which simply reflected a correlation between household decisions.

The findings suggest that households who participate in microfinance are much more likely to inform their friends that microfinance is available than households who choose not to participate. Thus, households that have a high fraction of friends who are participating are much more likely to hear about microfinance than those with a low fraction, all else held constant. Once we condition for the likelihood of becoming informed, the local endorsement or peer effects are no longer significant. Thus, the full-network analysis suggests that much of the peer interaction in this setting involves people making each other aware of microfinance and that peer influences beyond that play an insignificant role in participation decisions.

Of course, as with any structural modeling, one cannot be sure of causation. In the network setting, one has to worry about the fact that networks are endogenous (a concern I discuss further below). Here, we are relying on identifying assumptions that take advantage of variation across villages regarding who was first informed: a given household's position in the network relative to the first-informed households affects its probability of hearing about microfinance, but does not substantially alter how endorsement and peer influence operate. The richness of the data and a modern ability to compute models enable us to check the robustness of such estimations and rule out alternative explanations, to provide some confidence that the model is capturing real phenomena. This is especially important as policies need to be enacted and we do not always have the luxury of controlled experiments or exogenous variation.

Network Properties and their Implications for Behavior

Beyond disentangling peer effects, network structure can help in many other ways in the understanding of economic behaviors. Given the complexity of networks, it is helpful to work with some basic characteristics that capture important aspects of network structure, and which have intuitive implications for behavior. It is useful to break these into two main categories: *macro* (global or aggregate) characteristics of networks, such as the density of connections or the segregation patterns among nodes, and *micro* (local or individual) characteristics of networks, such as the frequency with which two friends of a given node are friends with each other or how specific nodes are positioned in a network.[4]

Macro patterns of a network play primary roles in processes of diffusion and social learning, as well as in determining the extent to which disparate norms or cultures can exist within a given society. To fix ideas, let us continue with diffusion in

---

[4] See Jackson, Rogers, and Zenou (2014) for extensive references and more discussion of some of the properties discussed here.

mind: Which network characteristics determine whether diffusion of a new product (for example, microfinance) or idea is extensive or limited?

The most basic and intuitive "macro" property of networks that relates to diffusion and contagion is network density. Denser networks, in terms of average numbers of connections per node (for example, other households with whom a household exchanges favors, advice, and so on in our example above), lead to more extensive diffusion or contagion, all else held equal. The roots of this implication are seen in early studies of the spread of infectious diseases, and its understanding was fundamental in fighting diseases such as smallpox (for example, Anderson and May 1991). In this literature, the "basic reproduction number" tracks how many other people are newly infected by a typical infected individual. If the basic reproduction number is above one, then the disease expands and becomes endemic, while if it is below one, then the disease tends to die out—an insight that guides vaccination policies. The basic reproduction number depends not only on the characteristics of the disease, such as how easily it spreads from one individual to another, but also upon the network structure of interactions among individuals, which determines how many other people a given individual comes into contact with. Denser networks lead to more interactions and greater basic reproduction numbers (holding fixed the probability of transmission via any given interaction).

This basic insight from epidemiology translates readily to the diffusion of ideas, information, and products, and even some aspects of transmission of economic shocks and financial contagion.[5] In the microfinance example discussed above, a household has, on average, about 15 connections in the network—which in this context refers to how many other households with whom it has at least some sort of regular interaction such as borrowing/lending kerosene or rice, seeking/providing medical help, exchanging advice, and other interactions. Our estimated probabilities of transmitting information from one household to another of .55 and .05 per period (for participating and nonparticipating households, respectively), lead to corresponding basic reproduction numbers of 8.25 and .75 per period. Given that these are per-period estimates, and households communicate over more than one period (typically three to six in the data), both basic reproduction numbers end up above one, and so nontrivial diffusion of information is feasible in the villages.

Although there is some variation in density across villages, they all end up with basic reproduction numbers that average above one, so variation in network density does not seem to be the answer as to why there is substantial differences in participation across villages. Nonetheless, the large difference in the reproduction numbers between participating and nonparticipating households suggests that whether the first-informed households participate could be important: there is randomness in

---

[5] Financial cascades add a twist to contagion processes since the amount of exposure of one institution to another varies along with the network structure. This can result in much richer interactions between network structure and cascades (for discussion, see Allen and Babus 2009; Gai and Kapadia 2010; Acemoglu, Ozdaglar, and Tahbaz-Salehi forthcoming; and Elliot, Golub, and Jackson forthcoming). Similarly, adding marketing or pricing to diffusion enriches the process, as in Campbell (2013).

the diffusion process and it does not go on indefinitely, and so differences in the probability of one household telling another could have consequences regarding the extent of the diffusion. Indeed, participation of the first-informed households turns out to be a significant predictor of eventual participation.

More generally, these sorts of insights help make predictions about which societies are more susceptible to contagions as well as widespread diffusion, and can help in designing policies on a wide range of things from vaccinations to the subsidization of technology adoption. The predictions of such theories are also born out in the field beyond epidemiological studies. For example, Alatas, Banerjee, Chandrasekhar, Hanna, and Olken (2012) use network information to assess how social learning is affected by network structure. Based on network information from more than 600 Indonesian villages, they find that network density and other characteristics correlate with how much people know about other people in their villages. This is also an example in which both the theoretical and the empirical research rely on measures of networks that cannot be proxied for, even with ideal geographic data—especially given the tight village formations.

