Diffusion, Strategic Interaction, and Social Structure

Matthew O. Jackson[*] and Leeat Yariv[†]

Final Revision: June, 2010

Abstract

We provide an overview and synthesis of the literature on how social networks influence behaviors, with a focus on diffusion. We discuss some highlights from the empirical literature on the impact of networks on behaviors and diffusion. We also discuss some of the more prominent models of network interactions, including recent advances regarding interdependent behaviors, modeled via games on networks.

Keywords: Diffusion, Learning, Social Networks, Network Games, Graphical Games, Games on Networks

JEL Classification Codes: D85, C72, L14, Z13

1 Introduction

How we act, as well as how we are acted upon, are to a large extent influenced by our relatives, friends, and acquaintances. This is true of which profession we decide to pursue, whether or not we adopt a new technology, as well as whether or not we catch the flu. In this chapter we provide an overview of research that examines how social structure impacts economic decision making and the diffusion of innovations, behaviors, and information.

We begin with a brief overview of some of the stylized facts on the role of social structure on diffusion in different realms. This is a rich area of study that includes a vast set of case studies suggesting some important regularities. With that empirical perspective, we then discuss insights from the epidemiology and random graph literatures that help shed light

---

[*]Department of Economics, Stanford University, Stanford CA 94305, USA. http://www.stanford.edu/~jacksonm e-mail: jacksonm@stanford.edu

[†]Division of the Humanities and Social Sciences, Caltech, Pasadena, CA 91125. http://www.hss.caltech.edu/~lyariv/index.htm e-mail: lyariv@hss.caltech.edu

on the spread of infections throughout a society. Contagion of this form can be thought of as a basic, but important, form of social interaction, where the social structure largely determines patterns of diffusion. This literature presents a rich understanding of questions such as: "How densely connected does a society have to be in order to have an infection reach a nontrivial fraction of its members?", "How does this depend on the infectiousness of the disease?", "How does it depend on the particulars of the social network in place?", "Who is most likely to become infected?", and "How widespread is an infection likely to be?", among others. The results on this apply beyond infectious diseases, and touch upon issues ranging from the spread of information to the proliferation of ideas.

While such epidemiological models provide a useful look at some types of diffusion, there are many economically relevant applications in which a different modeling approach is needed, and, in particular, where the interaction between individuals requires a game theoretic analysis. In fact, though disease and the transmission of certain ideas and bits of information can be modeled through mechanical or purely probabilistic sorts of diffusion processes, there are other important situations where individuals take decisions and care about how their social neighbors or peers behave. This applies to decisions of which products to buy, which technology to adopt, whether or not to become educated, whether to learn a language, how to vote, and so forth. Such interactions involve equilibrium considerations and often have multiple potential outcomes. For example, an agent might care about the proportion of neighbors adopting a given action, or might require some threshold of stimulus before becoming convinced to take an action, or might want to take an action that is different from that of his or her neighbors (e.g., free-riding on their information gathering if they do gather information, but gathering information him or herself if neighbors do not). Here we provide an overview of how the recent literature has modeled such interactions, and how it has been able to meld social structure with predictions of behavior.

2 Empirical Background: Social Networks and Diffusion

There is a large body of work that identifies the effects of social interactions on a wide range of applications spanning fields: epidemiology, marketing, labor markets, political science, and agriculture are only a few.

While some of the empirical tools for the analysis of social interaction effects have been described in Block, Blume, and Durlauf (Chapter ?, this volume), and many of their implementations for research on housing decisions, labor markets, addictions, and more, have been discussed in Ioannides (Chapter ?, this volume), Epple and Romano (Chapter ?, this volume), Topa (Chapter ?, this volume), Fafchamps (Chapter ?, this volume), Jackson (Chapter ?, this volume), and Munshi (Chapter ?, this volume), we now describe empirical work that ties directly to the models that are discussed in the current chapter. In particular, we discuss several examples of studies that illustrate how social structure impacts outcomes and behaviors.

The relevant studies are broadly divided into two classes. First, there are cross-sectional studies that concentrate on a snapshot of time and look for correlations between social interaction patterns and observable behaviors. This class relates to the analysis below of strategic games played by a network of agents. While it can be very useful in identifying correlations, it is important to keep in mind that identifying causation is complicated without the fortuitous exogenous variation or structural underpinnings. Second, there are longitudinal studies that take advantage of the inherent dynamics of diffusion. Such studies have generated a number of interesting observations and are more suggestive of some of the insights the theoretical literature on diffusion has generated. Nonetheless, these sorts of studies also face challenges in identifying causation because of potential unobserved factors that may contemporaneously influence linked individuals.

The empirical work on these topics is immense and we provide here only a narrow look of the work that is representative of the *type* of studies that have been pursued and relate

to the focus of this chapter.

2.1 The Effects of Networks from Static Perspectives

Studies that are based on observations at one point of time most often compare the frequency of a certain behavior or outcome across individuals who are connected as opposed to ones that are not. For example, Glaeser, Sacerdote, and Scheinkman (1996) showed that the structure of social interactions can help explain the cross-city variance in crime rates in the U.S.; Bearman, Moody, and Stovel (2004) examined the network of romantic connections in high-school, and its link to phenomena such as the spread of sexually transmitted diseases (see the next subsection for a discussion of the spread of epidemics). Such studies provide important evidence for the correlation of behaviors with characteristics of individuals' connections. In the case of diseases, they provide some direct evidence for diffusion patterns.

With regards to labor markets, there is a rich set of studies showing the importance of social connections for diffusing information about job openings, dating back to Rees (1966) and Rees and Schultz (1970). Influential studies by Granovetter (1973, 1985, 1995) show that even casual or infrequent acquaintances ("weak ties") can play a role in diffusing information. Those studies were based on interviews that directly ask subjects how they obtained information about their current jobs. Other studies, based on outcomes, such as Topa (2001), Conley and Topa (2002), and Bayer, Ross, and Topa (2008), identify local correlations in employment status within neighborhoods in Chicago, and consider neighborhoods that go beyond the geographic but also include proximity in other socioeconomic dimensions, examining the extent to which local interactions are important for employment outcomes. Bandiera, Barankay, and Rasul (2008) create a bridge between network formation (namely, the creation of friendships amongst fruit pickers) and the effectiveness of different labor contracts. The extensive literature on networks in labor markets[1] documents the important role of social connections in transmitting information about jobs, and also differentiates between

---

[1]For more references see the survey by Ioannides and Datcher-Loury (2004), Chapter 10 in Jackson (2008), and Jackson(Chapter ?, this volume).

different types of social contacts and shows that even weak ties can be important in relaying information.

There is further (and earlier) research that examines the different roles of individuals in diffusion. Important work by Katz and Lazarsfeld (1955) (building on earlier studies of Lazarsfeld, Berelson, and Gaudet (1944), Merton (1948), and others), identifies the role of "opinion leaders" in the formation of various beliefs and opinions. Individuals are heterogeneous (at least in behaviors), and some specialize in becoming well informed on certain subjects, and then information and opinions diffuse to other less informed individuals via conversations with these opinion leaders. Lazarsfeld, Berelson, and Gaudet (1944) study voting decisions in an Ohio town in the 1940 U.S. presidential campaign, and document the presence and significance of such opinion leaders. Katz and Lazarsfeld (1955) interviewed women in Decatur, Illinois, and asked about a number of things such as their views on household goods, fashion, movies, and local public affairs. When women showed a change in opinion in follow-up interviews, Katz and Lazarsfeld traced influences that led to the change in opinion, again finding evidence for the presence of opinion leaders.

Diffusion of new products is understandably a topic of much research. Rogers (1995) discusses numerous studies illustrating the impacts of social interactions on the diffusion of new products, and suggests various factors that impact which products succeed and which products fail. For example, related to the idea of opinion leaders, Feick and Price (1987) surveyed 1531 households and provided evidence that consumers recognize and make use of particular individuals in their social network termed "market mavens," those who have a high propensity to provide marketplace and shopping information. Whether or not products reach such mavens can influence the success of a product, independently of the product's quality. Tucker (2008) uses micro-data on the adoption and use of a new video-messaging technology in an investment bank consisting of 2118 employees. Tucker notes the effects of the underlying network in that employees follow the actions of those who either have formal power, or informal influence (which is, to some extent, endogenous to a social network).

In the political context, there are several studies focusing on the social sources of infor-

mation electors choose, as well as on the selective mis-perception of social information they are exposed to. A prime example of such a collection of studies is Huckfeldt and Sprague (1995), who concentrated on the social structure in South Bend, Indiana, during the 1984 elections. They illustrated the likelihood of socially connected individuals to hold similar political affiliations. In fact, the phenomenon of individuals connecting to individuals who are similar to them is observed across a wide array of attributes and is termed by sociologists *homophily* (for overviews see McPherson, Smith-Lovin, and Cook, 2001, Jackson, 2007, as well as the discussion of homophily in Jackson, Chapter ? in this volume).

While cross-sectional studies are tremendously interesting in that they suggest dimensions on which social interactions may have an impact, they face many empirical challenges. Most notably, correlations between behaviors and outcomes of individuals and their peers may be driven by common unobservables and therefore be spurious. Given the strong homophily patterns in many social interactions, individuals who associate with each other often have common unobserved traits, which could lead them to similar behaviors. This makes empirical analysis of the social impact on diffusion of behaviors based on cross sectional data difficult to draw (causal) conclusions from.[2]

Given some of the challenges with causal inference based on pure observation, laboratory experiments and field experiments are quite useful in eliciting the effects of real-world networks on fully controlled strategic interactions, and are being increasingly utilized. As an example, Leider, Mobius, Rosenblat, and Do (2009) elicited the friendship network among undergraduates at a U.S. college and illustrated how altruism varies as a function of social proximity. In a similar setup, Goeree, McConnell, Mitchell, Tromp, and Yariv (2010) elicited the friendship network in an all-girls school in Pasadena, CA, together with girls' characteristics and later ran dictator games with recipients who varied in social distance. They identified a "1/d Law of Giving," in that the percentage given to a friend was inversely re-

---

[2]In fact, Aral and Walker (2010) use different advertising methods on random samples of Facebook users and illustrate that the similarity in attributes may be an important component in observed patterns of network effects in diffusion. This is discussed at more length in Jackson (Chapter ?, this volume).

lated to her social distance in the network.[3] Various field experiments, such as those by Duflo and Saez (2003), Karlan, Mobius, Rosenblat, and Szeidl (2009), Dupas (2010), Beaman and Magruder (2010), and Feigenberg, Field, and Pande (2010), also provide some control over the process, while working with real-world network structures to examine network influences on behavior.[4]

Another approach that can be taken to infer causal relationships is via structural modeling. As an example, one can examine the implications of a particular diffusion model for the patterns of adoption that should be observed. One can then infer characteristics of the process by fitting the process parameters to best match the observed outcomes in terms of behavior. For instance, Banerjee, Chandrasekhar, Duflo, and Jackson (2010) use such an approach in a study of the diffusion of microfinance participation in rural Indian villages. Using a model of diffusion that incorporates both information and peer effects, they then fit the model to infer the relative importance of information diffusion versus peer influences in accounting for differences in microfinance participation rates across villages. Of course, in such an approach one is only as confident in the causal inference as one is confident that the model is capturing the essential underpinnings of the diffusion process.

