Computational Economics (2005) 25: 3–23
DOI: 10.1007/s10614-005-6245-1

Strategies for the Diffusion of Innovations on Social Networks

FLOORTJE ALKEMADE[1] and CAROLINA CASTALDI[2]
[1]*Department of Innovation Studies, Utrecht University, P.O. Box 80125, 3508 TC Utrecht, The Netherlands; E-mail: f.alkemade@geo.uu.nl*
[2]*Department of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands*

Accepted 25 April 2005

Abstract. We investigate the spread of innovations on a social network. The network consists of agents that are exposed to the introduction of a new product. Consumers decide whether or not to buy the product based on their own preferences and the decisions of their neighbors in the social network. We use and extend concepts from the literature on epidemics and herd behavior to study this problem. The central question of this paper is whether firms can learn about the network structure and consumer characteristics when only limited information is available, and use this information to evolve a successful directed-advertising strategy. In order to do so, we extend existing models to allow for heterogeneous agents and both positive and negative externalities. The firm can learn a directed-advertising strategy that takes into account both the topology of the social consumer network and the characteristics of the consumer. Such directed-advertising strategies outperform random advertising.

Key words: agents, learning, heterogeneous agents

1. Introduction

Network economics holds the view that individual actions and, in turn, aggregate outcomes, are in large part determined by the interaction structure between heterogeneous economic agents. The central question of this paper is whether firms can learn about social network structure and consumer characteristics when only limited information is available, and use this information to evolve a successful marketing strategy. We consider the decision problem of a firm that wishes to choose an advertising strategy to successfully introduce a new product in a network of consumers. There are in fact situations, which we discuss below, when consumers base their purchase decision on the behavior of other consumers. Under these circumstances, it may be of critical importance for a firm to get insights into the structure of the social network. However, marketing research charting consumer relations is typically expensive and difficult to perform (Scott, 2000; Wasserman and Faust, 1994). We investigate whether firms can learn targeted-advertising strategies, taking the social network structure into account, if only aggregate sales data are available. Is such a strategy of targeted advertising more effective than a random advertising strategy?

And does the best strategy change with respect to different topologies of networks? To address these questions we use insights from the existing literature on diffusion phenomena in networks, both in economics and epidemiology. We extend existing models to allow for more realistic modeling of consumer behavior and we study the diffusion of the innovation through agent-based simulation.

The simulation model allows us to study how word of mouth about an innovation (a new product or idea) spreads throughout a social network. More specifically, we look at the situation where the decisions of the consumers are strongly determined by the decisions of their neighbors in the network (cf. Cowan and Jonard, 2004 for a discussion of various empirical examples). There are two major situations in which this type of behavior can be considered a good strategy for consumers. First, when agents possess no reliable information about the new good, they look to the consumers around them as a way to extract information. Some of the other consumers may hold private information about the new good or, in the case of a dynamic setting, they may simply have purchased the good already and then inform the people to whom they are connected. Second, consumers may in fact assign a relatively low weight to the actual characteristics of the good itself and instead attach a higher value to the number of people purchasing the good. A good example is so-called ‘fashion goods’. Firms operating in such a market need to take the properties of the social network between consumers into account when they make their marketing decisions.

The rest of this paper discusses the details of our model and a discussion of the results. Section 2 provides an overview of the relevant literature. Section 3 describes the model used for our agent-based simulations. The experimental setup is described in Section 4. Results are given in Section 5, and Section 6 presents conclusions.

2. Models of Consumer Behavior in Social Networks

The literature on social interactions has studied thoroughly how a preference for conformity can explain herd behavior in consumers or the emergence of fashion, fads, and customs (see Banerjee, 1992; Brock and Durlauf, 1997; Bikhchandani, Hirshleifer and Welch, 1992; Bernheim, 1994). The diffusion of a new product, or innovation in a network, often follows a gradual pattern. In the first stage a few consumers (the innovators or early adopters) adopt, then consumers in contact with them adopt, then consumers in contact with those consumers adopt, and so forth until the innovation possibly spreads throughout the network reaching also the more conservative consumers (or ‘followers’). When the diffusion reaches the majority of the network, we call this a cascade. Such a cascade is associated with the commercial success of the new product. If innovation does not succeed in completely taking off, the firm may decide to file the product as a failure.

Thus, we explore whether or not the new product diffuses to the largest part of the network together with the time actually needed for the critical diffusion.

The initial positive feedback mechanism described above may be offset by a tendency of some consumers to distinguish themselves from the dominant trend. A form of negative externality may then make a few individuals revise their purchasing decision with the effect of limiting the diffusion of the innovation to the entire system. In this paper we will study both positive and negative feedback. We explicitly take into account that the consumption behavior of other people can have a positive externality ('people like to imitate other people'), but also negative feedback ('people like to be special'). One way for these two opposing forces to co-exist is to have 'imitation effects' being replaced by a tendency towards heterogeneity as soon as some critical level of diffusion of the product is attained.

A further useful distinction in the study of consumption behavior concerns the intrinsic purchasing attitude of consumers. Some may be considered 'innovative' consumers. They are the ones who first choose a new product and are responsible for its initial diffusion. Innovators usually represent a small portion of the set of consumers. Most people are instead simply 'followers', in that their strategy is to choose the novel good only after someone else has already tried it. Their strategy is a more conservative one and they are usually responsible for the actual spread of the innovation. These two consumption attitudes have both been shown to play a role in the diffusion processes of many products. One example is new software products (von Hippel, 1988), which are usually tried by a restricted group of experimental users and later, eventually, chosen by a higher number of more conservative consumers. In our model we allow for different consumer purchasing attitudes.

Next, we relate our model to the recent literature on diffusion in networks. Models of informational cascades analyze how the decisions of first movers in a sequential decision problem can lead to herd behavior (a cascade) of the whole system of agents, see Banerjee (1992). Herd behavior occurs when agents do not use any private information but instead only value the information provided by the decisions taken by the other agents. It should be noted that, in these models, decisions are made sequentially, so that agents look at the actions taken by the agents who decided before them. In this paper we take a different perspective and analyze the diffusion of information in networks of consumers, as in Watts (2002). A number of economic models exist that investigate how various diffusion processes in social networks (Ellison, 1993; Brock and Durlauf, 1997; Young, 2003). The focus is on local interactions and, in particular, on positive feedbacks. An exception is Cowan, Cowan and Swann (1997) where negative externalities are also considered, but in their model the influence of the topology of the social network is not investigated. Famous studies have been concerned with studying the choice between competing technologies, the specific properties of 'network technologies' entailing compatibility issues and the local self-reinforcement processes that allow rapid dominance of some new standard or product or institution (Arthur, 1994; David, 1985; Katz and Shapiro, 1985).

