MIT Sloan School of Management

MIT Sloan School Working Paper 4763-09

Social Learning in Social Networks

PJ Lamberson

Social Learning in Social Networks

P. J. Lamberson[a]

[a] *MIT Sloan School of Management, E53-339, 77 Massachusetts Ave, Cambridge, MA 02139.*
Voice: 617-252-1464, Fax: 617-258-7579

Abstract

This paper analyzes a model of social learning in a social network. Agents decide whether or not to adopt a new technology with unknown payoffs based on their prior beliefs and the experiences of their neighbors in the network. Using a mean-field approximation, we prove that the diffusion process always has at least one stable equilibrium, and we examine the dependance of the set of equilibria on the model parameters and the structure of the network. In particular, we show how first and second order stochastic dominance shifts in the degree distribution of the network impact diffusion. We find that the relationship between equilibrium diffusion levels and network structure depends on the distribution of payoffs to adoption and the distribution of agents' prior beliefs regarding those payoffs, and we derive the precise conditions characterizing those relationships.

Key words: social networks, learning, diffusion, mean-field analysis, stochastic dominance

Email addresses: pjl@mit.edu (P. J. Lamberson)

1. Introduction

When choosing whether or not to adopt a new technology, people often rely on information outside of their personal experience to make their decision. One potential source of information is other individuals who have already tried the technology. If the information from previous adopters is sufficiently positive, an initially skeptical individual may be convinced to adopt, making them a potential source of information for others in the future. Collectively, the population learns about the value of the new technology as it spreads through the market. This mechanism of *social learning* is a simple but compelling explanation for technology diffusion.

In seeking information from previous adopters, individuals most likely turn to their friends and acquaintances – in other words *neighbors* in their *social network* (Jackson, 2008). In this paper we analyze a model of social learning in which agents are embedded in a social network that dictates who interacts with whom. Each agent in the model employs a boundedly rational decision-making rule to determine whether or not to adopt a new technology. Specifically, agents combine information on the payoffs received by their adopting neighbors with their prior beliefs using Bayes rule, and they adopt the technology if their resulting posterior beliefs regarding the value of the technology exceed the known payoff of the status quo. Agents in the model may also discontinue using technology if payoffs observed later cause them to revise their beliefs about the technology's benefits.

To analyze the resulting diffusion we employ an approximation technique from statistical physics known as a mean-field approximation (Chandler, 1987), which has proven useful for studying network dynamics in several contexts (Newman et al., 2000; Pastor-Satorras and Vespignani, 2001a,b; Jackson and Yariv, 2005, 2007; Jackson and Rogers, 2007; López-Pintado, 2008; Lamberson, 2009). We prove that this approximation to the social learning process always has at least one stable equilibrium. In general there may be multiple stable equilibria. We derive conditions that guarantee a unique stable equilibrium for "costly" technologies, i.e. those with mean payoff less than that of the status quo, and explain why non-costly technologies are more likely to give rise to multiple equilibria. We then proceed to analyze how equilibrium levels of diffusion depend on the parameters of the model, specifically the distribution of payoffs to adoption and the distribution of agents' prior beliefs regarding those payoffs. Some of these relationships are the same as in models without network structure: higher payoffs and higher priors result in greater diffusion. However, the effects of changing the variance of the payoff distribution or the variance of the distribution of priors depends on the network.

The chief advantage of this model in comparison with previous studies of social learning is the inclusion of potentially complex network structures that govern agent interactions. We prove that networks matter for diffusion: changes in network structure cause predictable changes in diffusion levels. The effect of network structure on diffusion depends in subtle ways on the relationship between the costs or benefits of the new technology and agents' prior beliefs about those costs or benefits. Specifically, we consider the effects of two types of changes in the network: first and second order stochastic shifts in the degree distribution. Intuition might suggest that adding more links to the network (i.e. a first order stochastic shift in the degree distribution) would increase the diffusion of beneficial technologies and decrease diffusion for those that are costly. We confirm this intuition in some cases, but show that the opposite is true in others. When the agents' prior beliefs are sufficiently positive, adding links to the network can lead a beneficial technology to diffuse less. Similarly, when agents are strongly biased against adoption, adding links can lead more agents to adopt a costly technology. In these cases, agents would ultimately be better off with less information – their initial beliefs give rise to better decisions than those based on knowledge gained from neighbors' experiences. The effect of second order stochastic shifts is more complicated and often varies depending on the initial level of adoption. We illustrate this ambiguous relationship with an example comparing diffusion in a regular network and a scale-free network with the same average degree.

Finally, we extend the basic model by allowing agents to incorporate their observation of payoffs from the more distant past in their decision. We show that the number of past observations that agents consider shapes diffusion in a way that is analogous to the conditional effect on diffusion of first order stochastic shifts in the degree distribution.

Social learning has a rich history in both theoretical economics and empirical research.[1] The foundational social learning models of Ellison and Fudenberg (1993, 1995) were among the first to examine the collective outcome of individual agents employing simple boundedly rational decision

[1]For a sampling of the theoretical literature see Bikhchandani et al. (1992); Banerjee (1992); Kirman (1993); Ellison and Fudenberg (1993, 1995); Kapur (1995); Bala and Goyal (1998); Smith and Sørensen (2000); Chamley (2003); Chatterjee and Xu (2004); Banerjee and Fudenberg (2004); Manski (2004); Young (2006) and Golub and Jackson (2009). Foster and Rosenzweig (1995); Munshi (2004) and Conley and Udry (2005) present empirical studies supporting the theory.

rules. The most significant departure between our model and those of Ellison and Fudenberg (1993, 1995), and most other social learning models, is that in the model presented here, agents' interactions are limited by a social network. In all but one of the models considered by Ellison and Fudenberg agents interact randomly. The Ellison and Fudenberg (1993) model that includes structured interactions, does so in a particularly simple form: agents are located on a line and pay attention only to other agents located within a given distance. Despite the simplicity of that model, they find that the "window width," i.e. the number of neighbors from which each agent seeks information, affects both the efficiency and speed to convergence of the model. This hints at the importance of the structure of agents' interactions in diffusion. The model presented here allows us to analyze more complex network settings and the dependence of equilibria on the network structure. Beyond Ellison and Fudenberg's window width result, several empirical studies have argued that technologies and behaviors spread through social networks, and both computational and analytic models have illustrated that network structure can either facilitate or hamper diffusion.[2]

Bala and Goyal (1998) also tackle the problem of social learning in a social network. Their model makes a key assumption that we do not: agents have infinite memories and take into account all of their past observations when making their decision. In the model presented here, agents have finite memories, and we examine the dependance of diffusion equilibria on the length of that memory as discussed above. The assumption of infinite memory qualitatively changes the results of the model because it allows agents to take an action infinitely often, and thereby learn and communicate the true payoffs of the action. Ultimately this implies that in the limit all agents must receive the same utility and – if the utility to different actions differs – choose the same action (see also Jackson, 2008). Unlike Bala and Goyal's model, the agents' choices at stable equilibria in our model are always diverse; some agents will adopt while others do not. However, taking a limit as the number of time periods that agents consider in their decisions goes to infinity produces results in our model that agree with those of Bala and Goyal (1998).

The spirit and techniques of our analysis are most similar to several recent papers which also

[2]For a sampling of the empirical literature, see Coleman et al. (1966); Burt (1987); Christakis and Fowler (2007, 2008); Fowler and Christakis (2008) and Nickerson (2008). For theoretical explorations of network structure and diffusion, see Watts and Strogatz (1998); Newman et al. (2000); Pastor-Satorras and Vespignani (2001a,b); Sander et al. (2002); Jackson and Yariv (2005); Centola and Macy (2007); Jackson and Yariv (2007); Jackson and Rogers (2007) and López-Pintado (2008).

employ a mean-field approach to study network diffusion (Jackson and Yariv, 2005, 2007; Jackson and Rogers, 2007; López-Pintado, 2008). These models differ from the one presented in this paper in the specification of the individual decision rules. In the models of Jackson and Yariv (2005, 2007), Jackson and Rogers (2007), and López-Pintado (2008), the new technology or behavior spreads either by simple contact, like a disease, or through a social influence or "threshold model," in which agents adopt once a certain threshold number of their neighbors adopt. Our model adds a more sophisticated decision rule. As Young (2009) points out, of these three diffusion models – contagion, threshold models, and social learning – "social learning is certainly the most plausible from an economic standpoint, because it has firm decision-theoretic foundations: agents are assumed to make rational use of information generated by prior adopters in order to reach a decision." In addition to providing a microeconomic rationale for adopter decisions, the social learning model considered here also solves the "startup problem" of the contact and threshold models. In those models, no adoption is always a stable equilibrium. In order to start the diffusion process at least one agent must be exogenously selected to be an initial adopter. The model in this paper provides an endogenous solution to the startup problem: those agents with positive priors adopt the technology initially without need for an exogenous shock.