The types of conclusions that have been reached from these cross sectional studies can be roughly summarized as follows. First, in a wide variety of settings, associated individuals tend to have correlated actions and opinions. This does not necessarily embody diffusion or causation, but as discussed in the longitudinal section below, there is significant evidence of social influence in diffusion patterns as well. Second, individuals tend to associate with others who are similar to themselves, in terms of beliefs and opinions. This has an impact on the structure of social interactions, and can affect diffusion. It also represents an empirical quandary of the extent to which social structure influences opinions and behavior as opposed to the reverse (that can partly be sorted out with careful analysis of longitudinal data).

---

[3]For a look at a few network experiments that are not based on a real-world social structure, see Kosfeld (2004).

[4]Baccara, Imrohoroglu, Wilson, and Yariv (2009) use field data to illustrate how different layers of networks (social and professional) can affect outcomes differentially.

Third, individuals fill different roles in a society, with some acting as "opinion leaders," and being key conduits of information and potential catalysts for diffusion.

2.2 The Effects of Networks over Time

Longitudinal data can be especially important in diffusion studies, as they provide information on how opinions and behaviors move through a society over time. They also help sort out issues of causation as well as supply specific information about the extent to which behaviors and opinions are adopted dynamically, and by whom. Such data can be especially important in going beyond the documentation of correlation between social connections and behaviors, and illustrating that social links are truly the conduits for information and diffusion if one is careful to track what is observed by whom at what point in time, and can measure the resulting changes in behavior. For example, Conley and Udry (2008) show that pineapple growers in Ghana tend to follow those farmers who succeed in changing their levels of use of fertilizers. Through careful examination of local ties, and the timing of different actions, they trace the influence of the outcome of one farmer's crop on subsequent behavior of other farmers.

More generally, diffusion of new technologies is extremely important when looking at transitions in agriculture. Seminal studies by Ryan and Gross (1943) and Griliches (1957) examined the effects of social connections on the adoption of a new behavior, specifically the adoption of hybrid corn in the U.S. Looking at aggregate adoption rates in different states, these authors illustrated that the diffusion of hybrid corn followed an S-shape curve over time: starting out slowly and accelerating, and then ultimately decelerating.[5] Foster and Rosenzweig (1995) collected household-level panel data from a representative sample of rural Indian households having to do with the adoption and profitability of high-yielding seed varieties (associated with the Green Revolution). They identified significant learning-by-doing, where some of the learning was through neighbors' experience. In fact, the observation that adoption rates of new technologies, products, or behaviors exhibit S-shaped curves can

---

[5]See Young (2010) for a complementary analysis to that of Griliches (1957).

be traced to very early studies, such as Tarde (1903), who discussed the importance of imitation in adoption. Such patterns are found across many applications (see Mahajan and Peterson, 1985 and Rogers, 1995).

Understanding diffusion is particularly important for epidemiology and medicine for several reasons. For one, it is important to understand how different types of diseases spread in a population. In addition, it is crucial to examine how new treatments get adopted. Colizza, Barrat, Barthelemy, and Vespignani (2006, 2007) tracked the spread of severe acute respiratory syndrome (SARS) across the world combining census data with data on almost all air transit during the years 2002-2003. They illustrated the importance of structures of long-range transit networks for the spread of an epidemic. Coleman, Katz, and Menzel (1966) is one of the first studies to document the role of social networks in diffusion processes. The study looked at the adoption of a new drug (tetracycline) by doctors and highlighted two observations. First, as with hybrid corn, adoption rates followed an S-shape curve over time. Second, adoption rates depended on the density of social interactions. Doctors with more contacts (measured according to the trust placed in them by other doctors) adopted at higher rates and earlier in time.[6]

Diffusion can occur in many different arenas of human behavior. For example Christakis and Fowler (2007) document influences of social contacts on obesity levels. They studied the social network of 12,067 individuals in the U.S. assessed repeatedly from 1971 to 2003 as part of the Framingham Heart Study. Concentrating on body-mass index, Christakis and Fowler found that a person's chances of becoming obese increased by 57% if he or she had a friend who became obese, by 40% if he or she had a sibling who became obese, and by 37% if they had a spouse who became obese in a previous period. The study controls for various selection effects, and takes advantage of the direction of friendship nominations to help sort out causation. For example, Christakis and Fowler find a significantly higher increase of

---

[6] As a caveat, Van den Bulte and Lilien (2001) add controls having to do with marketing exposure of the doctors in the study and show that the social effects may be mitigated. Nonetheless, further studies such as Nair, Manchanda, and Bhatia (2006) have again found evidence of such effects after more carefully controlling for the marketing and other characteristics in a much larger data set.

an individual's body mass index in reaction to the obesity of someone that the individual named as a friend compared to someone who had named the individual as a friend. This is one method of sorting out causation, since if unobserved influences that were common to the agents were at work, then the direction of who mentioned the other as a friend would not matter, whereas direction would matter if it indicated which individual's react to which others. Based on this analysis, Christakis and Fowler conclude that obesity spreads very much like an epidemic with the underlying social structure appearing to play an important role.

It is worth emphasizing that even with longitudinal studies, one still has to be cautious in drawing causal inferences. The problem of homophily still looms, as linked individuals tend to have common characteristics and so may be influenced by common unobserved factors, for example, both being exposed to some external stimulus (such as advertising) at the same time. This then makes it appear as if one agent's behavior closely followed another's, even when it may simply be due to both having experienced some external event that prompted their behaviors. Aral, Muchnik, and Sundararajan (2009) provide an idea of how large this effect can be, by carefully tracking individual characteristics and then using propensity scores (likelihoods of having neighbors with certain behaviors) to illustrate the extent to which one can over-estimate diffusion effects by not accounting for common backgrounds of connected individuals.

Homophily not only suggests that linked individuals might be exposed to common influences, it also makes it hard to disentangle which of following two processes is at the root of observed similarities in behavior between connected agents. It could be that similar behavior in fact comes from a process of selection (assortative pairing), in which similarity precedes association. Alternatively, it could be a consequence of a process of socialization, in which association leads to similarity. In that respect, tracking connections and behaviors over time is particularly useful. Kandel (1978) concentrated on adolescent friendship pairs and examined the levels of homophily on four attributes (frequency of current marijuana use, level of educational aspirations, political orientation, and participation in minor delinquency) at

various stages of friendship formation and dissolution. She noted that observed homophily in friendship dyads resulted from a significant combination of both types of processes, so that individuals emulated their friends, but also tended to drop friendships with those more different from themselves and add new friendships to those more similar to themselves.[7]

In summary, let us mention a few of the important conclusions obtained from studies of diffusion. First, not only are behaviors across socially connected individuals correlated, but individuals do influence each other. While this may sound straightforward, it takes careful control to ensure that it is not unobserved correlated traits or influences that lead to similar actions by connected individuals, as well as an analysis of similarities between friends that can lead to correlations in their preferences and the things that influence them. Second, in various settings, more socially connected individuals adopt new behaviors and products earlier and at higher rates. Third, diffusion exhibits specific patterns over time, and specifically there are many settings where an "S"-shaped pattern emerges, with adoption starting slowly, then accelerating, and eventually asymptoting. Fourth, many diffusion processes are affected by the specifics of the patterns of the interaction.

3 Models of Diffusion and Strategic Interaction Absent Network Structure

We now turn to discussing various models of diffusion. As should be clear from our description of the empirical work on diffusion and behavior, models can help greatly in clarifying the tensions at play. Given the issues associated with the endogeneity of social relationships, and the substantial homophily that may lead to correlated behaviors among social neighbors, it is critical to have models that help predict how behavior should evolve and how it interacts with the social structure in place.

We start with some of the early models that do not account for the underlying network

---

[7]Of course there is also homophily based on non-malleable attributes, in which case homophily can only be due to the connection process. For example, Goeree, McConnell, Mitchell, Tromp, and Yariv (2010) observe homophily on height, and there is a rich literature on homophily based on ethnicity, gender, and other non-malleable attributes (see Jackson, Chapter ? in this volume, for references).

architecture per-se. These models incorporate the empirical observations regarding social influence through the particular dynamics assumed, or preferences posited, and generate predictions matching the *aggregate* empirical observations regarding diffusion over time of products, diseases, or behavior. For example, the so-called S-shaped adoption curves. After describing these models, we return to explicitly capturing the role of social networks.

3.1   Marketing Models

One of the earliest and still widely used models of diffusion is the Bass (1969) Model. This is a parsimonious model, which can be thought of as a "macro" model: it makes predictions about aggregate behavior in terms of the percentage of potential adopters of a product or behavior who will have adopted by a given time. The current rate of change of adoption depends on the current level and two critical parameters. These two parameters are linked to the rate at which people innovate or adopt on their own, and the rate at which they imitate or adopt because others have, thereby putting into (theoretical) force the empirical observation regarding peers' influence.

If we let $G(t)$ be the percentage of agents who have adopted by time t, and m be the fraction of agents in the population who are potential adopters, a discrete time version of the Bass model is characterized by the difference equation

$$G(t) = G(t-1) + p\left(m - G(t-1)\right) + q\left(m - G(t-1)\right)\frac{G(t-1)}{m},$$

where p is a rate of innovation and q is a rate of imitation. To glean some intuition, note that the expression $p\left(m - G(t-1)\right)$ represents the fraction of people who have not yet adopted and might potentially do so times the rate of spontaneous adoption. In the expression $q\left(m - G(t-1)\right)\frac{G(t-1)}{m}$, the rate of imitation is multiplied by two factors. The first factor, $(m - G(t-1))$, is the fraction of people who have not yet adopted and may still do so. The second expression, $\frac{G(t-1)}{m}$, is the relative fraction of potential adopters who are around to imitate. If we set m equal to 1, and look at a continuous time version of the above difference

equation, we get

$$g(t) = (p + qG(t))(1 - G(t)),$$ (1)

where g is the rate of diffusion (times the rate of change of G). Solving this when $p > 0$ and setting the initial set of adopters at 0, $G(0) = 0$, leads to the following expression:

$$G(t) = \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}}.$$

This is a fairly flexible formula that works well at fitting time series data of innovations. By estimating p and q from existing data, one can also make forecasts of future diffusion. It has been used extensively in marketing and for the general analysis of diffusion (e.g., Rogers, 1995), and has spawned many extensions and variations.[8]

If q is large enough,[9] then there is a sufficient imitation/social effect, which means that the rate of adoption accelerates after it begins, and so $G(t)$ is S-shaped (see Figure 1), matching one of the main insights of the longitudinal empirical studies on diffusion discussed above. The Bass model provides a clear intuition for why adoption curves would be S-shaped. Indeed, when the adoption process begins, imitation plays a minor role (relative to innovation) since not many agents have adopted yet and so the volume of adopters grows slowly. As the number of adopters increases, the process starts to accelerate as now innovators are joined by imitators. The process eventually starts to slow down, in part simply because there are fewer agents left to adopt (the term $1 - G(t)$ in (1) eventually becomes small). Thus, we see a process that starts out slowly, then accelerates, and then eventually slows and asymptotes.