Most of the cited theoretical contributions consider regular lattices, defined as symmetric structures where all nodes have the same 'degree', that is, the same number of links departing from each node. Instead, empirical evidence suggests that the degree distribution in social networks is highly right skewed. Specifically, social networks often display the properties of small-world graphs. As originally defined in Watts and Strogatz (1998), these networks are obtained from regular lattices by rewiring randomly chosen edges. In another version (Newman and Watts, 1999), they are graphs whose vertices are connected together in a regular lattice with the addition of a small number of connections bridging randomly chosen vertices. Small-world graphs show a higher level of clustering than random graphs. Their pattern is not as ordered as in a regular lattice, but they preserve short average path lengths proportional to the logarithm of system size. Small-world networks are thought to be a good model for many types of real social networks.

Many analytical results for diffusion phenomena on networks are available from the epidemiological literature that studies the influence of the topology of the underlying network of individuals on the dynamics of disease propagation, see for example Strogatz (2001). These models actually correspond to percolation problems on graphs. Specifically, they investigate threshold values for actual epidemic outbursts as opposed to limited localized spreading. Some models (Callaway et al., 2000) also explore the effectiveness of vaccination strategies that try to inactivate some of the nodes in the network. One limitation to the application of these exact solutions is that they are valid for very large networks, but they do not hold for small networks. They also require rather strict assumptions, see Moore and Newman (2000) for a definition of the general site and bond percolation problem. There are some analytical results on models of diffusion on networks, see for example Newman (2000). However, as the agents become more complex and we introduce a firm with a slightly more sophisticated advertising strategy, results soon become intractable. Therefore, we use an agent-based model to simulate the information and product diffusion process.

The cited literature also analyzes how an external agent can limit or enhance the diffusion process on a network. In the case of a network of consumers, one can envisage different advertising strategies for companies for diffusing a new product in a network of potential buyers. Advertising campaigns are costly and different strategies can be employed by the firms. Firms are boundedly rational and are not fully aware of the structure of the communication channels among consumers. We consider the situation where firms only have aggregate sales data. This can apply to products that are sold over the Internet (for example ring-tones), where the firms do not have information about the characteristics of the consumers. Figure 1 illustrates how the topology of the network may influence marketing strategies. A so-called 'Star Network' (see Figure 1) is easiest to penetrate for the firm. If the firm targets the center of the star, word of mouth about the product reaches all agents in the network very fast. The most difficult network for the firm is the regular network, where path lengths are very long (Figure 1). However, if consumers differ in their tendency to adopt new products, there is a trade-off between targeting consumers that are well connected (such as the center of a star) versus targeting consumers that have certain

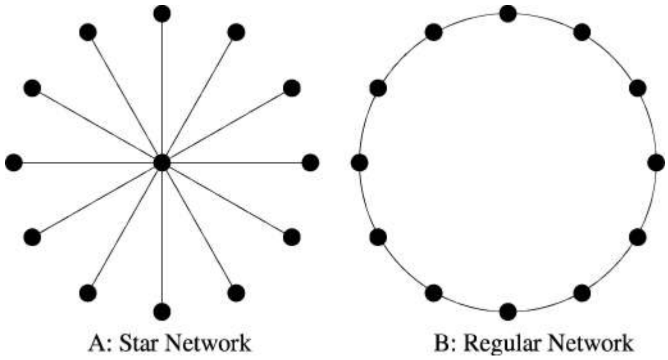

Figure 1. A Star Network (left) and a Regular Network (right).

characteristics, such as the 'opinion leaders' in Valente and Davis (1999). Our model allows us to investigate this trade-off and analyze good advertising strategies for the firm under different circumstances relative to both the social network topology and consumer characteristics.

3. The Diffusion Model

We consider the diffusion of an innovation over a network of consumers. The goal of this research is to investigate whether a firm can learn directed-advertising strategies that increase the size and speed of the diffusion. Consumer evaluation for the product depends only on the fraction of its neighbors that have already purchased the innovation. The consumers are connected through a social network and the innovation spreads through this consumer network. Furthermore, the firm that is pushing the product can advertise and thus inform consumers about the existence of the product. However, advertising is costly and firms can only target a limited number of consumers. Therefore the firm has to learn advertising strategies that produce the most effective diffusion of the product. Figure 2 gives an overview of the simulation model; the components are discussed in more detail below.

3.1. CONSUMERS

As already mentioned, we wish to look at products for which the consumer value depends strongly on the number of other consumers that use the product. If consumer value increases when the number of other consumers who have adopted the innovation becomes larger we call this *positive (network) externalities*, see for example Katz and Shapiro (1985), Farrel and Saloner (1986). *Negative (network) externalities* occur if consumer value decreases as the number of adopters becomes larger. In this paper we investigate both positive and negative externalities. For example,

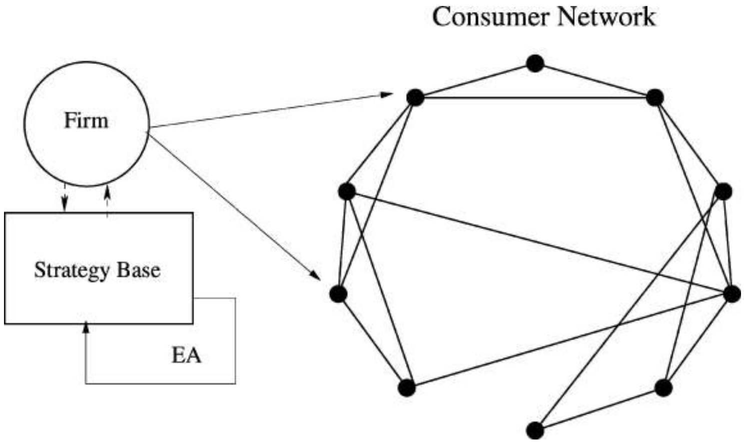

Figure 2. Outline of the simulation model. In this example the consumers are connected through a small-world network structure.

if we consider fashion goods, it might be unrealistic to assume that the positive externality continues to grow until all consumers use the product. Consumers may have a desire to be fashionable on the one hand, while they also want to be special on the other hand. This is particularly true for the innovator who always adopts the latest fashion or fad. These innovators may move on to the next fashion while more conservative consumers are still following a previous one. To be more precise, we also investigate the situation where the innovation is no longer attractive once it becomes adopted by too many people.