The paper proceeds as follows. Section 2 details the social learning model. Section 3 applies the mean-field analysis to approximate the dynamics of the model, and section 4 uses that approximation to find diffusion equilibria. In section 4, we also characterize stable and unstable equilibria and prove that for any set of parameters at least one stable equilibrium exists. Section 5 turns to analyzing the dependance of equilibrium levels of diffusion on the model parameters and the network structure. In Section 6 the model is extended to incorporate finite memory of arbitrary duration and proceeds to describe how the equilibria change with changes in the length of agents' memory. Section 7 concludes with a discussion of extensions for future research.

2. The Model

This section develops a simple model of social learning in a social network. Throughout the paper, we refer to the adoption of a new technology, but the model and results may apply equally to the diffusion of other behaviors that spread through social networks, such as smoking or political participation (Christakis and Fowler, 2008; Nickerson, 2008).

At each time in a discrete sequence of time steps, each agent in the model chooses whether or not to use a new technology with an unknown payoff. Each agent's decision is made by comparing her beliefs about the unknown payoffs against the known payoff of the status quo. If an agent believes the payoff to the new technology exceeds that of the status quo, then the agent will use it in the following time step. Conversely, if she believes the payoffs are less than the status quo, she will not use it. In the first time step, adoption decisions are made based solely on the agents' prior beliefs about the value of the technology. In each subsequent time step, an agent's beliefs about the technology's payoffs are formed by using Bayes rule to combine her prior beliefs with observation of the payoffs received in the previous period by her adopting neighbors in the social network (and her own if she also used the technology).[3]

Each period that an adopting agent continues to use the technology she receives a new payoff. Following Young (2009), we assume that the payoffs are independent and identically distributed across time and across agents.

There are several assumptions implicit in this model. First, the payoffs to neighboring agents are observable. This assumption stands in contrast to "herd models" in which agents' adoption decisions are observable, but their payoffs are not (e.g. Scharfstein and Stein, 1990; Bikhchandani et al., 1992). The assumption that agents know the payoffs of their adopting neighbors seems best justified in situations where the technology is sufficiently costly that agents would actively solicit payoff information instead of passively observing their neighbors' choices. For example, this assumption might apply in the decision of whether or not to adopt a hybrid electric vehicle or a new type of cellular phone.

Second, agents do not take into account some information implicit in their neighbors' adoption decisions – for instance, that the neighbors of their neighbors have had positive experiences. This departure from complete rationality is a common modeling assumption in the social learning literature (e.g. Ellison and Fudenberg, 1993, 1995; Young, 2009), which reflects our belief that the calculations involved in the fully rational Bayesian decision rule are unrealistically complicated.

Third, agents only attempt to maximize their next period expected payoffs. They will not

[3]Throughout the paper we refer to agents currently using the technology as adopters and those not using the technology as non-adopters. However, it is possible that an agent of either type has experienced a sequence of adoptions and disadoptions in the past.

experiment with the new technology just to gain information. This assumption is justified by the potential cost of adoption, as well as the limited complexity of realistic consumer decision rules.

Finally, the agents' decisions are based only on the recent past; they only incorporate payoffs from the previous period in updating their beliefs. In making this assumption, our model also follows Ellison and Fudenberg (1993, 1995). As they argue, the assumption of limited memory makes sense when adoption decisions are revised infrequently. We later relax this assumption and examine the effect of increasing the number of previous periods that the agents incorporate in their decision making.

Formally, at each time t each agent using the technology receives a payoff drawn from a normal distribution with mean μ and variance σ^2. The payoffs are assumed to be independent across time and across agents. Let $\tau = 1/\sigma^2$ be the precision of the payoff distribution. Following the standard Bayesian model, we assume that the agents have conjugate prior distributions regarding the unknown mean and variance of the payoff distribution (DeGroot, 1970; Gelman et al., 2004). Specifically, each agent i has prior distributions for μ and σ satisfying

$$\mu|\sigma^2 \sim N(\mu_{i0}, \sigma^2/\tau_{i0}) \tag{1}$$

$$\sigma^2 \sim \text{Inv}\chi^2(\nu_{i0}, \sigma_0^2). \tag{2}$$

The agents have heterogenous prior beliefs about the payoffs of the new technology, reflected by the individual specific parameters μ_{i0}, which are assumed to be distributed normally across the agents with mean m and variance s^2. To simplify the analysis we assume that all agents have identical precision on their prior distribution for μ given σ^2 and without loss of generality set this equal to one, so $\tau_{i0} = 1$ for all i. We also assume that the payoff to the status quo is constant and equal for all agents.

Agents update their beliefs and choices at each time t as follows. First, each agent seeks information on the value of the technology from their neighbors in the social network. Those neighbors that used the technology in the previous period report the payoff they received. Agents that did not use the technology at time $t-1$ provide no information. Then the agents reevaluate their adoption decision based on the information gained from their neighbors, and their own payoff from the previous period if they used the technology at time $t-1$, using Bayes rule. Finally, agents that choose to use the technology based on their updated beliefs receive a new payoff, which will inform their choice and their neighbors' choices in the following period.

Suppose that an agent i observes n_{it-1} payoffs at time $t-1$. Let $\bar{y}_{it-1}$ denote the mean realized payoff to those n_{it-1} adopters. Then i's posterior distribution for μ given σ^2 is normal with mean

$$\mu_{it} = \frac{n_{it-1}}{n_{it-1}+1}\bar{y}_{it-1} + \frac{1}{n_{it-1}+1}\mu_{i0} \tag{3}$$

(Gelman et al., 2004). This is agent i's expectation regarding the payoff of using the new technology given her prior beliefs and the data observed from her neighbors' experiences (and possibly her own). As specified in equation (3), this posterior is simply a weighted average of agent i's prior mean and the mean of the n_{it} payoffs that i observes, where the weight on the observed payoffs is equal to their number. Here network structure begins to play a role in the diffusion. Agents with more neighbors will on average have more observations on which to base their decision and will place greater weight on those observations relative to their prior beliefs.

We assume that the payoff of the status quo is equal for all agents and without loss of generality set this to zero. Thus, agent i uses the technology at time t if the mean of her posterior distribution for μ, μ_{it}, is greater than zero. If $\mu_{it} \leq 0$, she will not use the technology at time t.

3. Mean-field Analysis

Even this simple model of social learning in a network is rendered analytically intractable by the potential for multiple equilibria depending on the specifics of the network of connections and the distribution of prior beliefs and payoffs. Following previous studies (Jackson and Yariv, 2005, 2007; Jackson and Rogers, 2007; López-Pintado, 2008; Galeotti et al., 2009; Lamberson, 2009), we employ a mean-field analysis to approximate the dynamics.

Let P denote the degree distribution of the network, so $P(d)$ equals the probability that a randomly chosen agent is of degree d. We assume that the network is connected. Let π_{dt} denote the probability that a randomly chosen degree d agent is an adopter at time t and

$$\theta_t = \frac{1}{\bar{d}}\sum_d dP(d)\pi_{dt} \tag{4}$$

denote the probability that a randomly chosen link from any given agent points to an adopter. Following Jackson and Yariv (2005) we call this the *link-weighted fraction of adopters*.[4] The main assumption of the mean-field approximation is that the fraction of each agents' neighbors that

[4] This is different from the overall fraction of adopters in the network, $\sum_d P(d)\pi_{dt}$, because higher degree agents

are adopters at time t is given by (4). So, at time t a degree d agent observes the payoffs from $d\theta_{t-1}$ adopting neighbors. Agents currently using the technology also observe one additional payoff, their own, but for analytic convenience we assume that both those agents not currently using the technology and those currently using the technology observe the same number of payoffs.[5] Thus, in equation (3) we replace n_{it-1} with $d\theta_{t-1}$:

$$\mu_{dt} = \frac{d\theta_{t-1}}{d\theta_{t-1}+1}\bar{y}_{it-1} + \frac{1}{d\theta_{t-1}+1}\mu_{i0}. \tag{5}$$

Intuitively, the mean-field approximation can be thought of as follows. Rather than existing in a static network, each agent i has a type given by her degree d_i. At each time t, an agent of type d_i polls a sample of d_i other agents on their experience with the new technology, and in that sample the fraction of agents who have adopted matches the link-weighted fraction of adopters in the population as a whole. This approximation method ignores much of the structure in the network of connections. Nevertheless, as we show below, even with this simplified representation the structure of social interactions affects the technology diffusion.