3.2 Collective Action, Fashion, and Fads

The Bass model described above is mechanical in that adopters and imitators are randomly determined; they do not choose actions strategically. The empirical observation that individuals influence each other through social contact can be derived through agents' preferences, rather than through some exogenously specified dynamics.

---

[8] For some recent models, see Leskovec, Adamic, and Huberman (2007) and Young (2010).
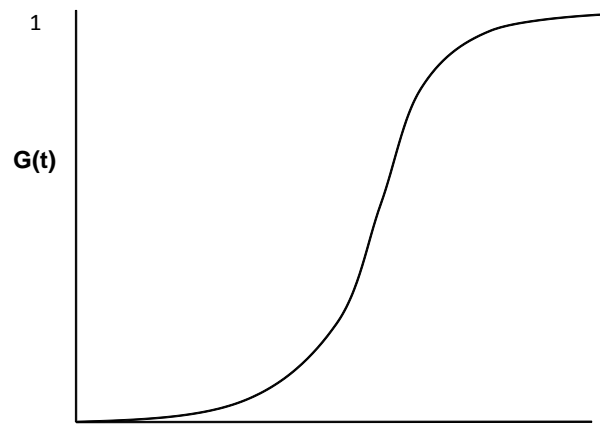[9] See Jackson (2008) for a more detailed discussion.

Figure 1: S-shape Adoption

Diffusion in a strategic context was first studied without a specific structure for interactions. Broadly speaking, there were two approaches taken in this early literature. In the first, all agents are connected to one another (that is, they form a complete network). Effectively, this corresponds to a standard multi-agent game in which payoffs to each player depend on the entire profile of actions played in the population. The second approach has been to look at interactions in which agents are matched to partners in a random fashion.

Diffusion on Complete Networks. Granovetter (1978) considered a model in which N agents are all connected to one another and each agent chooses one of two actions: 0 or 1. Associated with each agent i is a number n_i. This is a threshold such that if at least n_i other agents take action 1 then i prefers action 1 to action 0, and if fewer than n_i other agents take action 1 then agent i prefers to take action 0. The game has what are known as strategic complementarities. For instance, suppose that the utility of agent i faced with a profile of actions $(x_1, ..., x_N) \in \{0,1\}^N$ is described by:

$$u_i(x_1, ..., x_N) = \left[\frac{\sum_{j \neq i} x_j}{N-1} - c_i \right] x_i, \tag{2}$$

where c_i is randomly drawn from a distribution F over $[0,1]$. c_i can be thought of as a cost that agent i experiences upon choosing action 1 (e.g., a one time switching cost from one technology to the other, or potential time costs of joining a political revolt, etc.). The utility of agent i is normalized to 0 when choosing the action 0. When choosing the action 1, agent i experiences a benefit proportional to the fraction of other agents choosing the action 1 and a cost of c_i.

Granovetter considered a dynamic model in which at each stage agents best respond to the previous period's distribution of actions. If in period t there was a fraction x^t of agents choosing the action 1, then in period $t+1$ an agent i chooses action 1 if and only if his or her cost is lower than $\frac{Nx^t - x_i^t}{N-1}$, the fraction of other agents taking action 1 in the last period. For a large population, $\frac{Nx^t - x_i^t}{N-1} \simeq x^t$ and $x^{t+1} \simeq F(x^t)$. A fixed point $x^* = F(x^*)$ then corresponds to an (approximate) equilibrium of a large population.

The shape of the distribution F determines which equilibria are *tipping points*: equilibria

such that only a slight addition to the fraction of agents choosing the action 1 shifts the population, under the best response dynamics, to the next higher equilibrium level of adoption (we return to a discussion of tipping and stable points when we consider a more general model of strategic interactions on networks below).

Note that while in the Bass model the diffusion path was determined by $G(t)$, the fraction of adopters as a function of time, here it is easier to work with $F(x)$, corresponding to the fraction of adopters as a function of the previous period's fraction x.

Although Granovetter (1978) does not examine conditions under which the time series will exhibit attributes like the S-shape that we discussed above, by using techniques from Jackson and Yariv (2007) we can derive such results, as we now discuss. Keeping track of time in discrete periods (a continuous time analog is straightforward), the level of change of adoption in the society is given by

$$\Delta(x^t) = F(x^t) - x^t.$$

Thus, to derive an S-shape, we need this quantity to initially be increasing, and then eventually to decrease. Assuming differentiability of F, this corresponds to the derivative of $\Delta(x^t)$ being positive up to some $\bar{x}$ and then negative. The derivative of $F(x) - x$ is $F'(x) - 1$ and having an S-shape corresponds to F' being greater than 1 up to some point and then less than 1 beyond that point. For instance, if F is concave with an initial slope greater than 1 and an eventual slope less than 1, this is satisfied. Note that the S-shape of adoption over time does not translate into an S-shape of F - but rather a sort of concavity.[10] The idea is that we initially need a rapid level of change, which corresponds to an initially high slope of F, and then a slowing down, which corresponds to a lower slope of F.

Fashions and Random Matching.   A different approach than that of the Bass model is taken by Pesendorfer (1995), who considers a model in which individuals are randomly matched and new fashions serve as signaling instruments for the creation of matches. He

---

[10]Concavity, plus having a slope that is 1 at some point, is sufficient, but not necessary to have the positive and then negative property of $F'(x) - 1$.

identifies particular matching technologies that generate fashion cycles. Pesendorfer describes the spread of a new fashion as well as its decay over time. In Pesendorfer's model, the price of the design falls as it spreads across the population. Once sufficiently many consumers own the design, it is profitable to create a new design and thereby render the old design obsolete. In particular, demand for any new innovation eventually levels off as in the above two models.

Information Cascades and Learning. Another influence on collective behavior derives from social learning. This can happen without any direct complementarities in actions, but due to information flow about the potential payoffs from different behaviors. If people discuss which products are worth buying, or which technologies are worth adopting, books worth reading, and so forth, even without any complementarities in behaviors, one can end up with cascades in behavior, as people infer information from others' behaviors and can (rationally) imitate them. As effects along these lines are discussed at some length in Chamley (Chapter ?, this volume), Jackson (Chapter ?, this volume), and Goyal (Chapter ?, this volume), we will not detail them here. We only stress that pure information transmission can lead to diffusion of behaviors.

4 Models of Diffusion and Strategic Interaction in Network Settings

We now turn to models that explicitly incorporate social structure in examining diffusion patterns. We start with models that stem mostly from the epidemiology literature and account for the underlying social network, but are mechanical in terms of the way that disease spreads from one individual to another (much like the Bass model described above). We then proceed to models in which players make choices that depend on their neighbors' actions as embedded in a social network; for instance, only adopting an action if a certain proportion of neighbors adopt as well (as in Granovetter's setup), or possibly not adopting an action if enough neighbors do so.

4.1   A Unified Setting

Many models of diffusion and strategic interaction on networks have the following common elements.

There is a finite set of agents $N = \{1, \ldots, n\}$.

Agents are connected by a (possibly directed) network $g \in \{0,1\}^{n \times n}$. We let $N_i(g) \equiv \{j : g_{ij} = 1\}$ be the neighbors of i. The degree of a node i is the number of her neighbors, $d_i \equiv |N_i(g)|$.

When links are determined through some random process, it is often useful to summarize the process by the resulting distribution of degrees P, where $P(d)$ denotes the probability a random individual has a degree of d.[11, 12]

Each agent $i \in N$ takes an action x_i. In order to unify and simplify the description of various models, we focus on binary actions, so that $x_i \in \{0,1\}$. Actions can be metaphors for becoming "infected" or not, buying a new product or not, choosing one of two activities, and so forth.

4.2   Epidemiology Models

4.2.1   Random Graph Models

Some basic insights about the extent to which behavior or an infection can spread in a society can be derived from random graph theory. Random graph theory provides a tractable base for understanding characteristics important for diffusion, such as the structure and size of the components of a network, maximally connected subnetworks.[13]

---

[11]Such a description is not complete, in that it does not specify the potential correlations between degrees of different individuals on the network. See Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv (2010) for more details.

[12]In principle, one would want to calibrate degree distributions with actual data. The literature on network formation, see Bloch and Dutta (Chapter ?, this volume) and Jackson (Chapter ?, this volume), suggests some insights on plausible degree distributions $P(d)$.

[13]Formally, these are the subnetworks projected induced by maximal sets $C \subseteq N$ of nodes such any two distinct nodes in C are *path connected* within C. That is, for any $i, j \in C$, there exist $i_1, \ldots, i_k \in C$ such that $g_{ii_1} = g_{i_1 i_2} = \ldots = g_{i_{k-1} i_k} = g_{i_k j} = 1$.

Before presenting some results, let us talk through some of the ideas in the context of what is known as the Reed-Frost model.[14] Consider, for example, the spread of a disease. Initially, some individuals in the society get infected through mutations of a germ or other exogenous sources. Consequently, some of these individuals' neighbors are infected through contact, while others are not. This depends on how virulent the disease is, among other things. In this application, it makes sense (at least as a starting point) to assume that becoming infected or avoiding infection is not a choice; i.e., contagion here is non-strategic. In the simplest model, there is a probability $\pi \geq 0$ that a given individual is immune (e.g., through vaccination or natural defenses). If an individual is not immune, it is assumed that he or she is sure to catch the disease if one of his or her neighbors ends up with the disease. In this case, in order to estimate the volume of those ultimately infected, we proceed in two steps, depicted in Figure 2. First, we delete a fraction π of the nodes that will never be infected (these correspond to the dotted nodes in the Figure). Then, we note that the components of the remaining network that contain the originally infected individuals comprise the full extent of the infection. In particular, if we can characterize what the components of the network look like after removing some portion of the nodes, we have an idea of the extent of the infection. In Figure 2 we start with one large connected component (circumvented by a dotted line) and two small connected components. After removing the immune agents, there is still a large connected component (though smaller than before), and four small components.