A consumer is characterized by her *exposure threshold* (et) and by her neighbors. The threshold and the structure of the social network are exogenously given at the start of the simulation. Consumers that have already adopted the innovation talk about it to their neighbors, so consumers are aware of the purchase decisions of other consumers they are linked to. The *exposure threshold* models the tendency of a consumer to adopt the new product. A consumer with an exposure threshold of 0.5 will buy the product if and only if at least half of its neighbors have also bought the product. Note that this threshold is more easily reached in sparse (but connected) networks than in dense networks. In Section 5.4 we also consider negative externalities, where consumers want to be fashionable but still special. In this case consumers have both an *exposure threshold* (et) and an *over-exposure threshold* (oet) threshold. Consumers stop using the product if the fraction of their neighbors who have adopted the product exceeds their over-exposure threshold. This *over-exposure threshold* is used by the agents to discriminate between innovations that are attractive and innovations that are no longer fashionable because their user-base has become too large. The decision rule for the consumer can be characterized as follows:

Each trade period consumers take the following steps: **1.** *Consumers* who have already adopted the innovation talk about the product to their neighbors **2.** A *Consumer* decides to adopt the product if: Word of mouth received from its neighbors exceeds its Exposure Threshold. and (in case of negative externalities) Word of mouth received from its neighbors does **not** exceed its Over-Exposure Threshold.

3.2. The firm

The firm can engage in an advertising campaign to try to increase the size and the speed of the diffusion of its product (the innovation). Firms do not have any a priori knowledge about the structure of the consumer network and have to learn which advertising strategies are best. They have to choose which consumers to target to ensure fast diffusion of their new product. Advertising campaigns are costly and different strategies can be employed. Firms are boundedly rational and are not fully aware of the structure of the communication channels among consumers. As a result, they are likely to explore a range of targeted-advertising strategies. To model this search and learning process of the firm we use a simple genetic algorithm that is described in more detail below. An advertising strategy specifies which consumers are targeted at time 0. The success of an advertising strategy depends on (1) the number of consumers that have adopted the innovation after a specified period of time when that strategy was used, and (2) the cost of the advertising campaign. The number of products sold is the only information coming from the market that the firm obtains at the end of each period. The algorithm used by the firm can be summarized as follows:

The firm takes the following steps: **1.** Select an advertising strategy **2.** Calculate fitness of the strategy: *Fitness* = *Sales* − *Advertising costs* **3.** Update Strategies: Update strategy base using a GA **4.** Go to **1.**

We investigate the learning capabilities of the firm. We examine whether firms can learn the best directed-advertising strategy when a fixed number of consumers is targeted, that is, the firm has a fixed advertising budget. This allows us to compare the results to a random advertising strategy targeting the same number of consumers.

3.3. THE GENETIC ALGORITHM USED BY THE FIRMS

We use a genetic algorithm (GA) to model the strategy search and learning behavior of the firm. Genetic algorithms were first developed by Holland (1975) and are based on the principle of "survival of the fittest" from nature. A typical genetic algorithm can be described as follows (Mitchell, 1996): first, a population of randomly initialized strategies is generated. This is the population of chromosomes or genotypes. The behavior that is encoded in these chromosomes, that is, the behavior exhibited by an agent using that particular strategy can be seen as the phenotype. This population is improved in a number of generations by means of *selection*, *recombination* (crossover), and *mutation*. Selection chooses the better strategies to serve as parents for the next generation of strategies. This corresponds to the concept of survival of the fittest in nature. Offspring is then formed by pairwise recombination of the parents (crossover). Finally, the offspring strategies are slightly changed, with a small probability (mutation) and the new population replaces the old one.

A strategy is represented by a bit string of length l, whereas in this case, l is the number of consumers in the network. If the ith bit on the chromosome equals 1 this means that the strategy represented by the chromosome targets consumer i (if the bit equals 0 the consumer is not targeted). We use a simple genetic algorithm to update the strategies for the firm. However, using standard crossover and mutation operators, the number of bits set to 1 in each strategy (in our case the number of targeted consumers) may vary from generation to generation. Since we consider firms with a fixed budget, however, we use adapted mutation and crossover operators (see below) to ensure that a strategy targets exactly m consumers, where

$$m = \left\lfloor \frac{\text{Budget}\, b}{\text{Marginal cost of advertising}} \right\rfloor.$$

The fitness of a strategy is determined solely by the sales after t time steps, where t is the *training time*. High fitness strategies have a higher chance of being selected as parents for the next generation. The training time t is taken between 10 and 50 time steps. We have adapted the crossover and the mutation operator in order to ensure that a chromosome contains exactly m ones. Each chromosome thus consists of l bits representing the consumers, m of those bits are set to 1 and are called the *1-bits*. Similarly, *0-bits* are the $l - m$ bits that are set to 0. The algorithms for the adapted operators, which we call the one-preserving mutation and the one-preserving crossover, respectively, are given below.

One-preserving mutation:
For each chromosome that is selected for mutation:
1. Randomly select one of the *1-bits*
(one of the currently targeted consumers)
2. Change this bit to **0**
3. Randomly select one of the old *0-bits*
4. Change this bit to **1**
One-preserving crossover:
For two parent chromosomes (parent1 and parent2) we create
two offspring chromosomes (offspring1 and offspring2):
0. Set all offspring bits to zero.
First we consider the k *1-bits* that the two parents have in common
1. (Both strategies agree on those *1-bits)*
These k *1-bits* are copied onto the offspring chromosomes
(as would be the case with a regular crossover operator)
Second we consider the $m - k$ remaining *1-bits*
2. Select a random number C_{cross} between $[0 \ldots m - k - 1]$
C_{cross} is the crossover counter
3. Copy the first C_{cross} *1-bits* from parent1 to offspring1
Copy the remaining bits from parent1 to offspring2
4. Copy the first C_{cross} *1-bits* from parent2 to offspring2
Copy the remaining bits from parent2 to offspring1

In words, the one-preserving mutation replaces a currently targeted consumer with a consumer that is currently not targeted. Note that the one-preserving mutation thus works on the entire chromosome instead of on a single bit. An example illustrating both operators is given in Figure 3 for a chromosome of length 6. The first parent chromosome thus specifies that consumers 1, 3 and 5 receive targeted advertising (for example in the form of a free sample product).