Beyond the density of a network, another fundamental network property that has far-reaching implications concerns segregation patterns. In particular, a feature observed in many social networks, referred to as "homophily" (a term due to Lazarsfeld and Merton 1954), is that similar individuals tend to be linked to each other. For example, consider the network pictured in Figure 1, which is a network of connections between households in one of the villages from the Banerjee et al. (2013) study. A connection in this network is based on the borrowing and lending of kerosene and rice. The round (solid fill) nodes are households that fall within the "scheduled castes" and "scheduled tribes" (those considered for affirmative action by the Indian government), and the square (checkered fill) nodes are the remaining "general" and "otherwise backward" caste designations. We see strong patterns of association by caste designation. Although it may not be surprising to see some segregation in the network given the history of castes in India, the strength of the division is striking. Moreover, such strong homophily is not unusual, and is observed in network data based on all sorts of attributes, including age, race, gender, profession, religion, education level, and others (for background, see McPherson, Smith-Lovin, and Cook 2001).

Such segregation patterns have profound consequences for behaviors within a network. Such segregation could clearly slow or impede diffusion or contagion that begins in one group from reaching others. However, there are more subtle implications. To see this most starkly, consider two groups, Circles and Squares, as in Figure 1. Suppose that Circles are more susceptible to becoming infected (by a virus or opinion), while Squares have a lower probability of becoming infected, and that the virus or opinion is relatively difficult to transmit. For example, older people might be more susceptible to flu; similarly, some new idea or product, news of which spreads via word-of-mouth, might be more attractive to people of a particular age, ethnic background, religion, profession, or other shared characteristic. If a network is well-integrated, then that diffusion may never gain traction, because

Figure 1
Homophily in an Indian Village
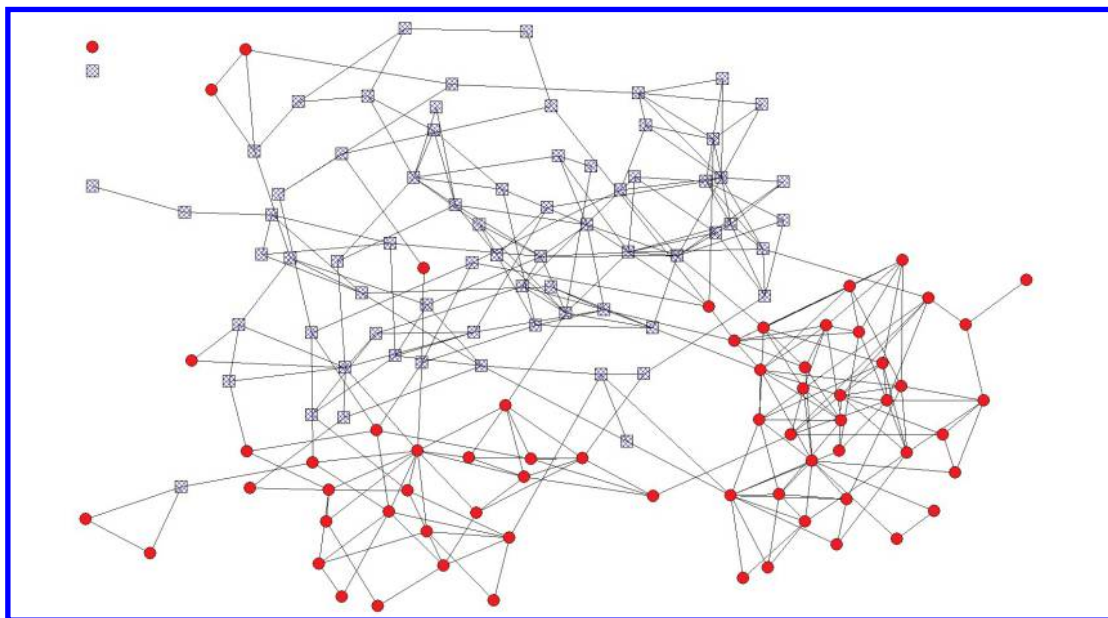
Notes: Nodes are households, and a link indicates that the pair of households report that at least one borrows kerosene and rice from the other. The round (solid fill) nodes are households that fall within the '"scheduled castes" and "scheduled tribes" (those considered for affirmative action by the Indian government), and the square (checkered fill) nodes are the remaining "general" and "otherwise backward" caste designations. The positioning of nodes is by a spring algorithm that groups nodes more closely together when they are linked to each other (and not based on geography, caste, or other node attribute). The frequency of links among pairs of households is .089 when both are within the same group, while it is only .006 when they are in different groups.

even if some Circle node becomes infected, the chance that it spreads to other nodes is small because the infected node has too few Circle neighbors and the virus spreads with low probability to any particular Circle neighbor, and with extremely low probability to any Square neighbors.

To see this most clearly, suppose that a typical node has four neighbors: two circles and two squares. Suppose also that the chance that an infected node infects any given neighbor is $1/3$ if that neighbor is a Circle and $1/8$ if the neighbor is a Square. So, a rough estimate of the basic reproduction number (discussed above) in a network in which a typical node has two Circle neighbors and two Square neighbors is $1/3 + 1/3 + 1/8 + 1/8 < 1$. In contrast, if the network exhibits the high level of homophily that we see in Figure 1, then many Circle nodes have at least four friends who are Circles, and thus have basic reproductive numbers of at least $4/3$, and so above 1. Therefore, in the network with high homophily, the diffusion or contagion can get traction in the Circle population and can even eventually infect

some of the Squares. The homophily helps to "incubate" the contagion or diffusion process, allowing it to gain hold, and then, once it becomes prevalent enough in one of the groups, it can spread more widely in the society.[6]