Thus, the estimation of the extent of infection of the society is reduced to the estimation of the component structure of the network. A starting point for the formal analysis of this sort of model uses the canonical random network model, where links are formed independently, each with an identical probability $p > 0$ of being present. This is sometimes referred to as a "Poisson random network" as its degree distribution is approximated by a Poisson distribution if p is not excessively large; and has various other aliases such as an "Erdös-Renyi random graph," a "Bernoulli random graph," or a "$G(n,p)$" random graph (see Jackson, Chapter ? in this volume, for more background). Ultimately, the analysis boils down to

---

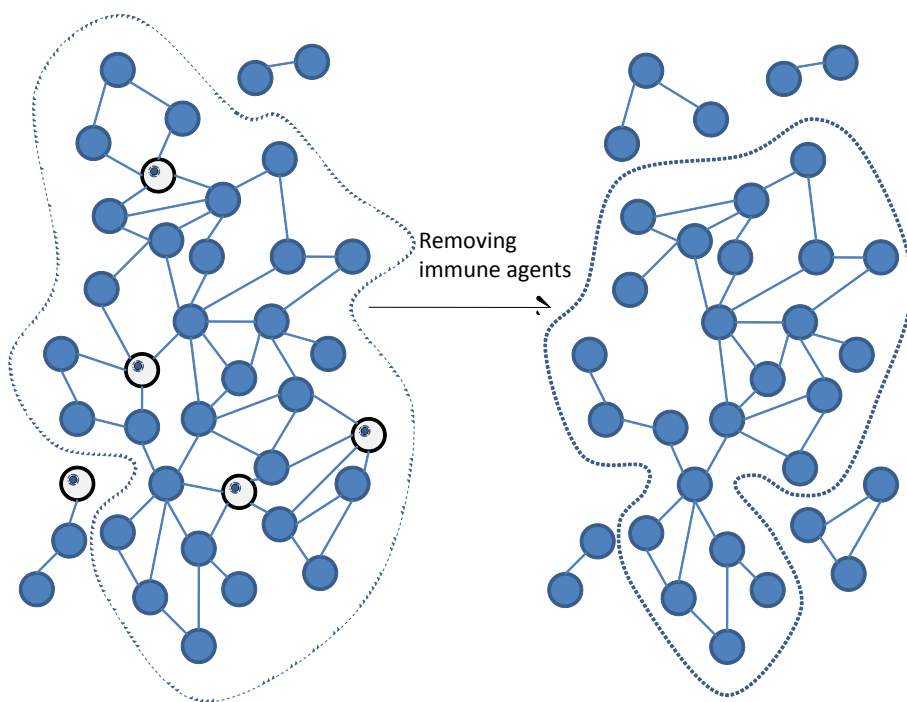[14]See Jackson (2008) for a more detailed discussion of this and related models.

Figure 2: Network Components and Immune Agents

considering a network on $(1 - \pi)n$ nodes with an independent link probability of p, and then measuring the size of the component containing a randomly chosen initially infected node.

Clearly, with a fixed set of nodes, and a positive probability p that lies strictly between 0 and 1, every conceivable network on the given set of nodes could arise. Thus, in order to say something specific about the properties of the networks that are "most likely" to arise, one generally works with large n where reasoning based on laws of large numbers can be employed. For example, if we think of letting n grow, we can ask for which p's (that are now dependent on n) a non-vanishing fraction of nodes will become infected with a probability bounded away from 0. So, let us consider a sequence of societies indexed by n and corresponding probabilities of links $p(n)$.

Erdös and Renyi (1959, 1960) proved a series of results that characterize some basic properties of such random graphs. In particular,[15]

- The threshold for the existence of a "giant component," a component that contains a nontrivial fraction of the population is $1/n$, corresponding to an average degree of 1. That is, if $p(n)$ over $1/n$ tends to infinity, then the probability of having a giant component tends to 1, while if $p(n)$ over $1/n$ tends to 0, then the probability of having a giant component tends to 0.

- The threshold for the network to be connected (so that every two nodes have a path between them) is $\log(n)/n$, corresponding to an average degree that is proportional to $\log(n)$.

The logic for the first threshold is easy to explain, though the proof is rather involved. To heuristically derive the threshold for the emergence of a giant component, consider following a link out of a given node. We ask whether or not one would expect to be able to find a link to another node from that one. If the expected degree is much smaller than 1, then following the few (if any) links from any given node is likely to lead to dead-ends. In contrast, when

---

[15]See Chapter 4 in Jackson (2008) for a fuller discussion and proofs of these results.

the expected degree is much higher than 1, then from any given node, one expects to be able to reach more nodes, and then even more nodes, and so forth, and so the component should expand outward.

Note that adjusting for the factor π of the number of immune nodes does not affect the above thresholds as they apply as limiting results, although the factor will be important for any fixed n.

Between these two thresholds, there is only one giant component, so that the next largest component is of a size that is a vanishing fraction of the giant component. This is intuitively clear, as to have two large components requires many links within each component but no links between the two components, which is an unlikely event. In that sense, the image that emerges from Figure 2 of one large connected component is reasonably typical for a range of parameter values.

These results then tell us that in a random network, if average degree is quite low (smaller than 1), then any initial infection is likely to die out. In contrast, if average degree is quite high (larger than $\log(n)$), then any initial infection is likely to spread to all of the susceptible individuals, i.e., a fraction of $1 - \pi$ of the population. In the intermediate range, there is a probability that the infection will die out and also a probability that it will infect a nontrivial, but limited, portion of the susceptible population. There, it can be shown that for such random networks and large n, the fraction of nodes in the giant component of susceptible nodes is roughly approximated by the nonzero q that solves

$$q = 1 - e^{-q(1-\pi)np}. \tag{3}$$

Here, q is an approximation of the probability of the infection spreading to a non-trivial fraction of nodes, and also of the percentage of susceptible nodes that would be infected.[16]

This provides a rough idea of the type of results that can be derived from random graph theory. There is much more that is known, as one can work with other models of random graphs (other than ones where each link has an identical probability), richer models of

---

[16] Again, see Chapter 4 in Jackson (2008) for more details.

probabilistic infection between nodes, as well as derive more information about the potential distribution of infected individuals.

It should also be emphasized that while the discussion here is in terms of "infection," the applications clearly extend to many of the other contexts we have been mentioning, such as the transmission of ideas and information. Fuller treatment of behaviors, where individual decisions depend in more complicated ways on neighbors' decisions, are treated in Section 4.3.

4.2.2 Diffusion with Recovery

The above analysis of diffusion presumes that once infected, a node eventually infects all of its susceptible neighbors. This misses important aspects of many applications. In terms of diseases, infected nodes can either recover and stop transmitting a disease, or die and completely disappear from the network. Transmission will also generally be probabilistic, depending on the type of interaction and its extent.[17] Similarly, if we think of behaviors, it might be that the likelihood that a node is still actively transmitting a bit of information to its neighbors decreases over time.

Ultimately, we will discuss models that allow for rather general strategic impact of peer behavior (a generalization of the approach taken by Granovetter). But first we discuss some aspects of the epidemiology literature that takes steps forward in that direction by considering two alternative models that keep track of the state of nodes and are more explicitly dynamic. The common terminology for the possible states that a node can be in are: *susceptible*, where a node is not currently infected or transmitting a disease but can catch it; *infected*, where a node has a disease and can transmit it to its neighbors; and *removed* (or *recovered*) where a node has been infected but is no longer able to transmit the disease and cannot be re-infected.

The first of the leading models is the "SIR" model (dating to Kermack and McKendrick,

---

[17]Probabistic transmission is easily handled in the above model by simply adjusting the link probability to reflect the fact that some links might not transmit the disease.

1927), where nodes are initially susceptible but can catch the disease from infected neighbors. Once infected, a node continues to infect neighbors until it is randomly removed from the system. This fits well the biology of some childhood diseases, such as the chicken pox, where one can only be infected once.

The other model is the "SIS" model (see Bailey, 1975), where once infected, nodes can randomly recover, but then they are susceptible again. This corresponds well with an assortment of bacterial infections, viruses, and flus, where one transitions back and forth between health and illness.

The analysis of the SIR model is a variant of the component-size analysis discussed above. The idea is that there is a random chance that an "infected" node infects a given "susceptible" neighbor before becoming "removed." Roughly, one examines component structures where instead of removing *nodes* randomly, one removes *links* randomly from the network. This results in variations on the above sorts of calculations, where there are adjusted thresholds for infection depending on the relative rates of how quickly infected nodes can infect their neighbors compared to how quickly they are removed.

In contrast, the SIS model involves a different sort of analysis. The canonical version of that model is best viewed as one with a random matching process rather than a social network. In particular, suppose that a node i in each period will have interactions with d_i other individuals from the population. Recall our notation of $P(d)$ describing the proportion of the population that has degree d (so d interactions per period). The matches are determined randomly, in such a way that if i is matched with j, then the probability that j has degree $d > 0$ is given by

$$\tilde{P}(d) = \frac{P(d)d}{\langle d \rangle}, \tag{4}$$

where $\langle \cdot \rangle$ represents the expectation with respect to P.[18] This reflects the fact that an agent is proportionally more likely to be matched with other individuals who have lots of connections. To justify this formally, one needs an infinite population. Indeed, with any

---

[18] We consider only individuals who have degree $d > 0$, as others do not participate in the society.

finite population of agents with heterogeneous degrees, the emergent networks will generally exhibit some correlation between neighbors' degrees.[19]

Individuals who have high degrees will have more interactions per period and will generally be more likely to be infected at any given time. An important calculation then pertains to the chance that a given meeting will be with an infected individual. If the infection rate of degree d individuals is $\rho(d)$, then the probability that any given meeting will be with an infected individual is θ, where

$$\theta = \sum_d \tilde{P}(d)\rho(d) = \frac{\sum P(d)\rho(d)d}{\langle d \rangle}. \tag{5}$$

The chance of meeting an infected individual in a given encounter then differs from the average infection rate in the population, which is just $\rho = \sum P(d)\rho(d)$, because θ is weighted by the rate at which individuals meet each other.

A standard version of contagion that is commonly analyzed is one where the probability of an agent of degree d becoming infected is

$$\nu\theta d, \tag{6}$$

where $\nu \in (0, 1)$ is a rate of transmission of infection in a given period, and is small enough so that this probability is less than one. If ν is very small, this is an approximation of getting infected under d interactions with each having an (independent) probability θ of being infected and then conditionally (and independently) having a probability ν of getting infected through contact with a given infected individual. The last part of the model is that in any given period, an infected individual recovers and becomes susceptible with a probability $\delta \in (0, 1)$.

If such a system operates on a finite population, then eventually all agents will become susceptible and that would end the infection. If there is a small probability of a new mutation and infection in any given period, the system will be ergodic and always have some probability of future infection.

---

[19]See the appendix of Currarini, Jackson, and Pin (2009) for some details along this line.

To get a feeling for the long run outcomes in large societies, the literature has examined a steady state (i.e., a situation in which the system essentially remains constant) of a process that is idealized as operating on an infinite (continuous) population. Formally, a steady-state is defined by having $\rho(d)$ be constant over time for each d. Working with an approximation at the limit (termed a "mean-field" approximation that in this case can be justified with a continuum of agents, but with quite a bit of technical detail), a steady-state condition can be derived to be

$$0 = (1 - \rho(d))\nu\theta d - \rho(d)\delta \qquad (7)$$

for each d. $(1 - \rho(d))\nu\theta d$ is the rate at which agents of degree d who were susceptible become infected and $\rho(d)\delta$ is the rate at which infected individuals of degree d recover. Letting $\lambda = \frac{\nu}{\delta}$, it follows that

$$\rho(d) = \frac{\lambda\theta d}{\lambda\theta d + 1}. \qquad (8)$$

Solving (5) and (8) simultaneously leads to a characterization of the steady-state θ:

$$\theta = \sum_d \frac{P(d)\lambda\theta d^2}{\langle d \rangle (\lambda\theta d + 1)}. \qquad (9)$$

This system always has a solution, and therefore a steady-state, where $\theta = 0$ so there is no infection. It can also have other solutions under which θ is positive (but always below 1 if λ is finite). Unless P takes very specific forms, it can be difficult to find steady states $\theta > 0$ analytically.