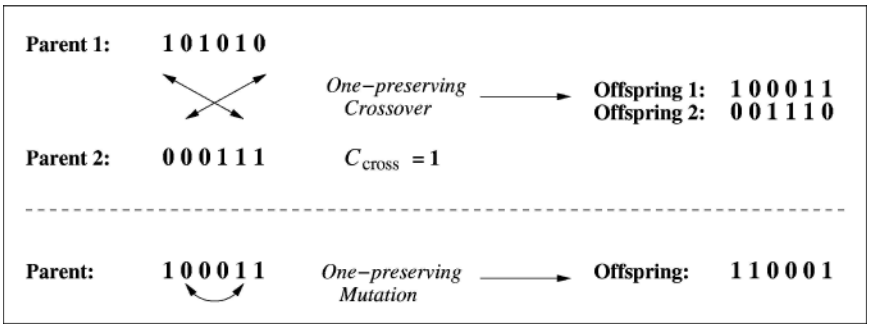

Figure 3. Example of one-preserving crossover and mutation.

3.4. THE SOCIAL NETWORK

We investigate three types of social network architectures: (i) the *k*-regular network, where all consumers have exactly *k* neighbors (here we restrict ourselves to regular networks that are modeled as a one-dimensional ring lattice, where each agent is connected to its *k* nearest neighbors by undirected edges); (ii) the random network, where links between consumers are constructed randomly; and (iii) the small-world network. A small-world network is constructed by taking a regular network and then rewiring a small fraction of the links (usually between 1 and 10 percent of all links Watts and Strogatz, 1998). This fraction of links is called the *rewiring constant* (rc).

The *degree* of a network is defined as the (average) number of neighbors of a given consumer in the network. For *k*-regular networks the degree thus corresponds to *k*. For small-world and random networks, the number of links per consumer may vary and the degree of such networks describes the average number of neighbors of a consumer.

We consider a social network of 1000 consumers. This is sufficiently large to observe the dynamics that occur for different network topologies (Cowan and Jonard, 2004). A *cascade* on the network occurs if, starting from a small fraction of the consumers (the initially targeted consumers), the diffusion spreads out to a large part of the population. In this work, we define a cascade as a diffusion that reaches at least 80% of all consumers, starting from a small initial number of targeted consumers. When the diffusion reaches all consumers in the network we call this a *global cascade*. Note that such a global cascade may not always be possible since random and small-world networks can be disconnected.

3.5. DIFFUSION DYNAMICS

Diffusion dynamics are driven by the topology of the network, the advertising strategy that is used by the firms, and the consumer characteristics. At each time step in the model, the agent actions described above are executed, resulting in the model cycle described below. Initially, the number of consumers as well as their characteristics, the network topology, and the advertising strategy used by the firm are exogenously given. The outcome of the agent-based simulations now depends on those initial conditions and the agent interactions.

Each cycle:

1. *Firms* choose an advertising strategy (from their strategy base)

2. *Consumers* who have already bought the product talk about the product to their neighbors

3. *Consumers* decide whether to (still) adopt the product

Go to 2. (Repeat for a given number of time steps)

4. Firms calculate their profits
5. Firms update their strategies for the next period

We are especially interested in two aspects of the diffusion dynamics: (1) the size of the diffusion, that is how many consumers eventually adopt the innovation, and (2) the speed of the diffusion, that is, how long it takes to reach this diffusion level. An important measure in assessing the properties of the diffusion is the *critical diffusion threshold*. The *critical diffusion threshold* in a network of homogeneous consumers is the highest *exposure threshold* for which a cascade is observed. All thresholds below this critical threshold will lead to cascades on the network.

4. Experimental Setup

The goal of the experiments is to investigate whether firms can learn directed-advertising strategies to increase the diffusion of their products. Furthermore, we want to investigate the properties of such strategies. The social network and the consumer thresholds are exogenously given at the start of the simulation. We vary the topology of the network with respect to degree and network architecture, as well as the exposure thresholds of the consumers, and study the diffusion process over time. We thus perform an agent-based computational study of the diffusion of an innovation over a social network. The diffusion dynamics are a result of the local interactions of autonomous agents over time. The model described above allows us to investigate the speed and the size of the innovation diffusion under different initial conditions. Table I gives an overview of the parameter values that are used in the simulations. We have chosen for a setup where learning is difficult, that is, only a few out of 1000 consumers are targeted by the firm and the firm is allowed only

Table I. Parameter values used in the simulations.

Parameter	Value
Number of consumers	1000
Degree	1–20
Rewiring constant (rc) (small-world network)	0.05
Exposure threshold	0.0–0.5
Number of initially targeted consumers	10
Diffusion time (during learning)	10–50
Generations	20
Number of strategies	50
Pone-mut	0.1
Pone-cross	1.0

very short time to learn. This allows for fast learning and strategies that can be used in an online setting. The rewiring constant is generally set at 0.05 for small-world networks (Watts and Strogatz, 1998). We have also chosen generally recommended values for the GA parameters (Back, Fogel and Michalewicz, 1997).

5. Results and Discussion

5.1. HOMOGENEOUS CONSUMERS

Here we consider the diffusion of information over the network when consumers are homogeneous with respect to their exposure threshold. Figure 4 shows critical diffusion thresholds for the different types of networks when consumers are homogeneous with respect to their exposure thresholds and firms have a fixed budget. Each network consists of 1000 consumers. The degree of the network specifies the (average) number of neighbors each consumer has. Points signify the maximum threshold for which an informational cascade is achieved after 1000 time steps starting from 10 nodes. The averages are over 20 runs; that is 20 different networks