The implications of homophily reach well beyond basic diffusion processes. For example, consider the choice of an individual to pursue higher education or even just to participate in the labor force. The payoff from doing either of these is dependent on the decisions of a person's friends and acquaintances. For example, it is well-documented that social contacts play an important role in obtaining jobs.[7] Thus, if an individual's friends and acquaintances are educated and employed, the individual has a greater incentive to become educated and be part of the work force as he or she will have greater opportunities to take advantage of the education and/or participation in the labor force. This is further amplified by other complementarities in decisions: having friends who study, take entrance exams, complete high school, apply and go to college, interview for jobs, and so forth, not only raises one's prospects for future employment from following suit, but also provides an individual with valuable sources of information about how to do so. All of these factors tie friends' and acquaintances' decisions together, and produce network effects whereby individuals who are connected in the network tend to make decisions that are correlated, beyond the influence of any of their characteristics. When we couple this with homophily, we then end up with strong correlations by ethnicity, age, gender, religion, and other characteristics in decisions to acquire human capital and participate in the work force, well beyond what would be predicted by other channels such as parental influence and investment, discrimination, and the like. Network segregation patterns, and homophily in particular, allow multiple conventions to endure in a society and can be an important factor in understanding sustained differences in employment and educational attainment across groups. This has important policy implications, because of the dependence of behaviors among linked-individuals. For example, understanding that decisions are tied to each other means that subsidizing education individual-by-individual can be less effective than targeting subsidies in ways that take advantage of local network effects, a point discussed in Calvó-Armengol and Jackson (2004, 2007).

Beyond the macro patterns of networks, there are two particularly prominent aspects of networks from the micro or individual side that are very useful in

---

[6] Diffusion and learning processes in networks with homophily suggest particular measures of segregation (Morris 2000; Golub and Jackson 2012), and the effects can depend on heterogeneity in the population (Dandekar, Goel, and Lee 2013; Jadbabaie, Molavi, and Tahbaz-Salehi 2013), and the interaction structure (Reluga 2009; Galeotti and Rogers 2013; Jackson and López-Pintado 2013).

[7] The extensive literature on this subject dates to Rees (1966); see Ioannides and Loury (2004) for an overview. For example, Beaman (2012) takes clever advantage of a program to resettle political refugees in the United States to study the impact of social connections on employment outcomes. She finds that if refugees were randomly placed into an area where there was a relatively larger group of residents with similar ethnic backgrounds who had been in place for sufficient time, then the resettled refugees had a significantly higher employment rate than similar refugees placed into areas with smaller groups of tenured residents.

understanding economic behaviors: one concerns the "centrality" of individual nodes in a network, and the other concerns local "clustering" patterns.

Let us begin with centrality, as that sheds substantial light on the puzzle of why there was substantial variation in participation in microfinance across villages—ranging from 7 to 44 percent. As we have already discussed, differences in network densities across villages do not seem to vary enough across villages to be the explanation in this instance. Also, even though there is substantial (even extreme) homophily in the villages, the bank initially approached enough households (roughly between 5 and 15) in each village that it tended to end up "seeding" the diffusion process within the different main groups in each village. As it turns out, centrality of the first-informed households does vary substantially across villages and in ways that account for a substantial amount of the variation in participation across villages—especially when coupled with whether those first-informed households participate (which affects how likely they are to tell other households).

The idea that centrality of individuals impacts what information they have access to, how they behave, and how their behavior affects that of others was an early one in the literature (for example, Simmel 1908; Katz and Lazarsfeld 1955). This has led to many different measures of the centrality or influence of nodes in a network, ranging from just counting a node's number of connections, to more sophisticated and iterative methods that track how well connected a node's friends are (for background, see Jackson 2008).

In the context of the diffusion of microfinance, the centrality concepts that are strong predictors of participation measure the first-informed nodes' positions in the network based on their reach, where reach is naturally defined based on a diffusion process (for details, see Banerjee, Chandrasekhar, Duflo, and Jackson 2013, 2014). The fact that centrality helps explain the variation in diffusion does not contradict our earlier observation that the basic reproduction numbers in the villages were all above one. The basic reproduction number is, in essence, a limit concept, and with a relatively small number of first-informed households, there is still significant randomness in the diffusion of information, and so the first-informed households' positions in their village networks can matter significantly.

Beyond diffusion processes, centrality also plays an important role in peer influence. For example, better-connected individuals can have greater opportunities for complementarities in behaviors and exude greater influence. For example, an individual's decision to undertake a given behavior can have ripple effects well beyond his or her immediate neighbors, operating via a sort of social multiplier. That social multiplier naturally translates into a centrality measure. Ballester, Calvó-Armengol, and Zenou (2006) and Bramoullé, Kranton, and D'Amours (2014) show that a specific centrality measure based on such a social multiplier calculation captures the influence of individuals in networks with complementarities in actions. Moreover, Calvó-Armengol, Patacchini, and Zenou (2009) find evidence that an individual's centrality can have significant effects on an individual's education attainment as well as the education of others. Similarly, more general models of interactions with complementarities (as in Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv 2010)

make broad predictions about how individual behavior depends on position in a network as well as how that translates into overall behavior in the network, and some of those predictions have recently been confirmed in a laboratory setting (Charness, Feri, Meléndez-Jiménez, and Sutter forthcoming).

"Clustering" is our last example of a prominent network property that has important implications for economic behaviors. This is another property that is prevalent in social networks (see, for example, Watts and Strogatz 1998), like homophily, but its implications are some of the most intricate that we have discussed. Clustering, and various related measures, track whether an individual's friends form a tightly knit group rather than being separate from each other. For example, one measure of clustering in a network is the frequency with which two friends of a given person are friends with each other. If household A borrows/lends kerosene and rice with households B and C, does that mean that B and C are also likely to borrow/lend kerosene with each other? The answer is often yes; and for a variety of reasons, connections in social networks tend to be correlated, so that friendships are significantly clustered.