Special cases have been analyzed, such as the case of a power distribution, where $P(d) = 2d^{-3}$ (e.g., see Pastor-Satorras and Vespignani, 2000, 2001). In that case, there is always a positive steady-state infection rate. More generally, Lopez-Pintado (2008) addresses the question of when it is that there will be a positive steady-state infection rate. To get some intuition for her results, let

$$H(\theta) = \sum \frac{P(d)d}{\langle d \rangle} \left(\frac{\lambda d\theta}{\lambda d\theta + 1} \right) = \sum \tilde{P}(d) \left(\frac{\lambda d\theta}{\lambda d\theta + 1} \right), \qquad (10)$$

so that the equation $\theta = H(\theta)$ corresponds to steady states of the system. We can now extend the analysis of Granovetter's (1978) model that we described above, with this richer model

in which $H(\theta)$ accounts for network attributes. While the fixed point equation identifying Granovetter's stable points allowed for rather arbitrary diffusion patterns (depending on the cost distribution F), the function H has additional structure to it that we can explore.

In particular, suppose we examine the infection rate that would result if we start at a rate of θ and then run the system on an infinite population for one period. Noting that $H(0) = 0$, it is clear that 0 is always a fixed point and thus a steady-state. Since $H(1) < 1$, and H is increasing and strictly concave in θ (which is seen by examining its first and second derivatives), there can be at most one fixed point besides 0. For there to be another fixed point (steady-state) above $\theta = 0$, it must be that $H'(0)$ is above 1, or else given the strict concavity, we would have $H(\theta) < \theta$ for all positive θ. Moreover, in cases where $H'(0) > 1$, a small perturbation away from a 0 infection rate will lead to increased infection. In the terminology we have introduced above, 0 would be a *tipping point*. Since

$$H'(0) = \lambda \frac{\langle d^2 \rangle}{\langle d \rangle}, \tag{11}$$

we have a simple way of checking whether we expect a positive steady-state infection or a 0 steady-state infection. This simply boils down to a comparison of the relative infection rate λ and $\frac{\langle d \rangle}{\langle d^2 \rangle}$ so that there is a positive infection rate if and only if

$$\lambda > \frac{\langle d \rangle}{\langle d^2 \rangle}. \tag{12}$$

Higher infection rates lead to the possibility of positive infection, as do degree distributions with high variances (relative to mean). The idea behind having a high variance is that there will be some "hub nodes" with high degree, who can foster contagion.

Going back to our empirical insights, this analysis fits the observations that highly-linked individuals are more likely to get infected and experience speedier diffusion. Whether the aggregate behavior exhibits the S-shape that is common in many real-world diffusion processes will depend on the particulars of H, much in the same way that we discussed how the S-shape in Granovetter's model depends on the shape of the distribution of costs F in that model. Here, things are slightly complicated since H is a function of θ, which is

the probability of infection of a neighbor, and not the overall probability of infection of the population. Thus, one needs to further translate how various θ's over time translate into population fractions that are infected.

Beyond the extant empirical studies, this analysis provides some intuitions behind what is needed for an infection to be possible. It does not, however, provide an idea of how extensive the infection spread will be and how that depends on network structure. While this does not boil down to as simple a comparison as (12), there is still much that can be deduced using (9), as shown by Jackson and Rogers (2007). While one cannot always directly solve

$$\theta = \sum_d \frac{P(d)\lambda\theta d^2}{\langle d \rangle \, (\lambda\theta d + 1)},$$

notice that $\frac{\lambda\theta d^2}{\langle d \rangle (\lambda\theta d+1)}$ is an increasing and convex function of d. Therefore, the right hand side of the above equality can be ordered when comparing different degree distributions in the sense of stochastic dominance (we will return to these sorts of comparisons in some of the models we discuss below). The interesting conclusion regarding steady-state infection rates is that they depend on network structure in ways that are very different at low levels of the infection rate λ compared to high levels.

4.3   Graphical Games

While the above models provide some ideas about how social structure impacts diffusion, they are limited to settings where, roughly speaking, the probability that a given individual adopts a behavior is simply proportional to the infection rate of neighbors. Especially when it comes to situations in which opinions or technologies are adopted, purchasing decisions are made, etc., an individual's decision can depend in much more complicated ways on the behavior of his or her neighbors. Such interaction naturally calls on game theory as a tool for modeling these richer interactions.

We start with static models of interactions on networks that allow for a rather general impact of peers' actions on one's own optimal choices.

The first model that explicitly examines games played on a network is the model of "graphical games" as introduced by Kearns, Littman, and Singh (2001), and analyzed by Kakade, Kearns, Langford, and Ortiz (2003), among others. The underlying premise in the graphical games model is that agents' *payoffs* depend on their own actions and the actions of their direct neighbors, as determined by the network of connections.[20]

Formally, the payoff structure underlying a graphical game is as follows. The payoff to each player i when the profile of actions is $x = (x_1, ..., x_n)$ is

$$u_i(x_i, x_{N_i(g)}),$$

where $x_{N_i(g)}$ is the profile of actions taken by the neighbors of i in the network g.

Most of the empirical applications discussed earlier entailed agents responding to neighbors' actions in roughly one of two ways. In some contexts, such as those pertaining to the adoption of a new product or new agricultural grain, decisions to join the workforce, or to join a criminal network, agents' conceivably gain more from a particular action the greater is the volume of peers who choose a similar action. That is, payoffs exhibit strategic complementarities. In other contexts, such as experimentation on a new drug, or contribution to a public good, when an agent's neighbors choose a particular action, the relative payoff the agent gains from choosing a similar action decreases, and there is strategic substitutability. The graphical games environment allows for the analysis of both types of setups, as the following example (taken from Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv, 2010) illustrates.

Example 1 (Payoffs Depend on the Sum of Actions) Player i's payoff function when he or she chooses x_i and her k neighbors choose the profile $(x_1, ..., x_k)$ is:

$$u_i(x_i, x_1, ..., x_k) = f\left(x_i + \lambda \sum_{j=1}^{k} x_j\right) - c(x_i), \tag{13}$$

---

[20]There are also other models of equilibria in social interactions, where players care about the play of certain other groups of players. See Glaeser and Scheinkman (2000) for an overview.

where $f(\cdot)$ is non-decreasing and $c(\cdot)$ is a "cost" function associated with own effort (more general but much in the spirit of (2)). The parameter $\lambda \in \mathbb{R}$ determines the nature of the externality across players' actions. The shape and sign of λf determine the effects of neighbors' action choices on one's own optimal choice. In particular, the example yields strict strategic substitutes (complements) if, assuming differentiability, $\lambda f''$ is negative (positive).

There are several papers that analyze graphical games for particular choices of f and λ. To mention a few examples, the case where f is concave, $\lambda = 1$, and $c(\cdot)$ is increasing and linear corresponds to the case of information sharing as a local public good studied by Bramoullé and Kranton (2007), where actions are strategic substitutes. In contrast, if $\lambda = 1$, but f is convex (with $c'' > f'' > 0$), we obtain a model with strategic complements, as proposed by Goyal and Moraga-Gonzalez (2001) to study collaboration among local monopolies. In fact, the formulation in (13) is general enough to accommodate numerous further examples in the literature such as human capital investment (Calvó-Armengol and Jackson, 2009), crime and other networks (Ballester, Calvó-Armengol, and Zenou, 2006), some coordination problems (Ellison, 1993), and the onset of social unrest (Chwe, 2000).

The computer science literature (e.g., the literature following Kearns, Littman, and Singh, 2001, and analyzed by Kakade, Kearns, Langford, and Ortiz, 2003) has focused predominantly on the question of when an efficient (polynomial-time) algorithm can be provided to compute Nash equilibria of graphical games. It has not had much to say about the properties of equilibria, which is important when thinking about applying such models to analyze diffusion in the presence of strategic interaction. In contrast, the economics literature has concentrated on characterizing equilibrium outcomes for particular applications, and deriving general comparative statics with respect to agents' positions in a network and with respect to the network architecture itself.

Information players hold regarding the underlying network (namely, whether they are fully informed of the entire set of connections in the population, or only of connections in some local neighborhood) ends up playing a crucial role in the scope of predictions generated by network games models. Importantly, graphical games are ones in which agents have complete information regarding the networks in place. Consequently, such models suffer from inherent multiplicity problems, as clearly illustrated in the following example. It is based on a variation of (13), which is similar to a model analyzed by Bramoullé and Kranton (2007).

Example 2 (Multiplicity – Complete Information) Suppose that in (13), we set $\lambda = 1$, choose $x_i \in \{0,1\}$ and have

$$f\left(x_i + \sum_{j=1}^{k} x_j\right) \equiv \min\left[x_i + \sum_{j=1}^{k} x_j, 1\right],$$

and $c(x_i) \equiv cx_i$, where $0 < c < 1$. This game, often labeled the *best-shot public goods game*, may be viewed as a game of local public-good provision. Each agent would choose the action 1 (say, experimenting with a new grain, or buying a product that can be shared with one's friends) if they were alone (or no one else experimented), but would prefer that one of their neighbors incur the cost c that the action 1 entails (when experimentation is observed publicly). Effectively, an agent just needs at least one agent in his or her neighborhood to take action 1 to enjoy its full benefits, but prefers that it be someone else given that the action is costly and there is no additional benefit beyond one person taking the action.

Note that, since $c < 1$, in any Nash equilibrium, for any player i with k neighbors, it must be the case that one of the agents in the neighborhood chooses the action 1. That is, if the chosen profile is $(x_1, ..., x_k)$, then $x_i + \sum_{j=1}^{k} x_j \geqslant 1$. In fact, there is a very rich set of equilibria in this game. To see this, consider a star network and note that there exist two equilibria, one in which the center chooses 0 and the spokes choose 1, and a second equilibrium in which the spoke players choose 0 while the center chooses
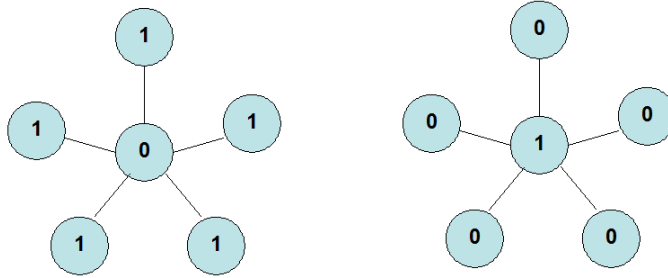
Figure 3: Multiplicity of Equilibria with Complete Information

1. Figure 3 illustrates these two equilibria. In the first, depicted in the left panel of the Figure, the center earns more than the spoke players, while in the second equilibrium (in the right panel) it is the other way round.

Even in the simplest network structures equilibrium multiplicity may arise and the relation between network architecture, equilibrium actions, and systematic patterns can be difficult to discover.