Since adopters' experiences are distributed $N(\mu, \sigma^2)$, the sample mean $\bar{y}_{it-1}$ from a sample of size $d\theta_{t-1}$ is distributed $N(\mu, \frac{\sigma^2}{d\theta_{t-1}})$. The prior beliefs μ_{i0} are distributed $N(m, s^2)$, so the posterior beliefs of the degree d agents determined by equation (5) are distributed

$$\mu_{dt} \sim N\left(\frac{d\theta_{t-1}\mu + m}{d\theta_{t-1}+1}, \frac{d\theta_{t-1}\sigma^2 + s^2}{(d\theta_{t-1}+1)^2}\right). \tag{6}$$

Since an agent will use the new technology at time t if the mean of her posterior distribution for μ is positive, the probability that a degree d agent will use the technology at time t is

$$\pi_{dt} = \Phi\left(\left(\frac{d\theta_{t-1}\mu + m}{d\theta_{t-1}+1}\right)\Big/\left(\frac{\sqrt{d\theta_{t-1}\sigma^2 + s^2}}{d\theta_{t-1}+1}\right)\right) = \Phi\left(\frac{d\theta_{t-1}\mu + m}{\sqrt{d\theta_{t-1}\sigma^2 + s^2}}\right), \tag{7}$$

where Φ is the standard normal cumulative distribution function.

---

are more likely to lie at the opposite end of a randomly chosen link than lower degree agents. Equation (4) correctly (assuming no correlation in neighboring agents' degrees) accounts for this by weighting $P(d)\pi_{dt}$ by d (see Jackson and Yariv, 2005; Jackson and Rogers, 2007).

[5]Without this assumption the analysis becomes substantially more complicated. One possible justification is that agents currently using the technology have less of an incentive to seek information from their neighbors, since they also observe their personal payoff, resulting in degree d adopters and degree d non-adopters observing an equal number of payoffs on average.

Steady states of the process occur when (7) determines a new link-weighted fraction of adopters θ_t that equals the previous link-weighted fraction of adopters θ_{t-1}. To simplify notation, let

$$h_d(\theta) = \Phi\left(\frac{d\theta\mu + m}{\sqrt{d\theta\sigma^2 + s^2}}\right). \tag{8}$$

Substituting $h_d(\theta)$ into equation (4), we see that an equilibrium link-weighted fraction of adopters is a solution to

$$\theta = \frac{1}{\bar{d}}\sum_d dP(d)h_d(\theta). \tag{9}$$

Given a solution θ^* to (9), the corresponding equilibrium (unweighted) fraction of adopters can be calculated from the equation (7) for the fraction of adopters of degree d, along with the degree distribution, and is given by

$$\sum_d P(d)h_d(\theta^*). \tag{10}$$

4. Equilibria

In this section we begin by proving that there is always at least one equilibrium of the diffusion process. Then we derive conditions that guarantee a unique equilibrium when the average payoff of the new technology is less than that of the status quo. Finally, we categorize equilibria as stable or unstable and show that every set of model parameters gives rise to at least one stable equilibrium.

4.1. Existence

Our first task is to prove that an equilibrium always exists. Define

$$G(\theta) = \frac{1}{\bar{d}}\sum_d dP(d)h_d(\theta). \tag{11}$$

Fixed points of G correspond to equilibria of the diffusion process. Since G is a continuous function defined on all of $[0, 1]$, existence is a consequence of the Brouwer fixed point theorem (e.g. Massey, 1991). We will call such a function G determined by the model parameters μ, σ^2, m and s^2 and the degree distribution P a *diffusion function*.

Several aspects of the function G should be noted. First,

$$G(0) = \frac{1}{\bar{d}}\sum_d dP(d)h_d(0) = \frac{1}{\bar{d}}\sum_d dP(d)\Phi(\frac{m}{s}) = \Phi(\frac{m}{s}) > 0. \tag{12}$$

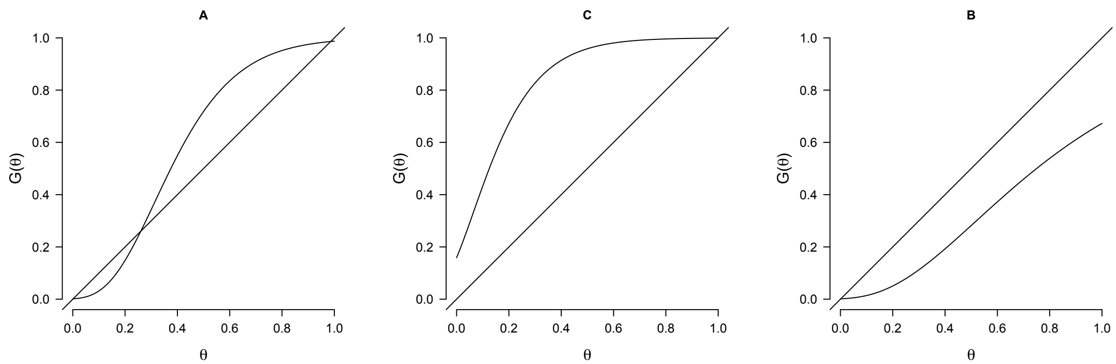

Figure 1: Three examples of a diffusion function $G(\theta)$ for a regular degree four network.

That is, $G(0)$ is simply the fraction of agents with positive priors. Thus, zero is never an equilibrium of the process. This highlights a difference between social learning and the contagion models of diffusion studied by Jackson and Yariv (2005), Jackson and Rogers (2007), and López-Pintado (2008), in which zero is always an equilibrium. Here, as soon as the technology is available some agents will adopt based solely on their prior beliefs. Similarly,

$$G(1) = \frac{1}{\bar{d}} \sum_d dP(d) h_d(1) < \frac{1}{\bar{d}} \sum_d dP(d) = 1, \tag{13}$$

so full adoption is also never an equilibrium. This stands in contrast to models of social learning in which the agents have infinite memories where typically all of the agents settle on the same action (e.g. Bala and Goyal, 1998). These observations have welfare implications. Because the payoff distribution is identical across agents and across time, the social optimum is always either no adoption or full adoption depending on the sign of the expected payoff μ. Equations (12) and (13) demonstrate that the agents never settle precisely on the optimal level of adoption. Depending on the parameter values there may be equilibria that are practically indistinguishable from no or full adoption, but in many cases there are not.

For some parameter values multiple equilibria exist. For example, as illustrated in Figure 1A, in a regular degree four network with mean payoff $\mu = 2$, payoff variance $\sigma^2 = 1$, mean prior $m = -3$, and variance of priors $s^2 = 1$, there are three equilibria: near zero, .26, or.99 (since this is a regular network, θ equals the actual fraction of adopters). Because the distribution of agents' prior beliefs is strongly biased against adoption we would expect the system to settle on the low adoption equilibrium, even though more than 97 percent of draws from the payoff distribution are positive. Unfortunately, from a social welfare perspective, without forcing many of the agents to adopt initially against their beliefs, the society will never gather enough evidence regarding the positive payoffs of the technology to overcome their skepticism and reach the high adoption equilibrium.

In other cases, only a single equilibrium exists. In the previous example, agents' were initially biased against adoption; only abut one in a thousand agents began with a positive prior. If agents have more favorable priors we observe a different pattern. Figure 1B depicts the graph of $G(\theta)$ for the same parameter values as Figure 1A except that m has been increased from -3 to -1. In this case, a greater number of agents adopt initially, and observing these early adopters' payoffs convinces much of the rest of the population to adopt. Figure 1C illustrates a third possibility for the diffusion function. The parameters that generate this curve are the same as those for Figure 1A except that now the mean payoff μ has been decreased from 2 to 1. In this case the system inevitably settles on near zero adoption.

With other network degree distributions there may be many more equilibria. Figure 2 plots an example with five equilibria. In this example $\mu = 2$, $\sigma^2 = .25$, $m = -3$ and $s^2 = .25$. In the network one percent of the nodes have degree fifty, while the rest of the nodes have an equal chance of having degree one, two, three, four, or five.

4.2. Uniqueness for Costly Technologies

When $\mu < 0$ the technology offers less utility than the status quo on average. We call such a technology *costly*. When a technology is costly, as long as the agents' priors are not overly biased against adoption, there is a unique equilibrium to the social learning process.