Theories of the implications of, and some reasons for, clustering are related to theories of social capital, such as those of Coleman (1988) and Putnam (2000). Those theories suggest that having high interconnectivity in a network on a *local* level is important in encouraging "cooperative" or "pro-social" behaviors. Recent game-theoretic models have helped clarify how local network patterns relate to behavior, and they shed light on measures related to clustering. Three main insights have emerged. First, in a highly-clustered network, if an individual misbehaves, then news of the misbehavior can quickly spread among that individual's friends—because the friends are also likely to be friends with each other—and so they can cooperate in retaliating against the misbehavior. Networks without such clustering patterns can insulate an individual from retaliation for bad behavior (Raub and Weesie 1990; Bloch, Genicot, and Ray 2008; Lippert and Spagnolo 2011; Ali and Miller 2012, 2013). Second, a pair of individuals who exchange favors, or who engage in any informal relationship that is not completely contractible, can have stronger incentives to behave efficiently if they have friends in common. Those common friends can react to misbehavior by ostracizing the individual, thereby providing incentives for that individual to behave cooperatively/efficiently (Jackson, Rodriguez-Barraquer, and Tan 2012). Third, high clustering can also affect diffusion. For instance, it may be that an individual will only undertake a new behavior if "enough" of his or her friends also do. For example, a choice of a video game platform may depend on whether one's friends are using the same platform. High clustering can allow groups of friends to coordinate on their behavior: as one example, all adopting a new technology that requires interaction with others (Centola 2011).

Some of the recent game-theoretic models of how local network patterns impact behavior are being examined in emerging empirical investigations. There is evidence consistent with some of the basic predictions, finding that various local link patterns are important predictors of money transfers and borrowing behavior (Karlan 2007; Karlan, Mobius, Rosenblat, and Szeidl 2009a, b; Blumenstock, Eagle,

and Fafchamps 2013; Kinnan and Townsend 2012); favor exchange (Jackson, Rodriguez-Barraquer, and Tan 2012); patterns of risk-sharing (Ambrus, Mobius, and Szeidl 2012); and product adoption (Centola 2010).

A Challenge: Endogenous Networks

Our discussion of clustering brings into focus another aspect of how social networks relate to behavior: networks are not only conduits for information or influence, but also adjust in reaction to behaviors. As one example, social norms can be very robust when people fear losing friends and their position in a network if they do not behave in prescribed ways. This symbiosis presents particular challenges for analyses of behaviors in social contexts because the coevolution means that it may be difficult to find exogenous sources of variation. One implication is that it becomes very important to understand how networks form. It is thus no surprise that the other broad branch of the economic networks literature, besides the one studying how network patterns of interactions determine behavior, concerns how networks form.

Network formation is important for various reasons, a couple of which are particularly germane to our discussion here. First, there is a fundamental question of whether the "right" networks form. Given that networks have important implications for behaviors, and that relationships have externalities, we need to know the extent to which networks that form in a decentralized manner end up being efficient from society's perspective. For example, when a researcher undertakes a new collaboration, that collaboration affects what he or she learns and can pass on to other researchers as well as how much time he or she can devote to other collaborations. These positive and negative externalities are not always incorporated in the decision to form the new relationship. The early economics literature was focused on network formation models and this question in particular. For example, a central theme in Jackson and Wolinsky (1996) was the differences between networks that form when individuals choose their relationships and the networks that maximize overall welfare.[8]

A second reason for studying network formation has come to the forefront with the new wave of studies of how networks affect behavior. People often form relationships because they wish to be connected to other individuals for economically relevant reasons like the benefits of collaboration, exchange, and sharing of information. As mentioned above, this endogeneity poses a huge challenge in analyzing how network structure affects behavior. Although selection and endogeneity issues

---

[8] The early literature explored this theme under a variety of formation processes (for example, Jackson and Wolinsky 1996; Dutta and Mutuswami 1997; Bala and Goyal 2000; Dutta and Jackson 2000; Currarini and Morelli 2000) and ultimately with a variety of stability definitions and in many different settings. Overviews appear in Jackson (2005, 2008) and Goyal (2007).

are well-known to economists working with observational data, the issues in network settings can be acute.

Although a researcher can control for observed characteristics, there could also be homophily driven by unobserved characteristics. For example, in the analysis of microfinance participation decisions in the Indian villages we discussed above, suppose that people who have similar levels of risk aversion are more likely to be friends with each other than people who have differing levels of risk aversion, all else held equal. Then correlations in loan-participation decisions across the network could be due to the correlations in risk preferences between linked individuals, not social influence. In our analysis such confounds for peer influence were likely not a major issue, as we found no peer effects after correcting for information passing (for which we had some identification), but it might not always work out that way. Indeed, failing to properly account for such homophily has been shown to lead to major biases in some imputed peer effects: Aral, Muchnik, and Sundararajan (2009) find such a bias in a network of estimating peer influence among 27 million users of an instant messaging network; and studies by Hsieh and Lee (2011) and Badev (2013) find such effects in friendship networks and decisions to smoke in US high schools.

One approach to dealing with this is to perform controlled experiments. This can be done by fully controlling the network of relationships within a lab (for example, Choi, Gale, and Kariv 2005; Kearns, Judd, Tan, and Wortman 2009), or by assigning subjects in the field positions in a network through which they must communicate (for example, Centola 2010, 2011; Goeree, McConnell, Mitchell, Tromp, and Yariv 2010). These techniques can test theories of how networks impact behaviors, which can then help inform further research as well as policy.

Although controlling networks themselves is not always possible, especially when dealing with the large-scale networks upon which some of the most interesting research questions apply, one can still exert some control over the interaction patterns and obtain robust conclusions. For example, Cai, de Janvry, and Sadoulet (forthcoming) examine the diffusion of a new form of insurance in rural China. After mapping out social networks, these researchers select some individuals to receive some financial education and then examine how that education (or lack thereof) affects decisions to participate in an insurance program by friends of farmers. By controlling who receives the education and the level of that education, they are able to see that the impact of having an additional "educated" friend is substantial—about half as much as the impact of directly receiving the education, and equivalent to a 15 percent drop in the price of insurance. Moreover, they are able to separate the effect of a friend's education from that friend's own decision to buy insurance and find that it is the information from the education, not the purchase decision, that impacts the friend's decision.