4.4   Network Games

While complete information regarding the structure of the social network imposed in graphical game models may be very sensible when the relevant network of agents is small, in large groups of agents (such as a country's electorate, the entire set of corn growers in the 50's, sites in the world-wide web, or academic economists), it is often the case that individuals have noisy perceptions of their network's architecture. As the discussion above stressed,

complete information poses many challenges because of the widespread occurrence of equilibrium multiplicity that accompanies it. In contrast, when one looks at another benchmark, where agents know how many neighbors they will have but not whom they will be, then the equilibrium correspondence is much easier to deal with. Moreover, this benchmark is an idealized model of settings in which agents make choices like learning a language or adopting a technology that they will use over a long time. In such contexts, agents have some idea of how many interactions they are likely to have in the future, but not exactly with whom the interactions will be.

A *network game* is a modification of a graphical game in which agents can have private and incomplete information regarding the realized social network at place. We describe here the setup corresponding to that analyzed by Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv (2010) and Jackson and Yariv (2005, 2007), restricting attention to binary action games.[21]

Uncertainty is operationalized by assuming the network is determined according to some random process yielding our distribution over agents' degrees, $P(d)$, which is common knowledge. Each player i has d_i interactions, but does not know how many interactions each neighbor has. Thus, each player knows something about his or her local neighborhood (the number of direct neighbors), but only the distribution of links in the remaining population.

Consider now the following utility specification, a generalization of (2). Agent i has a cost of choosing 1, denoted c_i. Costs are randomly and independently distributed across the society, according to a distribution F^c. Normalize the utility from the action 0 to 0 and let the benefit of agent i from action 1 be denoted by $v(d_i, x)$, where d_i is $i's$ degree and she expects each of her neighbors to independently choose the action 1 with probability x. Agent i's added payoff from adopting behavior 1 over sticking to the action 0 is then $v(d_i, x) - c_i$.

This captures how the number of neighbors that i has, as well as their propensity to choose the action 1, affects the benefits from adopting 1. In particular, i prefers to choose

---

[21]There are also other variations, such as Galeotti and Vega-Redondo (2006) and Sundararajan (2007), who study specific contexts, compatible with particular utility specifications.

the action 1 if

$$c_i \leqslant v(d_i, x). \tag{14}$$

This is a simple cost-benefit analysis generalizing Granovetter (1978)'s setup in that benefits can now depend on one's own degree (so that the underlying network is accounted for). Let $F(d, x) \equiv F^c(v(d, x))$. In words, $F(d, x)$ is the probability that a random agent of degree d chooses the action 1 when anticipating that each neighbor will choose 1 with an independent probability x.

Note that $v(d, x)$ can encompass all sorts of social interactions. In particular, it allows for a simple generalization of Granovetter's (1978) model to situations in which agents' payoffs depend on the expected number of neighbors adopting, dx.

Existence of symmetric Bayesian equilibria follows standard arguments. In cases where v is non-decreasing in x for each d, it is a direct consequence of Tarski's Fixed Point Theorem. In fact, in this case, there exists an equilibrium in pure strategies. In other cases, provided v is continuous in x for each d, a fixed point can still be found by appealing to standard theorems (e.g., Kakutani) and admitting mixed strategies.[22]

Homogeneous Costs. Suppose first that all individuals experience the same cost $c > 0$ of choosing the action 1 (much like in Example 2 above). In that case, as long as $v(d, x)$ is monotonic in d (non-increasing or non-decreasing), equilibria are characterized by a threshold. Indeed, suppose $v(d, x)$ is increasing in d, then any equilibrium is characterized by a threshold d^* such that all agents of degree $d < d^*$ choose the action 0 and all agents of degree $d > d^*$ choose the action 1 (and agents of degree d^* may mix between the actions). In particular, notice that the type of multiplicity that appeared in Example 2 no longer occurs (provided degree distributions are not trivial). It is now possible to look at comparative statics of equilibrium behavior and outcomes using stochastic dominance arguments on the network itself. For ease of exposition, we illustrate this in the case of non-atomic costs (see

---

[22]In such a case, the best response correspondence (allowing mixed strategies) for any (d_i, c_i) as dependent on x is upper hemi-continuous and convex-valued. Taking expectations with respect to d_i and c_i, we also have a set of population best responses as dependent on x that is upper hemi-continuous and convex valued.

Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv, 2010, for the general analysis).

Heterogeneous Costs. Consider the case in which F^c is a continuous function, with no atoms. In this case, a simple equation is sufficient to characterize equilibria. Let x be the probability that a randomly chosen neighbor chooses the action 1. Then $F(d, x)$ is the probability that a random (best responding) neighbor of degree d chooses the action 1. We can now proceed in a way reminiscent of the analysis of the SIS model. Recall that $\widetilde{P}(d)$ denoted the probability that a random neighbor is of degree d (see equation (4)). It must be that

$$x = \phi(x) \equiv \sum_d \widetilde{P}(d) F(d, x). \tag{15}$$

Again, a fixed point equation captures much of what occurs in the game. In fact, equation (15) characterizes equilibria in the sense that any symmetric[23] equilibrium results in an x that satisfies the equation, and any x that satisfies the equation corresponds to an equilibrium where type (d_i, c_i) chooses 1 if and only if inequality (14) holds. Given that equilibria can be described by their corresponding x, we often refer to some value of x as being an "equilibrium."

Consider a symmetric equilibrium and a corresponding probability of x for a random neighbor to choose action 1. If the payoff function v is increasing in degree d, then the expected payoff of an agent with degree $d+1$ is $v(d+1, x) \geqslant v(d, x)$ and so $F^c(v(d+1, x)) \geqslant F^c(v(d, x))$ and agents with higher degrees choose 1 with weakly higher probabilities. Indeed, an agent of degree $d+1$ can imitate the decisions of an agent of degree d and gain at least as high a payoff. Thus, if v is increasing (or, in much the same way, decreasing) in d for each x, then any symmetric equilibrium entails agents with higher degrees choosing action 1 with weakly higher (lower) probability. Furthermore, agents of higher degree have higher (lower) expected payoffs.

Much as in the analysis of the epidemiological models, the multiplicity of equilibria is determined by the properties of ϕ, which, in turn, correspond to properties of $\widetilde{P}$ and F. For

---

[23]Symmetry indicates that agents with the same degree and costs follow similar actions.

instance,

- if $F(d, 0) > 0$ for some d in the support of P, and F is concave in x for each d, then there exists at most one fixed point, and

- if $F(d, 0) = 0$ for all d and F is strictly concave or strictly convex in x for each d, then there are at most two equilibria - one at 0, and possibly an additional one, depending on the slope of $\phi(x)$ at $x = 0$.[24]

In general, as long as the graph of $\phi(x)$ crosses the 45 degree line only once, there is a unique equilibrium (see Figure 4 below).[25]

The set of equilibria generated in such network games is divided into stable and unstable ones (those we have already termed in Section 3.2 as *tipping points*). The simple characterization given by (15) allows for a variety of comparative statics on fundamentals pertaining to either type of equilibrium. In what follows, we show how these comparative statics tie directly to a simple strategic diffusion process. Indeed, it turns out there is a very useful technical link between the static and dynamic analysis of strategic interactions on networks.

4.5   Adding Dynamics – Diffusion and Equilibria of Network Games

An early contribution to the study of diffusion of strategic behavior allowing for general network architectures was by Morris (2000).[26] Morris (2000) considered coordination games played on networks. His analysis pertained to identifying social structures conducive to

---

[24]As before, the slope needs to be greater than 1 for there to be an additional equilibrium in the case of strict concavity, while the case of strict convexity depends on the various values of $F(d, 1)$ across d.

[25]Morris and Shin (2003, 2005) consider uncertainty on payoffs rather than on an underlying network. In coordination games, they identify a class of payoff shocks that lead to a unique equilibrium. Heterogeneity in degrees combined with uncertainty plays a similar role in restricting the set of equilibria. In a sense, the analysis described here is a generalization in that it allows studying the impact of changes in a variety of fundamentals on the set of stable and unstable equilibria, regardless of multiplicity, in a rather rich environment. Moreover, the equilibrium structure can be tied to the network of underlying social interactions.

[26]One can find predecessors with regards to specific architectures, usually lattices or complete mixings, such as Conway's (1970) "game of life," and various agent-based models that followed such as the "voter model" (e.g., see Clifford and Sudbury (1973) and Holley and Liggett (1975)), as well as models of stochastic stability (e.g., Kandori, Mailath, Robb (1993), Young (1993), Ellison (1995)).

contagion, where a small fraction of the population choosing one action leads to that action spreading across the *entire* population. The main insight from Morris (2000) is that maximal contagion occurs when the society has certain sorts of cohesion properties, where there are no groups (among those not initially infected) that are too inward looking in terms of their connections.

In order to identify the full set of stable of equilibria using the above formalization, consider a diffusion process governed by best responses in discrete time (following Jackson and Yariv, 2005, 2007). At time $t = 0$, a fraction x^0 of the population is exogenously and randomly assigned the action 1, and the rest of the population is assigned the action 0. At each time $t > 0$, each agent, *including the agents assigned to action* 1 *at the outset*, best responds to the distribution of agents choosing the action 1 in period $t - 1$, accounting for the number of neighbors they have and presuming that their neighbors will be a random draw from the population.

Let x_d^t denote the fraction of those agents with degree d who have adopted behavior 1 at time t, and let x^t denote the link-weighted fraction of agents who have adopted the behavior at time t. That is, using the distribution of neighbors' degrees $\tilde{P}(d)$,

$$x^t = \sum_d \widetilde{P}(d) x_d^t.$$

Then, as deduced before from equation (14), at each date t,

$$x_d^t = F(d, x^{t-1}).$$

and therefore

$$x^t = \sum_d \widetilde{P}(d) F(d, x^{t-1}) = \phi(x^{t-1}).$$

As we have discussed, any rest point of the system corresponds to a static (Bayesian) equilibrium of the system.

If payoffs exhibit complementarities, then convergence of behavior from any starting point is monotone, either upwards or downwards. In particular, once an agent switches behaviors,
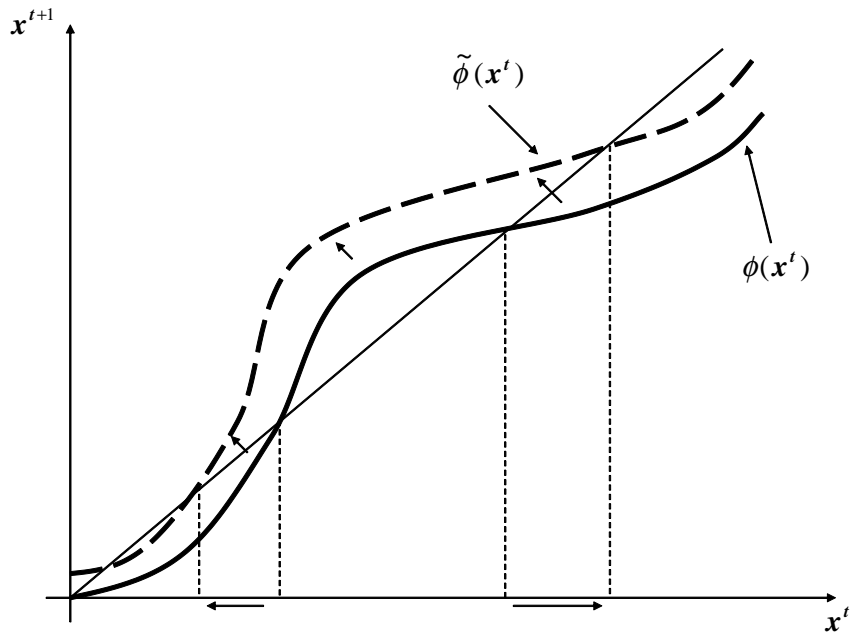
Figure 4: The Effects of Shifting $\phi(x)$ Pointwise

the agent will not want to switch back at a later date.[27] Thus, although these best responses are myopic, any eventual changes in behavior are equivalently forward-looking.