Theorem 1. *If* $\mu < 0$ *and*

$$2\frac{\mu}{\sigma^2} < \frac{m}{s^2}, \tag{14}$$

then a diffusion function G *with parameters* μ, σ^2, m, *and* s^2 *has a unique equilibrium regardless of the network structure.*

Proof. The reason for the unique equilibrium is that when $\mu < 0$ and (14) is satisfied, G is a decreasing function of θ and therefore has a unique fixed point. The derivative of G with respect

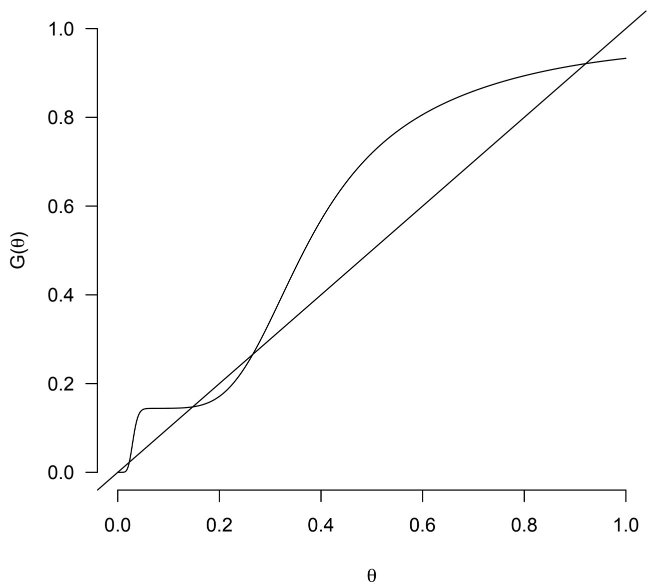

Figure 2: An example of a diffusion function $G(\theta)$ with five equilibria.

to θ is

$$G'(\theta) = \frac{1}{\bar{d}} \sum_d dP(d) \frac{\partial h_d}{\partial \theta}(\theta). \quad (15)$$

The derivative of h_d with respect to θ is

$$\frac{\partial h_d}{\partial \theta}(\theta) = d\Phi' \left(\frac{d\theta\mu + m}{\sqrt{d\theta\sigma^2 + s^2}} \right) \frac{d\theta\mu\sigma^2 + 2\mu s^2 - \sigma^2 m}{2(d\theta\sigma^2 + s^2)^{3/2}}. \quad (16)$$

Thus the sign of $\frac{\partial h_d}{\partial \theta}(\theta)$ and therefore of $G'(\theta)$ is the same as the sign of $d\theta\mu\sigma^2 + 2\mu s^2 - \sigma^2 m$. When $\mu < 0$,

$$d\theta\mu\sigma^2 + 2\mu s^2 - \sigma^2 m < 2\mu s^2 - \sigma^2 m, \quad (17)$$

which is negative whenever the condition (14) is satisfied. □

This theorem illustrates the asymmetry between costly and beneficial technologies. For costly technologies, an external shock that adds more adopters tends to be countered by a decrease in adoption as those new adopters learn and communicate that the technology is costly. This "negative feedback loop" tends to bring the system to equilibrium. For beneficial technologies, the system can come to rest at an equilibrium in which more agents would adopt if they knew about the benefits of adopting, but too few agents are currently adopting in order for the group to learn about those benefits. An external shock that adds more adopters increases the number of agents who know about the benefits, who in turn can communicate that knowledge to their neighbors leading to still further adoption. This results in a "positive feedback loop," which can move the system towards a higher adoption equilibrium.[6]

4.3. Stable and Unstable Equilibria

The mean-field analysis in the previous section identifies equilibria of the social learning process. In practice, randomness makes it unlikely for some of these equilibria to be maintained. For example, because agents' experiences are random draws from the payoff distribution, at any particular time or particular region of the network these draws will fluctuate around the true mean of the distribution in turn leading to fluctuations in the actual adoption pattern. Figure 3 illustrates this in the evolution of one simulated realization of the social learning process in a regular degree five network with 300 agents (with $\mu = 1$, $\sigma = 1$, $m = -1$, and $s = 1$). The adoption fraction in

[6] For discussions of positive and negative feedbacks and multiple equilibria see Arthur (1996) and Sterman (2000).

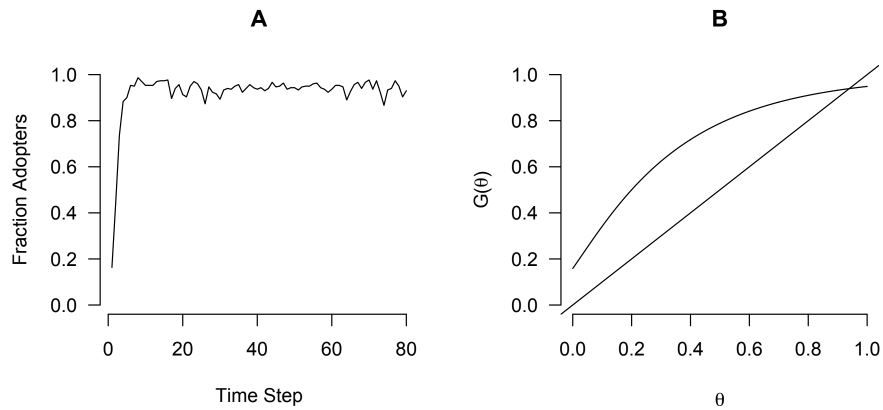

Figure 3: One realization of the social learning process and the corresponding diffusion function $G(\theta)$ in a regular degree five network. The adoption level converges to, and then fluctuates around, the unique stable equilibrium.

the simulated population quickly approaches the predicted equilibrium of 94% adoption and then fluctuates around that level.

In some cases, small perturbations away from an equilibrium tend to be countered by perturbations back towards the equilibrium as in Figure 3. We call these equilibria *stable*. In other cases, once a perturbation knocks the process off of the equilibrium the adoption path tends to diverge away from that equilibrium towards another. We call these equilibria *unstable*.

Stable and unstable equilibria can be identified by the local form of the function $G(\theta)$ near a fixed point. If the link-weighted fraction of adopters is θ and $G(\theta) > \theta$, then the link-weighted fraction of adopters will tend to increase. Conversely, if $G(\theta) < \theta$, then the link-weighted fraction of adopters will tend to decrease. This leads us to the following definition.

Definition 1. *A fixed point θ_s^* of G is a stable equilibrium if there exists an $\epsilon > 0$ such that for any $\theta \in (\theta_s^* - \epsilon, \theta_s^*)$, $G(\theta) > G(\theta_s^*)$ and for any $\theta \in (\theta_s^*, \theta_s^* + \epsilon)$, $G(\theta) < G(\theta_s^*)$. A fixed point θ_u^* of G is an unstable equilibrium if there exists an $\epsilon > 0$ such that for any $\theta \in (\theta_u^* - \epsilon, \theta_u^*)$, $G(\theta) < G(\theta_u^*)$ and for any $\theta \in (\theta_u^*, \theta_u^* + \epsilon)$, $G(\theta) > G(\theta_u^*)$ (c.f. Definition 1 of Jackson and Yariv, 2007).*[7]

[7] A fixed point θ^* of G may also be a degenerate fixed point, which under definition 1 is neither stable nor unstable, if $G'(\theta^*) = 1$. However, generically all fixed points of G are either stable or unstable. By this we mean that any G

For example, the fixed points in both Figure 1B and Figure 1C are stable. The smallest and largest fixed points in Figure 1A are stable, while the middle fixed point in Figure 1A is unstable. For the most part we are more interested in stable equilibria than unstable equilibria because it is highly unlikely that the stochastic process will settle on an unstable equilibrium. Instead a realization of the model will tend to hover near a stable equilibrium as in Figure 3.

The following theorem collects several observations on stable and unstable equilibria, which can be proven using simple Intermediate Value Theorem arguments along with the facts that G is continuous, $G(0) > 0$ and $G(1) < 1$.