Of course, field experiments that appropriately control interactions or treatments are often impractical, especially in some very interesting areas for networks research: financial markets, crime, international alliances, and many others. Rather than giving up on research in these areas, or waiting for some lucky source of

exogenous variation or a powerful instrument, we must still make progress as the importance of the applications require it. This requires modeling network formation and handling the endogeneity issue head-on.

One innovative technique is based on the "latent space" (meaning unobserved characteristic space) estimation that has roots in statistics and has seen some limited use in network settings before (as in Hoff, Raftery, and Handcock 2002). The ideas behind these techniques were recently put to new use in network settings by Goldsmith-Pinkham and Imbens (2013) to account for unobserved homophily.

As a simple illustration, suppose that we are interested in how networks of student friendships affect scholastic achievement and that two traits influence how they form friendships: their ethnicity and how ambitious they are. To keep the illustration simple, suppose that there are two ethnicities and two levels of ambition. Suppose that a student has a high probability of forming a friendship with another student who is the same on both traits, a medium probability of forming a friendship with another student who is the same on one trait and different on the other, and a low probability of forming a friendship with another student who is different on both traits. Now consider a researcher who observes the ethnicities of students and the friendship network but not the students' ambition level. Without correcting for ambition, which could affect scholastic achievement, we might misattribute some achievement effects to the network interactions and even to the ethnicities.

The key idea in this approach is that homophily can help us to draw inferences about the unobserved trait of ambition. For instance, if we see two students who are friends but of different ethnicities, then that helps us to infer something about their similarity with regards to ambition. Since they differ in ethnicity, it would be very unlikely for the friendship to have formed if they also differed on ambition levels. So, observing the friendship in spite of their difference in ethnicity leads us to believe it is more likely that these two have similar ambition levels than two people picked at random in the population. Similarly, if we see two people who are not friends but who are of the same ethnicity, then they are less likely to be similar on the ambition dimension than two individuals chosen at random from the population. The next step is then to use this insight to help in estimating peer effects. This method thus allows one to see whether the inferred missing characteristic correlates with achievement, and then accounting for that correction eliminates some of the bias in the attribution to other factors, such as peer influence. Even without fully estimating the impact of the unobserved characteristics, as Goldsmith-Pinkham and Imbens (2013) discuss, one can still use these ideas to test for possible missing correlated attributes.

This latent-space technique can be powerful, but is particularly sensitive to the way in which the model is specified. We need not know which unobserved factors might affect network formation and behavior. In the example, the unobserved factor was called "ambition," but it could have been any set of variables that acted similarly. The tricky part is that one has to guess the right specification, for instance the form that unobserved variables take (discrete, continuous) and how they impact behaviors (linearly, nonlinearly, in concert with other variables, or some other choice),

and (to date) the approach is a parametric one.[9] The technique can be thought of as an analog to revealed preference theory: we do not observe consumers' preferences, but can estimate them assuming that consumers are maximizing some class of utility function subject to errors, similar to here presuming that homophily along some unobserved traits exists and takes some functional form, subject to errors. The inferred traits and homophily, just as the inferred utility function and preference maximization, can then be used to help understand behaviors. The quality of the inference then depends on the appropriateness and versatility of the models.

This brings us back to one of the most important areas of network research: developing richer, but still tractable, models of network formation. By richer, I mean the following. In the latent-space example, formation took place entirely at the link level: the decisions by two individuals to form a friendship was based solely on their characteristics. However, as we saw in our discussion of clustering, link formation in a network is generally correlated. For example, people meet each other through friends of friends; there are advantages to having friends be friends with each other; networks with local clique structures are better suited for enforcing behaviors; and, most basically, the value of a relationship generally depends on to whom the other party is linked.

So, ideally, we would like a model of network formation to do three things: i) allow for network effects and admit dependencies at more than the link level; ii) capture endogenous decisions to form relationships; and iii) be tractable enough to take to data.

Satisfying all three desiderata is challenging, and models that do so are only now emerging. With interdependencies in relationships, a model can no longer be specified at the link level but must involve a more holistic view of the network. Specifying the likelihood at a network level then runs into tractability issues, as the number of possible networks is exponential in the number of nodes, and so even with a tiny number of nodes it is impossible to calculate the relative likelihood of different networks. The model that has become a workhorse in the sociology literature, the exponential random graph model, allows for quite rich sets of dependencies in relationships and can be adapted to allow for endogenous decisions to form relationships as functions both of network position and node characteristics. Unfortunately, in the raw form, these models have severe computability issues and deficiencies in accuracy even though there is software that provides estimates.[10]

Some models do satisfy all of the above desiderata, either by building the network in a sequential fashion that allows new links to depend on the network existing at the time of a new node's entry (Barabasi and Albert 1999; Jackson and

---

[9] For more discussion of this technique and other issues related to modeling and estimating peer interactions with endogenous relationships, see the comments that follow the Goldsmith-Pinkham and Imbens (2013) paper.

[10] The Markov Chain Monte Carlo simulations used in the estimation procedures generally do not mix in less than exponential time (Bhamidi, Bresler, and Sly 2008), and so even though estimates are offered by the software, the estimates and the bootstrapped random errors can be inaccurate (see the discussion in Chandrasekhar and Jackson 2014).