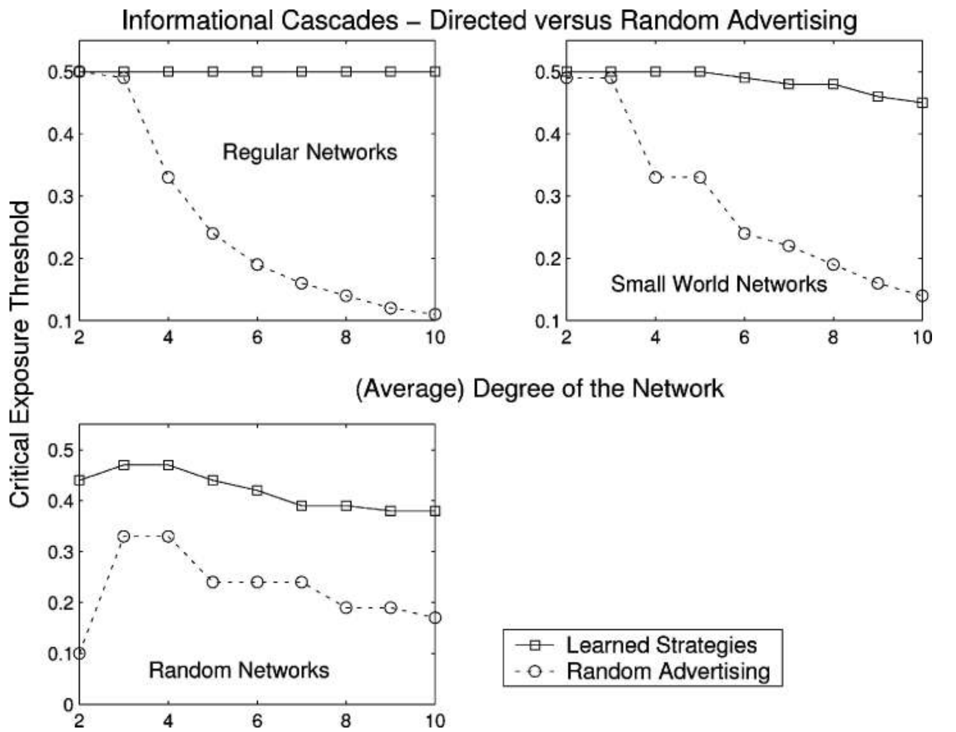

Figure 4. Critical diffusion thresholds for the different types of networks. Points signify the maximum threshold for which an informational cascade was achieved after 1000 time steps starting from 10 nodes. (Averages over 20 runs). The learning time is 50 time steps.

are tested. The consumers that are initially targeted can be different in each run, due to the use of a randomly selected advertising strategy. The learning time is 50 time steps. In the fixed budget experiments we use the one-preserving operators discussed in the previous section. Note that under these conditions the advertising strategy only has to be optimized with respect to the network topology. If we look at Figure 4, we see that firms are able to learn effective advertising strategies. In the case of the regular networks such a strategy has to target consumers that are evenly spread out over the network in order to ensure maximum diffusion. Using such a strategy, firms are able to achieve informational cascades even if consumers have a low tendency to buy the product (a high exposure threshold).

If we restrict our attention to the random strategies, we can observe two different regimes. In a sparse network, diffusion of information is limited by the global degree of the network, but cascades occur even when consumers are quite ‘resistant’ to being convinced by their neighbors, that is when their exposure threshold is high. On the other hand, if the network is sufficiently dense, the propagation is limited by the stability of individual nodes. In this case the critical exposure threshold is significantly smaller. Most nodes have a large number of neighbors, but with a random strategy it is unlikely that all these neighbors are buying the product at time 1, so the initial perturbation may not be able to diffuse at all. Note that critical thresholds are similar for small-world and regular networks. We can observe that there are no cascades on the random network for low degree. This results from the fact that the network may not be connected and the diffusion cannot reach all the components of the network. Notice that, as the degree increases, it becomes more difficult for a cascade to occur. Cascades are only observed in networks where the agents have a low exposure threshold. In a network of degree 2, only one neighbor needs to purchase the product to start the cascade, so a threshold of 0.5 is immediately obtained. Starting from one initial consumer, it becomes much more unlikely that a large fraction of neighbors has bought the good if the degree is high. Furthermore, we also observe cascades for some runs above the critical threshold. Figure 4 shows that in order for a cascade to occur it is necessary that the network is connected, that is, that there are not too many components. Furthermore, we see that as long as the network is connected, a cascade spreads more easily over less regular networks. This is in accordance with the literature, for example, Watts (2002).

With directed advertising, however, this effect disappears and cascades are achieved most often on regular networks. This can be explained by the fact that, on a regular network, firms only have to take into account the position of a consumer in the network. In small-world networks and random networks, however, not only the position of the consumer is important, but also the number of links a consumer has as well as the type of links (cross-network or only to close neighbors). Summarizing, we can say that firms are able to learn effective directed-advertising strategies in a network with homogeneous consumers. Using these strategies, cascades can be achieved in situations where random strategies lead only to limited diffusion. In the next section, we consider networks of heterogeneous consumers.

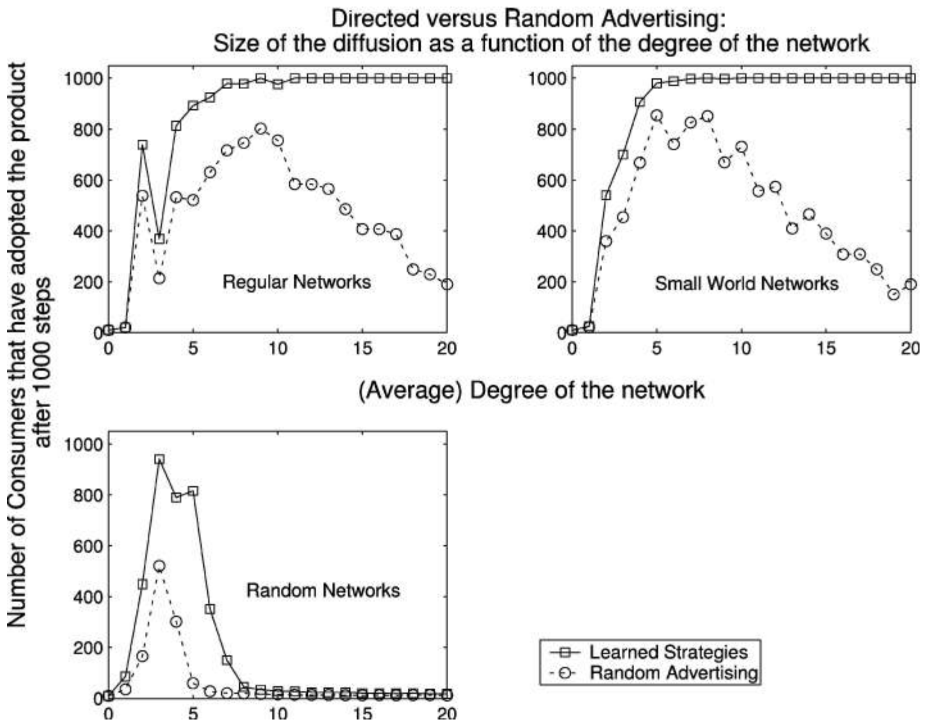

Figure 5. Directed and Random advertising when consumers are heterogeneous: Size of the diffusion as a function of the degree of the network for different types of networks. Each data-point represents the average over 50 runs. The learning time is 50 time steps.

5.2. HETEROGENEOUS CONSUMERS

In this section we consider the diffusion of information over the static network when consumers are heterogeneous with respect to their exposure threshold. Figure 5 shows the difference between random and directed advertising when consumer exposure thresholds are drawn from a normal $N(0.3, 0.1)$ distribution (based on results from the fixed threshold experiments this is the region where cascades are possible but sometimes difficult to achieve, and where firms may thus profit from learning behavior). Each point in the graph represents the average diffusion after 50 runs of 1000 time steps. In each run, a different consumer network is generated.