Figure 4 depicts a mapping ϕ governing the dynamics. Equilibria, and resting points of the diffusion process, correspond to intersections of ϕ with the 45 degree line.

The figure allows an immediate distinction between two classes of equilibria that we discussed informally up to now. Formally, an equilibrium x is *stable* if there exists $\varepsilon' > 0$ such that $\phi(x-\varepsilon) > x-\varepsilon$ and $\phi(x+\varepsilon) < x+\varepsilon$ for all $\varepsilon' > \varepsilon > 0$. An equilibrium x is *unstable* or a *tipping point* if there exists $\varepsilon' > 0$ such that $\phi(x-\varepsilon) < x-\varepsilon$ and $\phi(x+\varepsilon) > x+\varepsilon$ for all $\varepsilon' > \varepsilon > 0$. In the figure, the equilibrium to the left is a tipping point, while the equilibrium to the right is stable.

---

[27] If actions are strategic substitutes, convergence may not be guaranteed for all starting points. However, whenever convergence is achieved, the rest point is an equilibrium, and the analysis can therefore be useful for such games as well.

The composition of the equilibrium set hinges on the shape of the function ϕ. Furthermore, note that a point-wise shift of ϕ (as in the figure, to a new function $\tilde{\phi}$) shifts tipping points to the left and stable points to the right, loosely speaking (as sufficient shifts may eliminate some equilibria altogether), making adoption more likely. This simple insight allows for a variety of comparative statics.

For instance, consider an increase in the cost of adoption, manifested as a First Order Stochastic Dominance (FOSD) shift of the cost distribution F^c to $\bar{F}^c$. It follows immediately that:

$$\overline{\phi}(x) = \sum_d \widetilde{P}(d)\overline{F}^c(v(d,x)) \leqslant \sum_d \widetilde{P}(d)F^c(v(d,x)) = \phi(x)$$

and the increase in costs corresponds to an increase of the tipping points and decrease of the stable equilibria (one by one). Intuitively, increasing the barrier to choosing the action 1 leads to a higher fraction of existing adopters necessary to get the action 1 to spread even more.

This formulation also allows for an analysis that goes beyond graphical games regarding the social network itself, using stochastic dominance arguments (following Jackson and Rogers, 2007, and Jackson and Yariv, 2005, 2007). For instance, consider an increase in the expected degree of each random neighbor that an agent has. That is, suppose $\widetilde{P}'$ FOSD $\widetilde{P}$ and, for illustration, assume that $F(d,x)$ is non-decreasing in d for all x. Then, by the definition of FOSD,

$$\phi'(x) = \sum_d \widetilde{P}'(d)F(d,x) \geqslant \sum_d \widetilde{P}(d)F(d,x) = \phi(x),$$

and, under P', tipping points are lower and stable equilibria are higher.

Similar analysis allows for comparative statics regarding the distribution of links, by simply looking at Mean Preserving Spreads (MPS) of the underlying degree distribution.[28]

---

[28] In fact, Jackson and Yariv (2007) illustrate that if $F(d,x)$ is non-decreasing and convex, then power, Poisson, and regular degree distributions with identical means generate corresponding values of $\phi^{power}, \phi^{Poisson}$, and $\phi^{regular}$ such that

$$\phi^{power}(x) \geqslant \phi^{Poisson}(x) \geqslant \phi^{regular}(x)$$

for all x, thereby implying a clear ranking of the tipping points and stable equilibria corresponding to each type of network.

Going back to the dynamic path of adoption, we can generalize the insights that we derived regarding the Granovetter (1978) model. Namely, whether adoption paths track an S-shaped curve now depends on the shape of ϕ, and thereby on the shape of both the cost distribution F and agents' utilities.

5   Closing Notes

There is now a substantial and growing body of research studying the impacts of interactions that occur on a network of connections. This work builds on the empirical observations of peer influence and generates a rich set of individual and aggregate predictions. Insights that have been shown consistently in real-world data pertain to the higher propensities of contagion (of a disease, an action, or behavior) in more highly connected individuals, the role of "opinion leaders" in diffusion, as well as an aggregate S-shape of many diffusion curves. The theoretical analyses open the door to many other results, e.g., those regarding comparative statics *across* networks, payoffs, and cost distributions (when different actions vary in costs). Future experimental and field data will hopefully complement these theoretical insights.

A shortcoming of some of the theoretical analyses described in this chapter is that the foundation for modeling the underlying network is rooted in simple forms of random graphs in which there is little heterogeneity among nodes other than their connectivity. This misses a central observation from the empirical literature that illustrates again and again the presence of homophily, people's tendency to associate with other individuals who are similar to themselves. Moreover, there are empirical studies that are suggestive of how homophily might impact diffusion, providing for increased local connectivity but decreased diffusion on a more global scale (see Rogers (1995) for some discussion). Beyond the implications that homophily has for the connectivity structure of the network, it also has implications for the propensity of individuals to be affected by neighbors' behavior: for instance, people who are more likely to, say, be immune may be more likely to be connected to one another,

and, similarly, people who are more likely to be susceptible to infection may be more likely to be connected to one another.[29] Furthermore, background factors linked to homophily can also affect the payoffs individuals receive when making decisions in their social network. Enriching the interaction structure in that direction is crucial for deriving more accurate diffusion predictions. This is an active area of current study (e.g., see Baccara and Yariv, 2009, Bramoullé and Rogers, 2010, Currarini, Jackson, and Pin, 2009, 2010ab, and Peski, 2008).

Ultimately, the formation of a network and the strategic interactions that occur amongst individuals is a two-way street. Developing richer models of the endogenous formation of networks, together with endogenous interactions on those networks, is an interesting direction for future work, both empirical and theoretical.[30]

---

[29]The mechanism through which this occurs can be rooted in background characteristics such as wealth, or more fundamental personal attributes such as risk aversion. Risk averse individuals may connect to one another and be more prone to protect themselves against diseases by, e.g., getting immunized; similarly for wealth.

[30]As discussed above, there are some studies, such as that of Kandel (1978), that provide evidence for the back and forth interaction between behavior and network formation. There are also some models that study co-evolving social relationships and play in games with neighbors, such as Ely (2001), Mailath, Samuelson, and Shaked (2001), Jackson and Watts (2002, 2005), Droste, Gilles, and Johnson (2003), Corbae and Duffy (2003), and Goyal and Vega-Redondo (2005). These articles only begin to provide insight into such interplay.

References

[1] Aral, S. and Walker, D. (2010), "Creating Social Contagion through Viral Product Design: Theory and Evidence from a Randomized Field Experiment," mimeo.

[2] Aral, S., L. Muchnik, and Sundararajan, A. (2009) "Distinguishing Influence Based Contagions from Homophily Driven Diffusion in Dynamic Networks," *Proceedings of the National Academy of Sciences.*

[3] Baccara, M., Imrohoroglu, A., Wilson, A. J., and Yariv, L. (2009), "A Field Study on Matching with Network Externalities," mimeo.

[4] Baccara, M. and Yariv, L. (2009), "Similarity and Polarization in Groups," mimeo.

[5] Beaman, L. and Magruder, J. (2010) "Who gets the job referral? Evidence from a social networks experiment," mimeo.

[6] Bramoullé, Y. and Rogers, B. W. (2009), "Diversity and Popularity in Social Networks," mimeo.

[7] Bailey, N. T. J. (1975), *The Mathematical Theory of Infectious Diseases and Its Applications*, London Griffin.

[8] Ballester, C., Calvó-Armengol, A., and Zenou, Y. (2006), "Who's who in networks. Wanted: The Key Player," *Econometrica*, 74(5), 1403-1417.

[9] Bandiera, O., Barankay, I., and Rasul, I. (2008),"Social capital in the workplace: Evidence on its formation and consequences, *Labour Economics*, 15(4), 724-748.

[10] Banerjee, A., Chandrasekhar, A., Duflo, E., and Jackson, M. O. (2010) "The Diffusion of MicroFinance in Rural India," mimeo.

[11] Bass, F. (1969), "A new product growth model for consumer durables," *Management Science*, 15(5), 215-227

[12] Bayer, P., S. Ross, and Topa, G. (2008), "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *The Journal of Political Economy*, 116(6), 1150-1196.

[13] Bearman, P. S., Moody, J., and Stovel, K. (2004), "Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks," *American Journal of Sociology*, 110(1), 44-91.

[14] Bollobás, B. (1980):"A probabilistic proof of an asymptotic formula for the number of labelled regular graphs", *European Journal of Combinatorics* **1**, 311-16.

[15] Bramoullé, Y. and Kranton, R. (2007), "Strategic Experimentation in Networks," *Journal of Economic Theory*, 135(1), 478-494.

[16] Calvó-Armengol, A. and Jackson, M. O. (2009), "Like Father, Like Son: Labor Market Networks and Social Mobility," *American Economic Journal: Microeconomics*, 1(1), 124-150.

[17] Christakis, N. A. and Fowler, J. H. (2007), "The Spread of Obesity in a Large Social Network over 32 Years," *The New England Journal of Medicine*, 357(4), 370-379.

[18] Chwe, M. S-Y. (2000), "Communication and Coordination in Social Networks," *The Review of Economic Studies*, 67, 1-16.

[19] Coleman, J.S., Katz, E., and Menzel, H. (1966), *Medical Innovation: A Diffusion Study*, Bobbs-Merrill: Indianapolis.

[20] Colizza, V., Barrat, A., Barthélemy, M., and Vespignani, A. (2006), "The role of the airline transportation network in the prediction and predictability of global epidemics," Proceedings of the National Academy of Sciences USA, 103, 2015-2020.

[21] Colizza, V., Barrat, A., Barthelemy, M., and Vespignani, A. (2007),"Predictability and epidemic pathways in global outbreaks of infectious diseases: the SARS case study," *BMC Medicine*, 5, 34.

[22] Colizza, V. and Vespignani, A. (2007) "Invasion Threshold in Heterogeneous Metapopulation Networks," *Physical Review Letters*, 99, 148701.

[23] Conley, T. G. and Topa, G. (2002), "Socio-Economic Distance and Spatial Patterns in Unemployment," *Journal of Applied Econometrics*, 17(4), 303-27.