Rogers 2007; Christakis, Fowler, Imbens, and Kalyanaraman 2010; Chaney forthcoming); building the network in such a way that very specific dependencies are admitted but are restricted in ways so as to not cause estimation problems (Mele 2010); or by modeling at a "subgraph level," allowing for local interdependencies but allowing for sufficient independence on a larger scale so as to enable easy estimation (Chandrasekhar and Jackson 2014).

In addition, one needs models that address homophily and help sort out the various forces that mold individual decisions to associate with others with similar characteristics. Homophily depends not only on preferences, but also depends on opportunities to form relationships, on norms, and on other factors. There are some emerging models of network formation that pay particular attention to understanding homophily (for example, Currarini, Jackson, and Pin 2009, 2010; Baccara and Yariv 2010; Tarbush and Teytelboym 2014; Graham 2014), and such models may be useful as steps in building models suited for general application.

The Future

Economists cannot ignore that many human decisions are made in the context of, and shaped by, networks of interactions. The variety of settings in which network analyses are providing deep insights is already substantial and continues to expand, including: development economics, labor markets, risk sharing, local public good provision, crime, education, social learning, bargaining and exchange, technology adoption, marketing, international trade, financial markets, and political economy. As such, it is inevitable that economic research on social networks will continue to grow and become part of an economist's basic toolbox. Many models of networked interactions are not difficult to simulate, and thus can be estimated in ways that are quite familiar to social scientists who work with data. The main novelty is including information about network structure, such as density, homophily, centralities, or details of who interacts with whom; and such data is increasingly easy to find or collect. The healthy interface between empirical observation and theory is helping us develop richer network models that are tractable and versatile—helping us to answer questions as to why certain patterns of behavior appear and what the ultimate welfare and policy implications are.

Our changing world makes such analyses imperative. Indeed, advances in technology make it possible not only to interact with greater numbers of people, but also for niche groups built around specific interests to attract and maintain critical masses. News can spread around the globe in minutes, people can very cheaply keep in touch with others and collaborate on projects at great distances, and one can find a wanting audience for almost any type of knowledge or opinion. These potentially profound changes can lead to an increasingly dense network of interactions, but also could result in more segregated interactions, as it becomes easier to locate and stay in communication with others who have similar characteristics or interests. As we have seen, network density and homophily are network

properties that have different implications for behavior. Which one will win out: will the world become more interconnected or polarized? The answer may depend on context, since as we glimpsed above, these changes may have different implications for a pure contagion process, such as the spread of a disease, compared to one that is more interactive like collective action or political activism. Understanding the ultimate impact of such changes on our beliefs, decisions, and behaviors, will draw on a well-developed science of networked interactions and provides a rich agenda for an exciting field.

References

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. Forthcoming. "Systemic Risk and Stability in Financial Networks." *American Economic Review.*

Alatas, Vivi, Abhijit Banerjee, Arun G. Chandrasekhar, Rema Hanna, and Benjamin A. Olken. 2012. "Network Structure and the Aggregation of Information: Theory and Evidence from Indonesia." NBER Working Paper 18351.

Ali, S. Nageeb, and David A. Miller. 2012. "Ostracism." https://dl.dropboxusercontent.com/u/1258389/Website/ostracism.pdf.

Ali, S. Nabeeb, and David Miller. 2013. "Enforcing Cooperation in Networked Societies." http://www-personal.umich.edu/~econdm/files/AliMiller-Cooperation.html.

Allen, Franklin, and Ana Babus. 2009. "Networks in Finance." In *The Network Challenge: Strategy, Profit, and Risk in an Interlinked World*, edited by P. R. Kleindorfer and Y. Gerry, 367–82. Wharton School Publishing.

Ambrus, Attila, Markus Mobius, and Adam Szeidl. 2012. "Consumption Risk-Sharing in Social Networks." NBER Working Paper 15719.

Anderson, Roy M., and Robert M. May. 1991. *Infectious Diseases of Humans.* Oxford University Press.

Aral, Sinan, Lev Muchnik, and Arun Sundararajan. 2009. "Distinguishing Influence-Based Contagions from Homophily-Driven Diffusion in Dynamic Networks." *PNAS* 106(51): 21544–49.

Baccara, Mariagiovanna, and Leeat Yariv. 2013. "Homophily in Peer Groups." *American Economic Journal: Microeconomics* 5(3): 69–96.

Badev, Anton. 2013. "Discrete Games in Endogenous Networks: Theory and Policy." PSC Working Paper PSC 13-05, Population Studies Center, University of Pennsylvania.

Bala, Venkatesh, and Sanjeev Goyal. 2000. "A Noncooperative Model of Network Formation." *Econometrica* 68(5): 1181–29.

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou. 2006. "Who's Who in Networks, Wanted: the Key Player." *Econometrica* 74(5): 1403–17.

Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson. 2013. "The Diffusion of Microfinance." *Science* 341(6144).

Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson. 2014. "Gossip: Identifying Central Individuals in a Social Network." NBER Working Paper 20422.

Barabasi, Albert-László, and Réka Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286(5439): 509–512.

Beaman, Lori A. 2012. "Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S." *Review of Economic Studies* 79(1): 128–61.

Benhabib, Jess, Alberto Bisin, and Matthew O. Jackson. 2011. *Handbook of Social Economics*, 2 vols. North-Holland.

Bhamidi, Shankar, Guy Bresler, and Allan Sly. 2008 "Mixing Time of Exponential Random Graphs." *Foundations of Computer Science, 2008, IEEE 49th Annual IEEE Symposium on.* IEEE, pp. 803–812.

Bloch, Francis, Garance Genicot, and Debraj Ray. 2008. "Informal Insurance in Social Networks." *Journal of Economic Theory* 143(1): 36–58.