Theorem 2. *Consider a diffusion function G as in equation (11). Then the set of equilibria for G satisfy the following:*

1. *There is at least one stable equilibrium.*
2. *The smallest equilibrium is stable.*
3. *The largest equilibrium is stable.*
4. *The ordered set of equilibria alternates between stable and unstable equilibria.*

We are interested in how stable equilibria depend on the parameters of our model and the network structure; however, when there are multiple equilibria it is unclear what it means for certain parameters to generate more or less diffusion. To better describe this dependance we define a function $\phi_G : [0,1] \to (0,1)$, which we call the *equilibrium function* of the diffusion function G. For any $\theta \in [0,1]$ with $G(\theta) < \theta$ let $\phi_G(\theta)$ be the largest stable equilibrium of G that is less than θ. For any $\theta \in [0,1]$ with $G(\theta) \geq \theta$ let $\phi_G(\theta)$ be the smallest stable equilibrium of G that is greater than or equal to θ. The idea of the equilibrium function is that if we begin the social learning process specified by G with a link-weighted fraction of adopters θ then we expect the process to converge to near the stable equilibrium $\phi_G(\theta)$.[8]

with a degenerate fixed point θ^*can be transformed by an arbitrarily small perturbation into a function without a degenerate fixed point and for any G without a degenerate fixed point there exists an $\epsilon > 0$ such that perturbations of G that do not change values of G by more than ϵ have no degenerate fixed points. For the remainder of the paper we assume that G has no degenerate fixed points.

[8]When θ is an unstable equilibrium the choice to set $\phi_G(\theta)$ to be the next largest stable equilibrium as opposed to the next smaller stable equilibrium is arbitrary.

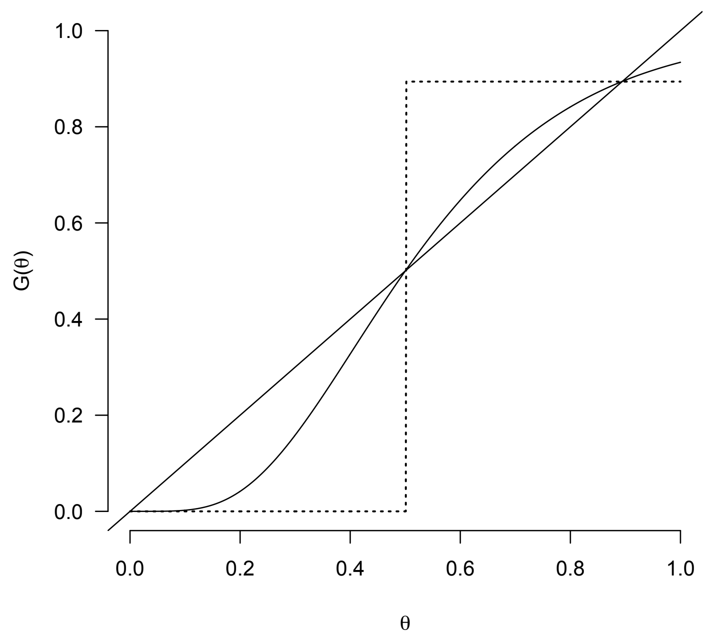

Figure 4: An example of a diffusion function $G(\theta)$ (solid) and the associated equilibrium function $\phi_G(\theta)$ (dashed).

Definition 2. *A diffusion function G* generates greater diffusion *than a diffusion function $\tilde{G}$ if $\phi_G(x) \geq \phi_{\tilde{G}}(x)$ for all $x \in [0,1]$ (c.f. Definition 3 of Jackson and Yariv, 2007).*

Intuitively, a diffusion function G generates more diffusion than another $\tilde{G}$ if, regardless of the initial fraction of adopters, we expect the process specified by G to converge to an equilibrium with a greater fraction of adopters than that specified by $\tilde{G}$.

5. Comparative Statics

In this section we examine how changes in the model parameters and the social network affect equilibrium levels of diffusion.

5.1. Dependance on Model Parameters

First, we examine how the stable equilibria of a diffusion function G depend on the non-network parameters of the model, μ, σ^2, m, and s^2. We begin with the following lemma, which shows that changes that increase the diffusion function lead to greater diffusion.

Lemma 1. *If $G(\theta) \geq \tilde{G}(\theta)$ for all θ then G generates greater diffusion than $\tilde{G}$ (c.f. Proposition 1 of Jackson and Yariv, 2007).*

Proof. For $t \in [0,1]$, define $\varphi_t : [0,1] \to [0,1]$ by

$$\varphi_t(x) = G(x) + t(\tilde{G}(x) - G(x)). \tag{18}$$

So $\varphi_0(x) = G(x)$ and $\varphi_1(x) = \tilde{G}(x)$ (i.e. φ_t is a homotopy from G to $\tilde{G}$). We can determine how $\phi_{\tilde{G}}$ relates to ϕ_G by examining how solutions to $\varphi_t(x) = x$ change as t moves from zero to one. We extend the definition of stable and unstable equilibria to stable and unstable fixed points of $\varphi_t(x)$ in the obvious way and for each t define a function $\phi_t(x)$ corresponding to $\varphi_t(x)$ in the same way that $\phi_G(x)$ is defined from G. Small increases in t can result in three changes in the ordered set of fixed points of φ_t:

A. Stable fixed points increase and unstable fixed points decrease.

B. A stable fixed point and the next highest unstable fixed point vanish.

C. A new unstable fixed point and consecutive stable fixed point are introduced.

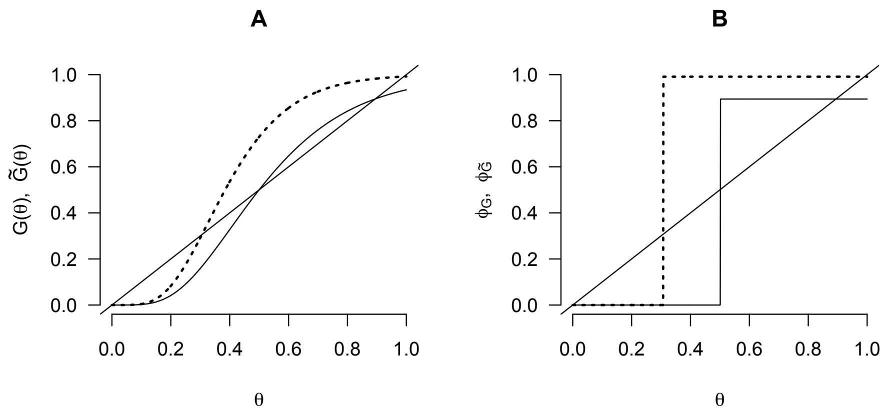

Figure 5: Two diffusion functions (A) and the associated equilibrium functions (B). The process illustrated by the dotted line has a higher mean payoff μ than that illustrated by the solid line.

It is straightforward to check that each of these three changes causes an increase in $\phi_t(x)$. Thus, $\varphi_1 = \phi_{\tilde{G}} \geq \varphi_0 = \phi_G$. □

Figure 5 illustrates a specific example of Lemma 1. Panel A on the left side of the figure plots two diffusion functions G and $\tilde{G}$ with $\tilde{G}(\theta) \geq G(\theta)$ for all θ. Panel B on the right plots the associated equilibrium functions. As we can see, $\phi_{\tilde{G}}(\theta) \geq \phi_G(\theta)$ for all θ.

Any change in parameters that increases the values of the function $h_d(\theta)$ also increases the value of the function $G(\theta)$. Returning to the definition of $h_d(\theta)$ from equation (8) we see that $h_d(\theta)$ increases when μ increases or m increases. Thus, we have proved the following theorem.

Theorem 3. *Increasing the mean of the payoff distribution, μ, or the mean of the distribution of priors, m, generates greater diffusion.*

The example depicted in Figure 5 is generated by a mean shift as described in Theorem 3. The solid lines plot the diffusion function G and the associated equilibrium function ϕ_G with $\mu = 1$, $\sigma^2 = 1$, $m = -5$ and $s = 1$ for a regular degree ten network. The dotted lines show the diffusion function $\tilde{G}$ and the associated equilibrium function $\phi_{\tilde{G}}$ which has the same parameters and network as for G but with μ increased to 1.3.

The effects of changes in σ^2 and s^2 are conditional on the sign of $d\theta\mu + m$. Since $d\theta > 0$, if μ

and m are either both positive or both negative, then the sign of $d\theta\mu + m$ also positive or negative, respectively. Thus, if both μ and m are positive then $h_d(\theta)$ increases when σ^2 decreases or s^2 decreases. This proves the following theorem.

Theorem 4. *If both μ and m are positive, decreasing the variance of the payoff distribution σ^2 or the variance of the distribution of priors s^2 generates greater diffusion. If both μ and m are negative, increasing the variance of the payoff distribution σ^2 or the variance of the distribution of priors s^2 generates greater diffusion.*

The results in Theorem 3 and 4 do not depend on the network structure. The same relationships would hold if there was no structure to agent interactions. However, when μ and m have opposite signs, the effect of increasing or decreasing the variance in the payoff or prior distributions depends on the degree distribution of the network. Depending on the network, changing the variance of the payoff or prior distribution can generate greater or less diffusion or have an ambiguous effect.