Note that the learned strategies outperform the random-advertising strategies with respect to the size of the maximum diffusion. Looking at Figure 6, we also notice that (1) directed-advertising strategies are able to achieve cascades when the random strategies are not, (2) the size of the diffusion is larger for directed-advertising strategies (even if no cascade is achieved), and (3) the speed of the diffusion is larger for directed-advertising strategies of networks. Each network consisted of 1000 consumers with a degree of 10. Initially 10 consumers were

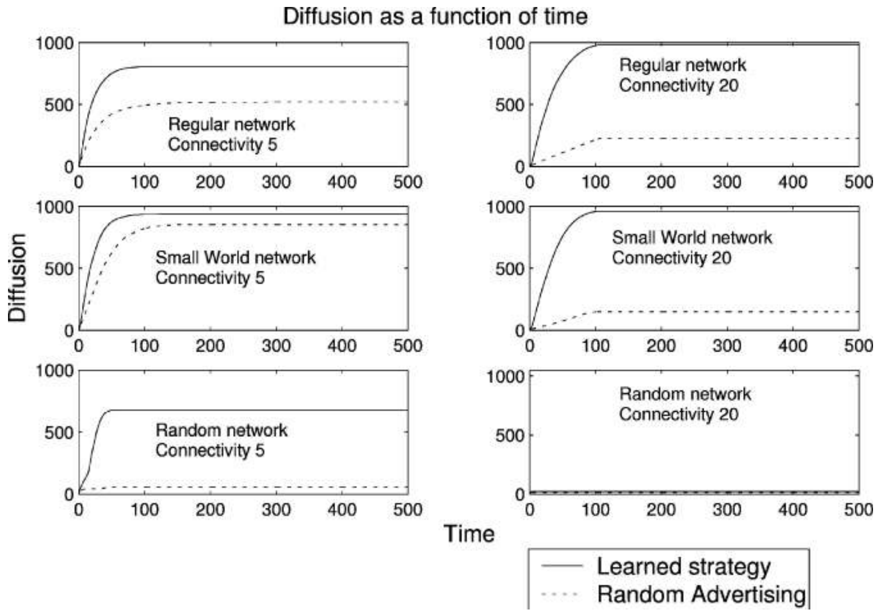

Figure 6. Directed versus Random advertising when consumers are heterogeneous: diffusion as a function of time for different types of networks.

perturbed. In these simulations we consider a homogeneous consumer population with exposure threshold of 0.0 and 0.3 respectively. We can observe that in the no threshold case, the cascade happens fastest for the random network and slowest for the regular network again this is in accordance with the existing theory. If we look at the case of a homogeneous population with an exposure threshold of 0.3, i.e., a consumer decides to adopt the innovation if the fraction of its neighbors that has already adopted the product is larger than 0.3, we see a similar pattern. However, initially the diffusion is slower for the random network, this is caused by the effect that not all consumers have an equal amount of neighbors in the random network. In both cases the small-world network exhibits both the fast start of the diffusion (due to the high regularity of the network) and the fast occurrence of a cascade (due to short average path lengths).

5.3. DYNAMICS

To gain more insight in the nature of the directed-advertising strategies, Figure 7 shows three evolved strategies for a smaller network. We see evolved strategies for three types of networks, with twelve heterogeneous consumers and an average degree of three. The experiments are conducted for a budget of 4. The numbers at the nodes represent the exposure thresholds of the consumers. Black nodes are targeted by the evolved strategies. All three strategies lead to a cascade within 10

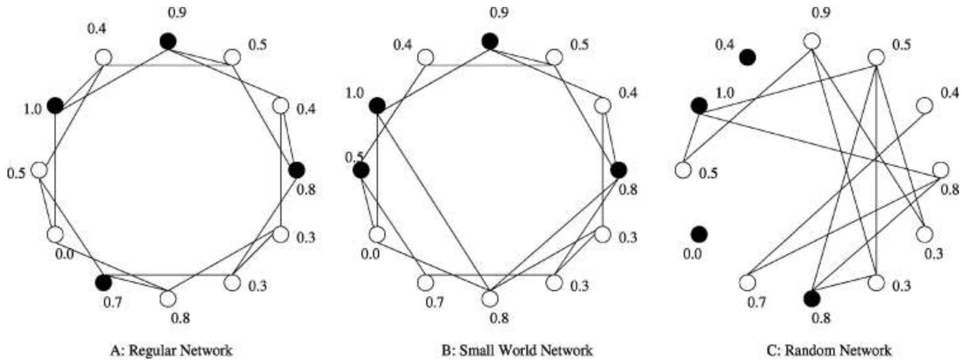

Figure 7. Learned directed-advertising strategies. Black nodes are targeted by the learned strategy. Numbers at the nodes state the exposure thresholds of the individual consumers.

time steps. As we examine the evolved strategies more closely, we notice several interesting features. First, the firms target consumers with a high exposure threshold, that is, the most conservative consumers. Second, they target consumers with a high number of neighbors, and third they target isolated consumers (these consumers can otherwise never be reached by the cascade). Furthermore, we notice that in the case of the small-world network, both nodes that have a "rewired" link are targeted, ensuring a short path length for the diffusion. In the case of random networks, the disconnected nodes (or components) have to be targeted individually.

Most of these effects were also observed in the simulations with the large, 1000 consumer network. More specifically, we found that (1) in case of small-world and random networks, 79% of the targeted nodes had a higher number of neighbors than the average degree of the network. Furthermore, we found that (2) in all three types of networks 92% of the targeted nodes had exposure thresholds greater than 0.5. In contrast with the results for small networks, we found that (3) isolated consumers were not often targeted in the 1000 consumer setting. This can be explained by the fact that if total diffusion is not possible within the limited learning time it is not always profitable to target a single isolated consumer, which only increases the size of the diffusion by only one. In other words, isolated consumers are only relevant when a cascade is possible.

5.4. INTRODUCING NEGATIVE EXTERNALITIES

This section describes the network dynamics when negative as well as positive externalities are present in the model. We consider two scenarios. In the first we assume that the attractiveness of the product depends on the number of other consumers that have adopted it. A consumer decides to adopt the product if this number exceeds its exposure threshold, but discontinues using the product if its over-exposure threshold is exceeded. However, once the number of other users decreases, the user may decide to use the product again. In the second scenario, a consumer never returns to the product once he has abandoned it. This models a situation where the consumers abandon a fashion or fad and move on to the next trendy product.