Blumenstock, Joshua, Nathan Eagle, and Marcel Fafchamps. 2013. "Motives for Mobile Phone-Based Giving: Evidence in the Aftermath of Natural Disasters." http://www.jblumenstock.com/files/papers/jblumenstock_mobilequakes.pdf.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. 2009. "Identification of Peer Effects through Social Networks." *Journal of Econometrics* 150(1): 41–55.

Bramoullé, Yann, Rachel Kranton, and Martin D'Amours. 2014. "Strategic Interaction in Networks." *American Economic Review* 104(3): 898–930.

Burt, Ronald S. 1992. *Structural Holes: The Social Structure of Competition.* Harvard University Press.

Cai, Jing, Alain de Janvry, and Elisabeth Sadoulet. Forthcoming. "Social Networks and the Decision to Insure." *American Economic Journal: Applied Economics.*

Calvó-Armengol, Antoni, and Matthew O. Jackson. 2004. "The Effects of Social Networks on Employment and Inequality." *American Economic Review* 94(3): 426–54.

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

Calvó-Armengol, Antoni, Eleonara Patacchini, and Yves Zenou. 2009. "Peer Effects and Social Networks in Education." *Review of Economic Studies* 76(4): 1239–67.

Campbell, Arthur. 2013. "Word-of-Mouth Communication and Percolation in Social Networks." *American Economic Review* 103(6): 2466–98.

Centola, Damon. 2010. "The Spread of Behavior in an Online Social Network Experiment." *Science* 329(5996): 1194–97.

Centola, Damon. 2011. "An Experimental Study of Homophily in the Adoption of Health Behavior." *Science* 334(6060): 1269–72.

Chandrasekhar, Arun, and Matthew O. Jackson. 2014. "Tractable and Consistent Random Graph Models." Available at SSRN: http://ssrn.com/abstract=2150428.

Chaney, Thomas. Forthcoming. "The Network Structure of International Trade." *American Economic Review.*

Charness, Gary, Francesco Feri, Miguel Meléndez-Jiménez, and Matthias Sutter. Forthcoming. "Experimental Games on Networks: Underpinnings of Behavior and Equilibrium Selection." *Econometrica.*

Choi, Syngjoo, Doublas Gale, and Shachar Kariv. 2005. "Behavioral Aspects of Learning in Social Networks: An Experimental Study." *Advances in Applied Microeconomics: A Research Annual,* vol. 13, pp. 25–61.

Christakis, Nicholas A., James H. Fowler, Guido W. Imbens, and Karthik Kalyanaraman. 2010. "An Empirical Model for Strategic Network Formation." NBER Working Paper 16039.

Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94(Supplement): S95–S120.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin. 2009. "An Economic Model of Friendship: Homophily, Minorities, and Segregation." *Econometrica* 77(4): 1003–1045.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin. 2010. "Identifying the Roles of Race-based Choice and Chance in High School Friendship Network Formation." *Proceedings of the National Academy of Sciences* 107(11): 4857–61.

Currarini, Sergio, and Massimo Morelli. 2000. "Network Formation with Sequential Demands." *Review of Economic Design* 5(3): 229–50.

Dandekar, Pranav, Ashish Goel, and David T. Lee. 2013. "Biased Assimilation, Homophily, and the Dynamics of Polarization." *Proceedings of the National Academy of Sciences* 110(15): 5791–96.

Demange, Gabrielle, and Myrna Wooders. 2005. *Group Formation in Economics; Networks, Clubs and Coalitions.* Cambridge University Press.

Dutta, Bhaskar, and Matthew O. Jackson. 2000. "The Stability and Efficiency of Directed Communication Networks." *Review of Economic Design* 5(3): 251–72.

Dutta, Bhaskar, and Suresh Mutuswami. 1997. "Stable Networks." *Journal of Economic Theory* 76(2): 322–44.

Elliott, Matthew, Benjamin Golub, and Matthew O. Jackson. Forthcoming. "Financial Networks and Contagion." *American Economic Review.*

Furusawa, Taiji, and Hideo Konishi. 2007. "Free

Trade Networks." *Journal of International Economics* 7(2): 310–35.

Gai, Prasanna, and Sujit Kapadia. 2010. "Contagion in Financial Networks." *Proceedings of the Royal Society A*, 466(2120): 2401–23.

Galeotti, Andrea, Sanjeev Goyal, Matthew O. Jackson, Fernando Vega-Redondo, and Leeat Yariv. 2010. "Network Games." *Review of Economic Studies* 77(1): 218–44.

Galeotti, Andrea, and Brian W. Rogers. 2013. "Strategic Immunization and Group Structure." *American Economic Journal: Microeconomics* 5(2): 1–32.

Goeree, Jacob K., Margaret A. McConnell, Tiffany Mitchell, Tracey Tromp, and Leeat Yariv. 2010. "The 1/d Law of Giving." *American Economic Journal: Microeconomics* 2(1): 183–203.

Goldsmith-Pinkham, Paul, and Guido W. Imbens. 2013. "Social Networks and the Identification of Peer Effects." *Journal of Business and Economic Statistics* 31(3): 253–64.

Golub, Benjamin, and Matthew O. Jackson. 2012. "How Homophily Affects the Speed of Learning and Best-Response Dynamics." *Quarterly Journal of Economics* 127(3): 1287–1338.

Goyal, Sanjeev 2007. *Connections: An Introduction to the Economics of Networks.* Cambridge University Press.

Graham, Bryan S. 2014. "An Empirical Model of Network Formation: Detecting Homophily when Agents Are Heterogeneous." Berkeley.

Granovetter, Mark. 2005 "The Impact of Social Structure on Economic Outcomes." *Journal of Economic Perspectives* 19(1): 33–50.

Hagenbach, Jeanne, and Frédéric Koessler. 2010. "Strategic Communication Networks." *Review of Economic Studies* 77(3): 1072–99.