5.2. Dependance on Network Structure

This section examines the effect of changes in the network structure, as specified by the degree distribution P, on the extent of the technology diffusion. We examine the effects of two types of changes in the network: first and second order stochastic dominance shifts in the degree distribution (Rothschild and Stiglitz, 1970). A distribution P *strictly first order stochastically dominates* a distribution $\tilde{P}$ if for every nondecreasing function $u : \mathbb{R} \to \mathbb{R}$,

$$\sum_{d=0}^{D_{max}} u(d)\tilde{P}(d) < \sum_{d=0}^{D_{max}} u(d)P(d), \tag{19}$$

where D_{max} is the maximum degree of any node in the network. A distribution P *strictly second order stochastically dominates* a distribution $\tilde{P}$ if for every nondecreasing concave function $u : \mathbb{R} \to \mathbb{R}$,

$$\sum_{d=0}^{D_{max}} u(d)\tilde{P}(d) < \sum_{d=0}^{D_{max}} u(d)P(d). \tag{20}$$

First order stochastic dominance implies second order stochastic dominance, but not vice versa. If P and $\tilde{P}$ have the same mean, then the statement P second order stochastically dominates $\tilde{P}$ is equivalent to the statement $\tilde{P}$ is a *mean preserving spread* of P. Intuitively, one network first order stochastically dominates another if agents have more neighbors in the former than the latter. A network second order stochastically dominates another if there is less heterogeneity in the number of neighbors that agents have in the former than the latter.[9]

In our case, the role of the function u in equations (19) and (20) is played by $h_d(\theta)$ and the role of the distribution P is played by $dP/\bar{d} = dP/\mathrm{E}_P[d]$. In order to understand the consequences of stochastic shifts in the degree distribution, we need to understand when h is increasing and decreasing as well as its concavity when viewed as a function of d. Since throughout this section we will be interested in $h_d(\theta)$ as a function of d, we will abuse notation by suppressing the dependance on θ and simply write $h(d)$ for $h_d(\theta)$, $h'(d)$ for $\frac{\partial h_d(\theta)}{\partial d}$ and so on. Examining the first derivative of h, we see that

$$h'(d) = \theta\Phi'\left(\frac{d\theta\mu + m}{\sqrt{d\theta\sigma^2 + s^2}}\right)\frac{d\theta\mu\sigma^2 + 2\mu s^2 - \sigma^2 m}{2(d\theta\sigma^2 + s^2)^{3/2}}. \tag{21}$$

Thus the sign of $h'(d)$ depends on the sign of

$$d\theta\mu\sigma^2 + 2\mu s^2 - \sigma^2 m. \tag{22}$$

If $d\theta\mu\sigma^2 + 2\mu s^2 - \sigma^2 m > 0$ then $h'(d)$ is positive. Suppose that $\mu > 0$. Then

$$d\theta\mu\sigma^2 + 2\mu s^2 - \sigma^2 m \geq 2\mu s^2 - \sigma^2 m, \tag{23}$$

since $d\theta\mu\sigma^2 \geq 0$. The right hand side of (23) is greater than zero when

$$2\frac{\mu}{\sigma^2} > \frac{m}{s^2}. \tag{24}$$

So, when $\mu > 0$ and (24) holds, h is an increasing function of d for any $\theta > 0$. A similar argument shows that when $\mu < 0$ and

$$2\frac{\mu}{\sigma^2} < \frac{m}{s^2} \tag{25}$$

h is a decreasing function of d. In this case, Theorem 1 guarantees that there is a unique equilibrium level of diffusion in both P and $\tilde{P}$. Combining this with the definition of first order stochastic dominance and Lemma 1 proves:

Theorem 5. *Suppose that $dP/\mathrm{E}_P[d]$ strictly first order stochastically dominates $d\tilde{P}/\mathrm{E}_{\tilde{P}}[d]$. If $\mu > 0$ (i.e. on average adopting the technology is beneficial) and*

$$2\frac{\mu}{\sigma^2} > \frac{m}{s^2}, \tag{26}$$

[9]For an introduction to stochastic dominance and its role in network diffusion see Jackson (2008) or Lamberson (2009).

then P generates greater diffusion than $\tilde{P}$. If $\mu < 0$ (i.e. on average adopting the technology is costly) and

$$2\frac{\mu}{\sigma^2} < \frac{m}{s^2}, \tag{27}$$

then the unique equilibrium level of diffusion in the network with degree distribution $\tilde{P}$ is greater than the unique equilibrium level of diffusion in the network with degree distribution P.

We can think of a network that first order stochastically dominates another as providing the agents with more information, since on average the agents have more links to other agents. We would expect that for beneficial technologies, more information would aid diffusion. Theorem 5 confirms this intuition, but only if the agents are not overly optimistic about the technology to begin with, as captured by condition (26). If the agents' prior beliefs about the payoffs of the technology are sufficiently positive, so that (26) is violated, adding more links to the network can hinder diffusion. This stands in contrast to contagion models in which adding links always aids diffusion (Jackson, 2008; López-Pintado, 2008).

On reflection, we might expect that when agents' priors tend to be more positive than the payoffs, adding links could decrease diffusion. That logic leads to a condition that says if the fraction of payoffs that are positive is greater than the fraction of agents with positive priors, i.e.

$$\frac{\mu}{\sigma^2} > \frac{m}{s^2}, \tag{28}$$

then first order stochastic shifts lead to greater diffusion. But the actual condition (26) is more subtle. The intuitive condition (28) differs from the actual condition (26) by a factor of two on the left hand side. If we fix the distribution of priors, and consider (26) as a condition on the payoffs, then the actual condition (26) is weaker than the intuitive condition (28). In other words, relative to the distribution of priors, the payoff distribution contributes more to the marginal effect of degree on diffusion than we might expect.

This discrepancy arises because there is a non-trivial interaction between the effect of adding links to the network and of changing the payoff distribution due to the fact that only adopting agents can communicate payoff information. Increasing the payoffs increases the number of adopting agents, which makes the effect of adding links stronger because those additional links are more likely to connect to agents that have payoff information to share. Conversely, decreasing the payoff distribution weakens the effect of adding links, because those additional links are more likely to connect to non-adopting agents who do not contribute any additional information.

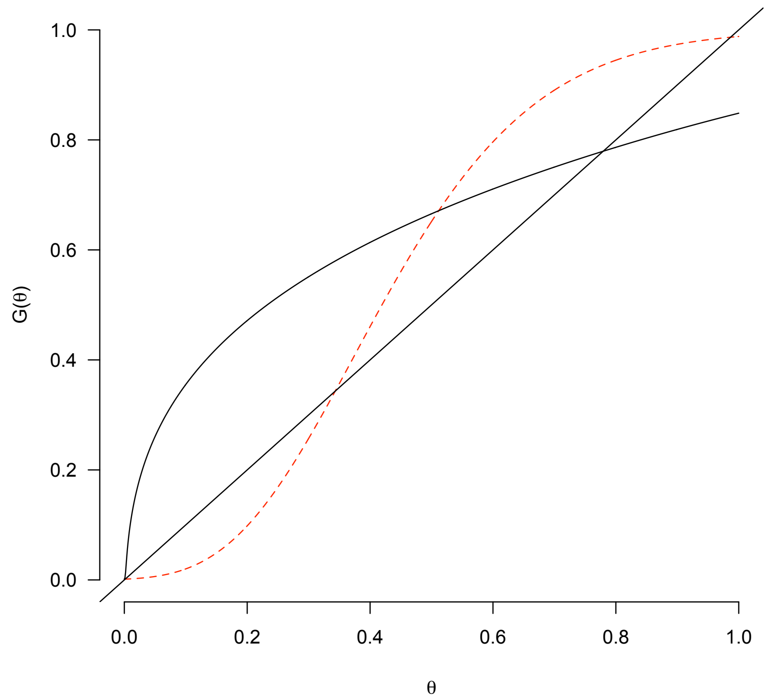

Figure 6: A diffusion function with the same model parameters in two different networks with the same average degree: a regular network (dashed line) and a scale-free network (solid line).