Scenario 1: Figure 8 shows results for heterogeneous agents when negative externalities are also present. In these experiments, the exposure threshold is drawn from uniform [0, 0.5]. The exposure threshold (et) and the over-exposure threshold (oet) are related, namely oet = et + 0.5. This reflects the fact that the innovators may also be the first consumers who abandon the product if the next innovation reaches the market or if they no longer consider the item fashionable once too many consumers are using it. Again, we have used a network of 1000 consumers and look at the size of the diffusion after 100 periods (the size of the diffusion remained constant if we consider 1000 time steps). First, we note that the learned strategy performs much better than the random strategy in regular networks. This can be explained by looking at the path of the diffusion. In networks with high clustering (such as regular networks and small-world networks with a low rewiring

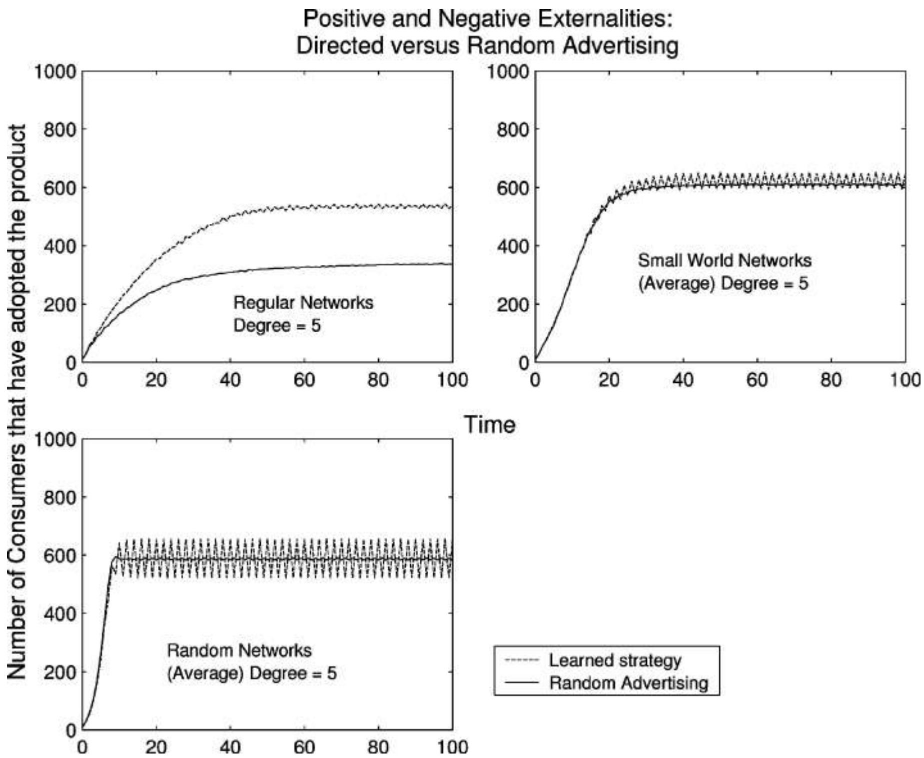

Figure 8. Diffusion of the product when negative as well as positive externalities are present. Averages over 20 runs.

Table II. Average number of consumers that have adopted the innovation, when negative externalities are present, averaged over 20 runs and taken over time step 100–1000.

Network topology	Average size of the diffusion *Learned/random strategy*
Regular networks	535/337
Small-world networks	623/610
Random networks	589/586

constant), the diffusion progresses from neighbor to neighbor. But since, in such networks, most neighbors of one consumer are also neighbors of each other, the *over-exposure* threshold is reached sooner than in networks with low clustering.

This explains why the average size of the diffusion is lower in regular networks than for less clustered networks. Table II gives the average size of the diffusion for the different types of networks. Figure 9 shows a typical run for a small-world network. Note that the average size of the diffusion is a little higher for the learned strategy. The oscillating behavior is caused by a small group of consumers who continually switch between using and not using the product. Note that

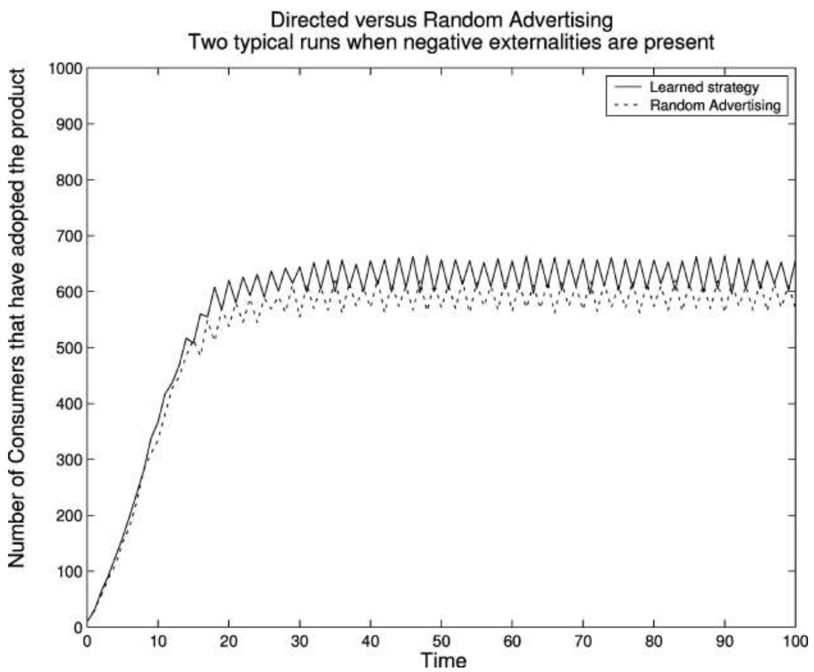

Figure 9. Diffusion of the product in a small-world network when negative as well as positive externalities are present. Two typical runs.