Hoff, Peter, Adrian E. Raftery, and Mark S. Handcock. 2002. "Latent Space Approaches to Social Network Analysis." *Journal of the American Statistical Association* 97(460): 1090–98.

Hsieh, Chih-Sheng, and Lung-fei. Lee. 2011. "A Social Interactions Model with Endogenous Friendship Formation and Selectivity." https://www.american.edu/cas/economics/info-metrics/pdf/upload/Chin-Sheng-Hsieh-paper-March-2012-conference.pdf.

Ioannides, Yannis M., and Linda Datcher Loury. 2004. "Job Information Networks, Neighborhood Effects and Inequality." *Journal of Economic Literature* 42(4): 1056–93.

Jackson, Matthew O. 2005. "A Survey of Network Formation Models: Stability and Efficiency." Chap. 1 in *Group Formation in Economics; Networks, Clubs, and Coalitions,* edited by Gabrielle Demange and Myrna Wooders. Cambridge University Press.

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

Jackson, Matthew O. 2011 "An Overview of Social Networks and Economic Applications." Chap. 12 in *Handbook of Social Economics,* vol. 1, edited by Jess Benhabib, Alberto Bisin, and Matthew O. Jackson. San Diego: North Holland.

Jackson, Matthew O., and Dunia López-Pintado. 2013. "Diffusion and Contagion in Networks with Heterogeneous Agents and Homophily." *Network Science* 1(1): 49–67.

Jackson, Matthew O., Tomas Rodriguez-Barraquer, and Xu Tan. 2012. "Social Capital and Social Quilts: Network Patterns of Favor Exchange." *American Economic Review* 102(5): 1857–97.

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

Jackson, Matthew. O., Brian W. Rogers, and Yves Zenou. 2014. "Economic Consequences of Social Network Structure." Available soon at SSRN.

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

Jackson, Matthew O., and Yves Zenou. 2014. "Games on Networks." In [*Handbook of*] *Game Theory with Economic Applications,* vol. 4, edited by Young, H. Peyton and Shmuel Zamir. Elsevier.

Jadbabaie, Ali, Pooya Molavi, and Alireza Tahbaz-Salehi. 2013. "Information Heterogeneity and the Speed of Learning in Social Networks." *Columbia Business School Research Paper 13-28.*

Karlan, Dean S. 2007. "Social Connections and Group Banking." *Economic Journal* 117(517): F52–F84.

Karlan, Dean, Markus M. Mobius, Tanya S. Rosenblat, and Adam Szeidl. 2009a. "Measuring Trust in Peruvian Shantytowns." http://www.personal.ceu.hu/staff/Adam_Szeidl/papers/trust_peru.pdf.

Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl. 2009b. "Trust and Social Collateral." *Quarterly Journal of Economics* 124(3): 1307–61.

Katz, Elihu, and Paul F. Lazarsfeld. 1955. *Personal Influence: The Part Played by People in the Flow of Mass Communication.* Glencoe, IL: Free Press.

Kearns, Michael, Stephen Judd, Jinsong Tan, and Jennifer Wortman. 2009. "Behavioral Experiments on Biased Voting in Networks." *PNAS* 106(5): 1347–52.

Kinnan, Cynthia, and Robert Townsend. 2012. "Kinship and Financial Networks, Formal Financial Access, and Risk Reduction." *American Economic Review* 102(3): 289–93.

Lazarsfeld, Peter F., and Robert K. Merton.

1954. "Friendship as a Social Process: A Substantive and Methodological Analysis." In *Freedom and Control in Modern Society,* edited by Morroe Berger, 18–66. New York: Van Nostrand.

Lippert, Steffen, and Giancarlo Spagnolo. 2011. "Networks of Relations and Word-of-Mouth Communication." *Games and Economic Behavior* 72(1): 202–17.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60(3): 531–42.

Marvel, Seth A., Travis Martin, Charles R. Doering, David Lusseau, and M. E. J. Newman. 2013. "The Small-World Effect is a Modern Phenomenon." arXiv:1310.2636 [physics.soc-ph].

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27: 415–44.

Mele, Angelo. 2010. "A Structural Model of Segregation in Social Networks." NET Institute Working Paper 10-16.

Morris, Stephen. 2000. "Contagion." *Review of Economic Studies* 67(1): 57–78.

Newman, M. E. J. 2010. *Networks: An Introduction.* Oxford University Press.

Putnam, Robert D. 2000. *Bowling Alone: The Collapse and Revival of American Community.* New York: Simon and Schuster.

Raub, Werner, and Jeroen Weesie. 1990. "Reputation and Efficiency in Social Interactions: An Example of Network Effects." *American Journal of Sociology* 96(3): 626–54.

Rees, Albert. 1966. "Information Networks in Labor Markets." *American Economic Review* 56(1/2): 559–66.

Reluga, Timothy C. 2009. "An SIS Epidemiology Game with Two Subpopulations." *Journal of Biological Dynamics* 3(5): 515–31.

Simmel, Georg. 1908. *Sociology: Investigations on the Forms of Sociation.* Leipzig: Duncker and Humblot.

Tarbush, Bassel, and Alexander Teytelboym. 2014. "Friending." http://t8el.com/wp-content/uploads/2013/11/TarbushTeytelboymFriending.pdf.

Vega-Redondo, Fernando. 2007. *Complex Social Networks.* (Econometric Society Monographs.) Cambridge University Press.

Wasserman, Stanley, and Katherine Faust. 1994. "Social Network Analysis: Methods and Applications." Cambridge University Press.

Watts, Duncan J., and Steven H. Strogatz. 1998. "Collective Dynamics of 'Small-World' Networks." *Nature*, 393(June, 4): 440–42.