Turning to second order stochastic dominance shifts, we have a similar theorem:

Theorem 6. *Suppose that* $dP/\mathrm{E}_P[d]$ *strictly second order stochastically dominates* $d\tilde{P}/\mathrm{E}_{\tilde{P}}[d]$*. If*

$$h''(d) > 0, \tag{29}$$

then P *generates greater diffusion than* $\tilde{P}$*. If*

$$h''(d) < 0, \tag{30}$$

then $\tilde{P}$ *generates greater diffusion than* P*.*

In Theorem 5 we were able to express conditions (26) and (27) in terms of the social learning parameters in an interpretable fashion. In the case of second order stochastic changes in the degree distribution, as examined in Theorem 6, the analogous conditions become too complex to decipher when written out in terms of the model parameters.[10] Moreover, in many cases second order stochastic shifts in the degree distribution do not have a consistent effect on diffusion because h is convex for some values of θ and concave for others.

This highlights another difference between network diffusion via social learning and via an epidemic model as considered by Jackson and Rogers (2007) or López-Pintado (2008). In the model by Jackson and Rogers (2007), for example, a second order stochastically dominant degree distribution always has a lower highest stable equilibrium. This holds because in the epidemic model the effect of adding edges is convex, essentially because adding a link to an agent increases both the chances that she becomes infected and the chances that she spreads the infection. In the social learning model the contribution of adding edges depends on the level of adoption, the distribution of payoffs, and the distribution of prior beliefs as well as the degree distribution.

Figure 6 illustrates the phenomenon. The figure plots the diffusion function G with the same model parameters ($\mu = 3$, $\sigma^2 = 1$, $m = -3$ and $s^2 = 1$) for a regular network (dashed line) and a scale-free network (i.e. one with a power law degree distribution, solid line).[11] The regular network second order stochastically dominates the scale-free network, but both have the same average degree. Despite having the same average degree, these two degree distributions generate vastly different dynamics. The scale free network has a single equilibrium link-weighted fraction of adopters of 87.9%, which by equation (10) corresponds to an actual adoption fraction of only 49.6%. Regardless of the initial fraction of adopters, in the scale-free network we would expect the process to converge to near 49.6% adoption. The regular network gives rise to two stable equilibria, one nearly indistinguishable from no adoption and another at approximately 98.6% adoption, as well as one unstable equilibrium at 34.3%. For this network, unless the fraction of adoption is exogenously pushed beyond the unstable equilibrium at 34.3% adoption, the process settles on the equilibrium near zero. However, if the population begins at a point above the unstable equilibrium,

[10] The second derivative of h with respect to d is

$$\frac{\theta^2}{4}\left[\Phi''\left(\frac{d\theta\mu + m}{\sqrt{d\theta\sigma^2 + s^2}}\right)\frac{(d\theta\mu\sigma^2 + 2\mu s^2 - \sigma^2 m)^2}{(d\theta\sigma^2 + s^2)^3} + \Phi'\left(\frac{d\theta\mu + m}{\sqrt{d\theta\sigma^2 + s^2}}\right)\frac{-\sigma^2(d\theta\mu\sigma^2 + 4\mu s^2 - 3\sigma^2 m)}{(d\theta\sigma^2 + s^2)^{5/2}}\right]. \tag{31}$$

[11] For this computation the maximum degree is fixed at 500. The power law exponent is 2.3, the same as the exponent in the network of movie actors measured by Barabasi and Albert (1999).

it then moves to the equilibrium at 98.6%, which is 49% higher than the equilibrium in the scale-free network. Thus, depending on the initial adoption level, the regular network, which second order stochastically dominants the scale-free network, can generate more or less diffusion.

6. Memory

Up to this point, agents' adoption and disadoption decisions in the model are based solely on payoffs from the previous period. In this respect, our model follows those considered by Ellison and Fudenberg (1993, 1995). Ellison and Fudenberg (1993, p. 618) justify this assumption by supposing that "individual players revise their choices too infrequently to want to keep track of each period's results and, more strongly, that the market at this particular 'location' is too small for a record-keeping agency to provide this service." While individual agents have no memory of the past beyond the previous period, in some sense the population retains a memory of the past in the form of the overall fraction of adopters which in turn influences future adoption decisions.

In contrast, the model by Bala and Goyal (1998) allows agents to have infinite memories. At each stage in Bala and Goyal's model the agents update their priors based on new observations, and their new posterior becomes the prior for the following round. As described in the introduction, the finite and infinite memory cases are qualitatively different. In the infinite memory case, the population tends towards conformity, while the finite memory case considered here always maintains some diversity.

While we take the one period memory approach of Ellison and Fudenberg in our analysis above, we now extend the model to incorporate finite memory of arbitrary length. Suppose that agents base their adoption decision on observations of payoffs from the previous k periods.[12] Then, in equation (5) we would replace $d\theta_{t-1}$ with $\sum_{j=1}^{k} d\theta_{t-j}$. Carrying through the mean-field analysis this leads to a new definition of the function $h_d(\theta)$ in equation (8),

$$h_d(\theta) = \Phi\left(\frac{dk\theta\mu + m}{\sqrt{dk\theta\sigma^2 + s^2}}\right). \tag{32}$$

None of the comparative statics analyzed in Theorems 3 and 4 are affected by this change. Furthermore, differentiating h with respect to d as in equation (21), we obtain

$$h'(d) = \theta k\Phi'\left(\frac{d\theta k\mu + m}{\sqrt{d\theta k\sigma^2 + s^2}}\right)\frac{d\theta k\mu\sigma^2 + 2\mu s^2 - \sigma^2 m}{2(d\theta k\sigma^2 + s^2)^{3/2}}. \tag{33}$$

[12] We implicitly assume that each agent observes at least k periods worth of payoffs before updating her decision.

Following the analysis in section 5.2, adding the memory parameter k also has no effect on the conditions (26) and (27) or on the relationship between network structure and diffusion described in Theorem 5 or 6.

While the inclusion of memory does not change any of the other comparative statics, it does itself have an effect on the equilibrium. As is evident from equation (32), the role of the memory term is similar to the role of degree. As with changes in degree, increases in the memory parameter k cause increases in h when $\mu > 0$ and (26) is satisfied. Increases in k cause h to decrease when $\mu < 0$ and (27) is satisfied.

Combining these observations and applying Lemma 1 proves the following theorem.

Theorem 7. *Consider two diffusion functions G and $\tilde{G}$ with all of the same parameters with the exception that the memory parameter k for G is greater than the memory parameter $\tilde{k}$ for $\tilde{G}$. If $\mu > 0$, so on average adopting the technology is beneficial, and*

$$2\frac{\mu}{\sigma^2} > \frac{m}{s^2}, \tag{34}$$

then G generates greater diffusion than $\tilde{G}$. If $\mu < 0$, so on average adopting the technology is costly, and

$$2\frac{\mu}{\sigma^2} < \frac{m}{s^2}, \tag{35}$$

then $\tilde{G}$ generates greater diffusion than G.

So, the condition and direction of the effect of increases in memory on diffusion are the same as for first order stochastic shifts in the degree distribution.

We can also ask, what happens in the limit as k approaches infinity? If $\mu > 0$ then as k approaches infinity G approaches one. Conversely, if $\mu < 0$ then G approaches zero. Thus, in the infinite memory limit the population converges to the social optimum: all agents adopt if the technology has a positive average payoff and no agents adopt if the technology has a negative average payoff. In the infinite memory model of Bala and Goyal (1998), the population always converges to a consensus, but that consensus may not be the optimal one. The reason for the discrepancy between our results and theirs lies in the distribution of prior beliefs. Their model allows for the possibility, for example, that all agents are sufficiently biased against adoption of a technology that none of them ever try it. However, if at least one agent has a sufficiently positive prior when $\mu > 0$, or a sufficiently negative prior when $\mu < 0$, then the population in the Bala and Goyal model also converges to the "correct" equilibrium. Because the model here assumes a normal distribution of prior beliefs and the mean-field approximation assumes an infinite population, this condition is always satisfied. Thus, taking the limit as k approaches infinity, the model here reproduces the results of Bala and Goyal (1998) under the assumption that the distribution of agents' priors has sufficiently wide support.