[24] Conley, T. G. and Udry, C. (2008), "Learning About a New Technology: Pineapple in Ghana," *The American Economic Review*, forthcoming.

[25] Corbae, D. and Duffy, J. (2003) "Experiments with Network Formation," forthcoming : *Games and Economic Behavior*

[26] Currarini, S., Jackson, M. O., and Pin, P. (2009), "An Economic Model of Friendship: Homophily, Minorities and Segregation," *Econometrica*, 77(4), 1003–1045.

[27] Currarini, S., Jackson, M. O., and Pin, P. (2010), "Identifying the roles of race-based choice and chance in high school friendship network formation," in the *Proceedings of the National Academy of Sciences* 107(11): 4857–4861, March 16, 2010.

[28] Duflo, E. and Saez, E. (2003) "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence From a Randomized Experiment," *Quarterly Journal of Economics*,

[29] Dupas, P. (2010) "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment," mimeo: UCLA.

[30] Ellison, G. (1993), "Learning, Local Interaction, and Coordination," *Econometrica*, 61, 1047-1071.

[31] Ely, J. C. (2002) "Local Conventions," *Advances in Theoretical Economics*: Vol. 2 : Iss. 1, Article 1.

[32] Erdős, P. and Rényi, A. (1959), "On random graphs," *Publicationes Mathematicae,* 6, 290–297.

[33] Erdős, P. and Rényi, A. (1960), "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5, 17–61.

[34] Feick, L. F. and Price, L. L. (1987), "The Market Maven: A Diffuser of Marketplace Information," *Journal of Marketing*, 51(1), 83-97.

[35] Feigenberg, B., E. Field, and Pande, R. (2010) "Building Social Capital through Microfinance," mimeo: Harvard University.

[36] Foster, A. D. and Rosenzweig, M. R. (1995), "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103(6), 1176-1209.

[37] Galeotti, A., Goyal, S., Jackson, M. O., Vega-Redondo, F., and Yariv, L. (2010), "Network Games," *The Review of Economic Studies*, 77(1), 218-244.

[38] Galeotti, A. and Vega-Redondo, F. (2006), "Complex Networks and Local Externalities: A Strategic Approach," mimeo.

[39] Glaeser, E. L. and Scheinkman, J. A. (2000), "Non-market Interactions," NBER Working Paper Number 8053.

[40] Glaeser, E. L., Sacerdote, B., and Scheinkman, J. A. (1996), "Crime and Social Interactions," *The Quarterly Journal of Economics*, 111(2), 507-548.

[41] Goeree, J. K., McConnell, M. A., Mitchell, T. , Tromp, T., and Yariv, L. (2010), "The 1/d Law of Giving," *American Economic Journal: Microeconomics*, 2(1), 183-203.

[42] Goyal, S. and Moraga-Gonzalez, J. L. (2001), "R&D Networks," *Rand Journal of Economics*, 32(4), 686-707.

[43] Goyal, S. and Vega-Redondo, F. (2005) "Network Formation and Social Coordination,' *Games and Economic Behavior*, 50, 178–207.

[44] Granovetter, M. (1973), "The Strength of Weak Ties," *American Journal of Sociology*, 78(6), 1360-1380.

[45] Granovetter, M. (1978), "Threshold Models of Collective Behavior," *American Journal of Sociology*, 83(6), 1420-1443.

[46] Granovetter, M. (1985), "Economic action and social structure: the problems of embeddedness," *American Journal of Sociology*, 91(3), 481-510.

[47] Granovetter, M. (1995), *Getting a Job: A Study in Contacts and Careers*, University of Chicago Press.

[48] Griliches, Z. (1957), "Hybrid corn: an exploration of the economics of technological change," *Econometrica*, 25, 501-522.

[49] Huckfeldt, R. R. and Sprague, J. (1995), *Citizens, Politics and Social Communication*, Cambridge Studies in Public Opinion and Political Psychology.

[50] Ioannides, Y. M. and Datcher-Loury, L. (2004). "Job information networks, neighborhood effects and inequality," *Journal of Economic Literature*, 424, 1056 - 93

[51] Jackson M.O. (2007) "Social Structure, Segregation, and Economic Behavior," Nancy Schwartz Memorial Lecture, given in April 2007 at Northwestern University, printed version: http://www.stanford.edu/∼jacksonm/schwartzlecture.pdf

[52] Jackson, M. O. (2008), *Social and Economic Networks*, Princeton University Press: NJ.

[53] Jackson, M. O. and Rogers, B. W. (2007), "Relating Network Structure to Diffusion Properties through Stochastic Dominance," *The B.E. Press Journal of Theoretical Economics*, 7(1) (Advances), 1-13.

[54] Jackson, M. O. and Watts, A. (2002) "On the Formation of Interaction Networks in Social Coordination Games,' *Games and Economic Behavior* 41, 265-291.

[55] Jackson, M. O. and Watts, A. (2005) "Social Games: Matching and the Play of Finitely Repeated Games,' forthcoming: *Games and Economic Behavior.*

[56] Jackson, M. O. and Yariv, L. (2005), "Diffusion on Social Networks," *Economie Publique*, 16(1), 69-82.

[57] Jackson, M. O. and Yariv, L. (2007), "Diffusion of Behavior and Equilibrium Properties in Network Games," *The American Economic Review (Papers and Proceedings)*, 97(2), 92-98.

[58] Kakade, S., Kearns, M., Langford, J., and Ortiz, L. (2003), "Correlated Equilibria in Graphical Games, " *ACM Conference on Electronic Commerce*, New York.

[59] Kandel, D. B. (1978), "Homophily, Selection, and Socialization in Adolescent Friendships," *The American Journal of Sociology*, 84(2), 427-436.

[60] Karlan, D., Mobius, M. M., Rosenblat, T. S., and Szeidl, A. (2009) "Measuring Trust in Peruvian Shantytowns" mimeo.

[61] Katz, E. and Lazarsfeld, P. F. (1955), *Personal influence: The part played by people in the flow of mass communication*, Glencoe, IL: Free Press.

[62] Kearns, M., Littman, M., and Singh, S. (2001), Graphical Models for Game Theory, in J.S. Breese, D. Koller (eds.), *Proceedings of the 17th Conference on Uncertainty in Artificial Intelligence,* San Francisco: Morgan Kaufmann University of Washington, Seattle, Washington, USA, 253-260.

[63] Kermack, W. O. and McKendrick, A. G. (1927), "A Contribution to the Mathematical Theory of Epidemics," *Proceedings of the Royal Society of London*, A 115, 700-721.

[64] Kosfeld, M. (2004), "Economic Networks in the Laboratory: A Survey," *Review of Network Economics*, 30, 20-42.

[65] Lazarsfeld, P. E., Berelson, B., and Gaudet, G. (1944), *The people's choice: How the voter makes up his mind in a presidential campaign*, New York: Duell, Sloan and Pearce.

[66] Leider, S., Mobius, M. M., Rosenblat, T., and Do, Q-A. (2009), "Directed Altruism and Enforced Reciprocity in Social Networks," *The Quarterly Journal of Economics*, forthcoming.

[67] Leskovec, J., Adamic, L. A., and Huberman, B. A. (2007) "The dynamics of viral marketing," *ACM Transactions on the Web (TWEB) archive*, 1:1, Article No. 5.

[68] Lopez-Pintado, D. (2008) "Contagion in Complex Social Networks," *Games and Economic Behavior*, 62(2), 573-590 .

[69] Mahajan, V. and Peterson, R. A. (1985), *Models for Innovation Diffusion (Quantitative Applications in the Social Sciences)*, Sage Publications, Inc.

[70] Manski, C. F. (2000), "Economic Analysis of Social Interactions," *The Journal of Economic Perspectives*, 14(3), 115-136.

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

[72] Merton, R. (1948), "Patterns of Influence: a study of interpersonal influence and of communications behavior in a local community," In Lazarsfeld, Paul and Stanton, Frank (eds.), *Communication Research*, NewYork: Harper and Brothers. pp180-215.

[73] Morris, S. (2000), "Contagion," *The Review of Economic Studies*, 67(1), 57-78.

[74] Morris, S. and Shin, H. (2003), "Global Games: Theory and Applications,", in *Advances in Economics and Econometrics* (Proceedings of the Eighth World Congress of the Econometric Society), edited by M. Dewatripont, L. Hansen and S. Turnovsky. Cambridge: Cambridge University Press, 56-114.

[75] Morris, S. and Shin, H. (2005), "Heterogeneity and Uniqueness in Interaction Games," in *The Economy as an Evolving Complex System III*, edited by L. Blume and S. Durlauf; Oxford University Press, Santa Fe Institute Studies in the Sciences of Complexity.

[76] Nair, H., Manchanda, P., and Bhatia, T. (2006), "Asymmetric Effects in Physician Prescription Behavior: The Role of Opinion Leaders," mimeo.

[77] Pastor-Satorras, R. and Vespignani, A. (2000) "Epidemic Spreading in Scale-Free Networks," *Physical Review Letters*, 86(14), 3200-3203.

[78] Pastor-Satorras, R. and Vespignani, A. (2001) "Epidemic dynamics and endemic states in complex networks," *Physical Review E*, 63, 066117.

[79] Pesendorfer, W. (1995), "Design Innovation and Fashion Cycles," *The American Economic Review*, 85(4), 771-792.

[80] Peski, M. (2008), "Complementarities, Group Formation, and Preferences for Similarity," mimeo.

[81] Rees, A. (1966), "Information Networks in Labor Markets," *The American Economic Review*, 56 (2), 559-566.

[82] Rees, A. and Schultz, G. P. (1970), *Workers and Wages in an Urban Labor Market*, University of Chicago Press.

[83] Rogers, E. M. (1995), *Diffusion of Innovations*, Free Press.

[84] Ryan, B. and Gross, N. C. (1943) "The diffusion of hybrid seed corn in two Iowa communities," *Rural Sociology*, 8, 15–24.

[85] Schelling, T. C. (1978), *Micromotives and Macrobehavior*, W. W. Norton.

[86] Solomonoff, R. and Rapoport, A. (1951), "Connectivity of Random Nets," *Bulletin of Mathematical Biophysics*, 13, 107-117.

[87] Sundararajan, A. (2007), "Local network effects and network structure," *The B.E. Journal of Theoretical Economics*: Vol. 7 : Iss. 1 (Contributions), Article 46.

[88] Tarde, G. (1903), *Les Lois de L'Imitation: Etude Sociologique*, Elibron Classics, translated to English in *The Laws of Imitation*, H. Holt and company.

[89] Topa, G. (2001), "Social Interactions, Local Spillovers and Unemployment," *Review of Economic Studies*, 68(2), 261-295.

[90] Tucker, C. (2008), "Identifying Formal and Informal Inuence in Technology Adoption with Network Externalities," *Management Science*, 55(12), 2024-2039.

[91] Van den Bulte, C. and Lilien, G. L. (2001), "Medical Innovation Revisited: Social Contagion versus Marketing Effort," *American Journal of Sociology*, 106(5), 1409-1435.

[92] Young, P. (2010), "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning," *The American Economic Review*, forthcoming.