7. Conclusion

In this paper we analyze a model of social learning in a social network. The paper contributes to two streams of literature – the literature on social learning as a mechanism for diffusion and the literature on diffusion in social networks – which were until now largely separate.[13] Incorporating social network structure in a standard social learning model adds realism; we would expect that individuals seek information from their friends and family. We prove that adding this network structure affects the diffusion. To the diffusion literature, the model presented here adds a microeconomic rationale for agents' decisions, as opposed to a simple contagion or threshold model. Not surprisingly, we find that the collective dynamics of rational actors are more complex than the physics of disease spread. For example, in a contagion model, first order stochastic shifts in the degree distribution always increase diffusion (Jackson and Rogers, 2007). In contrast, in the model presented here, the effect of a first order stochastic shift depends on the payoffs to adoption and the agents' prior beliefs regarding those payoffs. We derive precise conditions for the relationships between first and second order stochastic shifts in the degree distribution and equilibrium levels of diffusion. In some cases we find these conditional effects surprising. For example, adding links to a network can decrease diffusion even when the social optimum is for all agents to adopt.

To analyze this model we employ a mean-field approximation, which requires assumptions that may not always be appropriate. For example, the approximation results are likely to be less accurate in small networks or networks in which the degrees of neighboring agents are highly correlated. However, in many cases simulation results confirm that mean-field techniques provide a good approximation to discrete dynamics (e.g. Newman and Watts, 1999; Newman et al., 2000; Newman, 2002).

[13] The papers by Bala and Goyal (1998, 2001) are notable exceptions.

The model and analysis employed in this paper open the door to the exploration of other questions. First, in the model presented here, while agents' prior beliefs differ their preferences do not. Extending this model to include heterogenous preferences is a logical next step. We may also consider the possibility that those preferences are correlated with agents' positions in the network to reflect the fact that agents are more likely to have social ties with agents that are similar to them (i.e. the network exhibits homophily (McPherson et al., 2001)). Second, one could add a dynamic to the "supply side" of the model to investigate how the results may be affected if the payoffs to the new technology changed over time or if multiple technologies competed for market share. The model raises the possibility of using information on network structure to tailor firm strategies to specific network contexts. Finally, the model offers a potential explanation for why technologies and behaviors may diffuse to a greater extent in one community than another. This could provide the basis for an empirical test of the model's predictions and help us to better understand the mechanisms of diffusion and the role of social structure in the process.

References

Arthur, W. Brian, "Increasing Returns and the New World of Business," *Harvard Business Review*, July-August 1996.

Bala, V. and S. Goyal, "Learning from Neighbours," *Review of Economic Studies*, 1998, *65*, 595–621.

_ **and** _ , "Conformism and Diversity under Social Learning," *Economic Theory*, 2001, *17* (1), 101–120.

Banerjee, A. and D. Fudenberg, "Word-of-mouth learning," *Games and Economic Behavior*, 2004, *46* (1), 1–22.

Banerjee, Abhijit, "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 1992, *107*, 797–817.

Barabasi, A.L. and R. Albert, "Emergence of Scaling in Random Networks," *Science*, 1999, *286*, 509–512.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *The Journal of Political Economy*, 1992, *100* (5), 992–1026.

Burt, R.S., "Social Contagion and Innovation: Cohesion Versus Structural Equivalence," *American Journal of Sociology*, 1987, *92* (6), 1287.

Centola, D. and M. Macy, "Complex Contagions and the Weakness of Long Ties," *American Journal of Sociology*, 2007, *113* (3), 702–734.

Chamley, C., *Rational Herds: Economic Models of Social Learning*, Cambridge: Cambridge Univ Press, 2003.

Chandler, D., *Introduction to Modern Statistical Mechanics*, New York: Oxford, 1987.

Chatterjee, K. and S.H. Xu, "Technology Diffusion by Learning from Neighbours," *Advances in Applied Probability*, 2004, pp. 355–376.

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

Christakis, Nicholas A. and James H. Fowler, "The Colletive Dynamics of Smoking in a Large Social Network," *The New England Journal of Medicine*, 2008, *358*, 2248–2258.

Coleman, J.S., E. Katz, and H. Menzel, *Medical Innovation: A Diffusion Study*, Bobbs-Merrill Co, 1966.

Conley, T. and C. Udry, "Learning about a New Technology: Pineapple in Ghana," *Proceedings of the Federal Reserve Bank of San Francisco*, 2005.

DeGroot, Morris H., *Optimal Statistical Decisions*, New York: McGraw-Hill, 1970.

Ellison, G. and D. Fudenberg, "Rules of Thumb for Social Learning," *Journal of Political Economy*, 1993, *101* (4), 612–643.

Ellison, Glenn and Drew Fudenberg, "Word-of-mouth Communication and Social Learning," *The Quarterly Journal of Economics*, 1995, *110* (1), 93–125.

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

Fowler, J.H. and N.A. Christakis, "Dynamic Spread of Happiness in a Large Social Network: Longitudinal Analysis over 20 years in the Framingham Heart Study," *British Medical Journal*, 2008, *337* (dec04 2), a2338.

Galeotti, A., S. Goyal, Matthew O. Jackson, F. Vega-Redondo, and L. Yariv, "Network Games," *Review of Economic Studies*, 2009, *Forthcoming.*

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin, *Bayesian Data Analysis*, second ed., Boca Raton, FL: Chapman and Hall/CRC, 2004.

Golub, Benjamin and Matthew O. Jackson, "Naive Learning in Social Networks: Convergence, Influence and the Wisdom of Crowds," *American Economic Journal Microeconomics*, 2009, *Forthcoming.*

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

_ **and Brian W. Rogers**, "Relating Network Structure to Diffusion Properties through Stochastic Dominance," *The B.E. Journal of Theoretical Economics*, 2007, *7* (1 (Advances)).

_ **and L. Yariv**, "Diffusion on Social Networks," *Économie publique*, 2005, *16*, 69–82.

_ **and** _ , "Diffusion of Behavior and Equilibrium Properties in Network Games," *American Economic Review*, 2007, *97* (2), 92–98.

Kapur, S., "Technological Diffusion with Social Learning," *The Journal of Industrial Economics*, 1995, *43* (2), 173–195.

Kirman, A., "Ants, Rationality, and Recruitment," *The Quarterly Journal of Economics*, 1993, *108* (1), 137–156.

Lamberson, P J, "Linking Network Structure and Diffusion Through Stochastic Dominance," *MIT Sloan Research Paper*, 2009, *No. 4760-09.*

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

Manski, C.F., "Social Learning from Private Experiences: the Dynamics of the Selection Problem," *Review of Economic Studies*, 2004, *71* (2), 443–458.

Massey, William S., *A Basic Course in Algebraic Topology*, Vol. 127 of *Graduate Texts in Mathematics*, New York: Springer, 1991.

McPherson, M., L. Smith-Lovin, and J.M. Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, 2001, *27* (1), 415–444.

Munshi, K., "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution," *Journal of Development Economics*, 2004, *73* (1), 185–213.

Newman, M. E. J., "The Spread of Epidemic Disease on Networks," *Physical Review E*, 2002, *66*, 016128.

_ and Duncan J. Watts, "Scaling and Percolation in the Small-world Network Model," *Physical Review Volume E*, 1999, *60* (6), 7332–7342.

_ , C. Moore, and Duncan J. Watts, "Mean-field Solution of the Small-world Network Model," *Physical Review Letters*, 2000, *84* (14), 3201–3204.

Nickerson, D.W., "Is Voting Contagious? Evidence from Two Field Experiments," *American Political Science Review*, 2008, *102* (01), 49–57.

Pastor-Satorras, R. and A. Vespignani, "Epidemic Dynamics and Endemic States in Complex Networks," *Physical Review E*, 2001, *63* (6).

_ and _ , "Epidemic Spreading in Scale-free Networks," *Physical Review Letters*, 2001, *86* (14), 3200–3203.

Rothschild, M. and J. Stiglitz, "Increasing Risk: I. A Definition," *Journal of Economic Theory*, 1970, *2*, 225–243.

Sander, LM, CP Warren, IM Sokolov, C. Simon, and J. Koopman, "Percolation on Heterogeneous Networks as a Model for Epidemics," *Mathematical Biosciences*, 2002, *180* (1-2), 293–305.

Scharfstein, D.S. and J.C. Stein, "Herd Behavior and Investment," *The American Economic Review*, 1990, *80* (3), 465–479.

Smith, L. and P. Sørensen, "Pathological Outcomes of Observational Learning," *Econometrica*, 2000, *68* (2), 371–398.

Sterman, J.D., *Business Dynamics*, Irwin/McGraw-Hill, 2000.

Watts, D.J. and S.H. Strogatz, "Collective Dynamics of 'Small-world' Networks," *Nature*, 1998, *393*, 440–442.

Young, H. Peyton, "The Spread of Innovations by Social Learning," *Working Paper*, 2006.

_ , "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning," *American Economic Review*, 2009, *Forthcoming.*