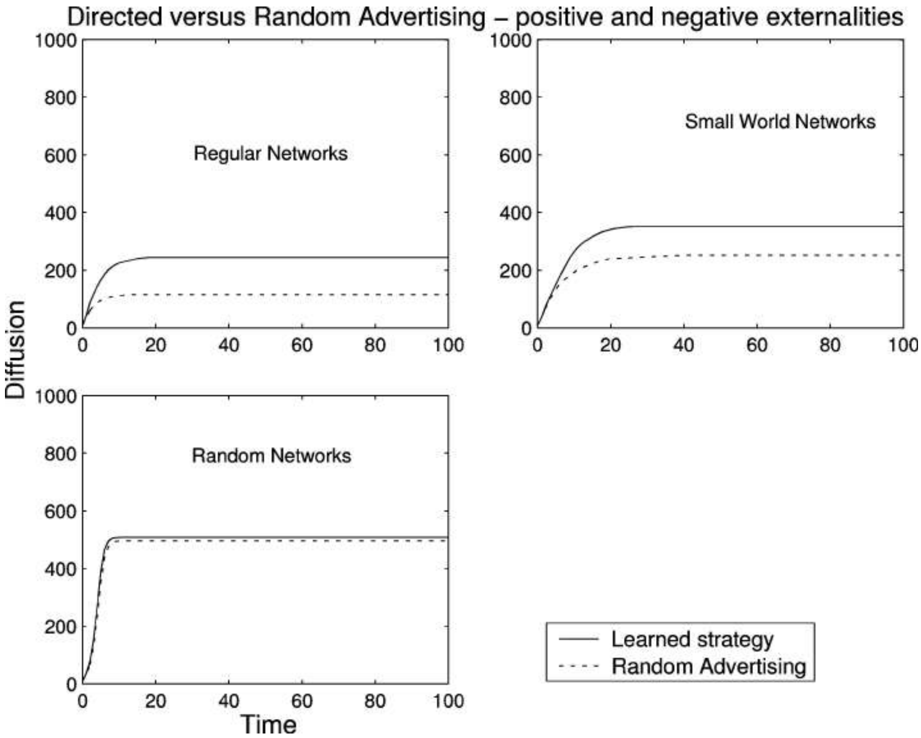

Figure 10. Diffusion patterns when negative externalities are present: scenario 2.

the oscillating behavior also occurs in the random scenario but the fluctuations are much smaller. The group of 'switching' consumers is larger when strategies are learned. If we look at the second scenario, we see that the fluctuations disappear.

Scenario 2: Under this scenario, the negative externality has a permanent effect and a consumer who decides to discontinue using the product never resume using it. Figure 10 gives results for different types of networks under these conditions. First, we notice that the fluctuating behavior disappears. Furthermore, we can observe that the total size of the diffusion is smaller than in the first scenario. This results from the fact that the diffusion dies off because of the negative externality, before all potential consumers have been reached. This effect is stronger for the more regular networks than for the random network. Moreover, the effect of learning is also much clearer for the regular and the small-world network. This reflects the fact that the diffusion goes very fast in random networks.

6. Conclusions

This paper investigates the spread of information in a social network. The network consists of agents that are exposed to the introduction of a new product. Consumers decide whether or not to buy the product based on their own preferences and the decisions of their neighbors in the social network. We use and extend concepts from the literature on epidemics and herd behavior to study this problem. The central question of this paper is whether firms can learn about the network structure and consumer characteristics when only limited information is available, and use this information to evolve a successful directed-advertising strategy. To do this, we have extended existing models to allow for heterogeneous agents and positive as well as negative externalities. The firm can learn a directed-advertising strategy that takes into account both the topology of the social consumer network and the characteristics of the consumer. Such directed-advertising strategies outperform random advertising.

These results are a first step towards enabling online strategic learning with sparse information by the firms. Currently we are extending our model to allow for non-static consumer networks which complicates learning even further.

Acknowledgements

We would like to thank Jason Woodard and other participants from the Santa Fe Complex Systems Summer School 2002 and the 2004 SCE conference on Computational Economics and Finance as well as two anonymous referees for useful comments. Part of this work was performed while the first author was working at the Centre for Computer Science and Mathematics (CWI) in Amsterdam, The Netherlands, supported by NWO-grant 612-60-008.

References

Arthur, W.B. (1994). *Increasing Returns and Path Dependence in the Economy*. University of Michigan Press, Ann Arbor, MI.

Bäck, T., Fogel, D. and Michalewicz, Z. (eds.) (1997). *Handbook of Evolutionary Computation*. Oxford University Press, Oxford.

Banerjee, A.V. (1992). A simple model of herd behaviour. *Quarterly Journal of Economics*, **107**(3), 797–818.

Bernheim, D. (1994). A theory of conformity. *Journal of Political Economy*, **105**, 841–877.

Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, **100**(5), 992–1026.

Brock, W.A. and Durlauf, S.N. (1997). Discrete choice with social interaction. *Review of Economic Studies*, **68**, 235–260.

Callaway, D.S., Newman, M.E.J., Strogatz, S.H. and Watts, D.J. (2000). Network robustness and fragility: Percolation on random graphs. *Physical Review Letters*, **85**, 5468–5471.

Cowan, R. and Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control*, **28**, 1557–1575.

Cowan, R., Cowan, W. and Swann, P. (1997). A model of demand with interactions among consumers. *International Journal of Industrial Organization*, **15**, 711–732.

Cownan, R., Cowan, W. and Jonard, N. (2004). Waves in consumption with interdependence among consumers. *Canadian Journal of Economics*, **37**, 149–177.

David, P.A. (1985). Clio and the economics of qwerty. *The American Economic Review*, **75**, 332–337.

Ellison, G. (1993). Learning, social interaction and coordination. *Econometrica*, **61**, 1047–1071.

Farrell, J. and Saloner, G. (1986). Installed base and compatibility: Innovation, product preannouncements, and predation. *The American Economic Review*, **76**(5), 940–955.

von Hippel, E. (1988). *The Sources of Innovation*. Oxford University Press, New York, NY.

Holland, J.H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI.

Katz, M.L. and Shapiro, C. (1985). Network externalities, competition, and compatibility. *The American Economic Review*, **75**, 424–440.

Mitchell, M. (1996). *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, MA.

Moore, C. and Newman, M.E.J. (2000). Epidemics and percolation in a small-world networks. *Physical Review E*, **61**, 5678–5682.

Newman, M.E.J. (2002). The spread of epidemic disease on networks. *Physical Review E*, **66**.

Newman, M.E.J. and Watts, D.J. (1999). Scaling and percolation in the small-world network model. *Physical Review E*, **60**, 7332–7342.

Scott, J. (2000). *Social Network Analysis*, 2nd edn. Sage Publications, London.

Strogatz, S.H. (2001). Exploring complex networks. *Nature*, **410**, 268–276.

Valente, T.W. and Davis, R.L. (1999). Accelerating the diffusion of innovations using opinion leaders. *Annals of the American Academy of Political and Social Sciences*, **566**, 55–67.

Wasserman, S. and Faust, K. (1994). *Social Network Analysis*. Cambridge University Press, Cambridge, MA.

Watts, D.J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences (PNAS)*, **99**(9), 5766–5771.

Watts, D.J. and Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. *Nature*, **393**, 440–442.

Young, H.P. (2003). The diffusion of innovations in social networks. In L.E. Blume and S.N. Durlauf (eds.), *The Economy as an Evolving Complex System III*. Oxford University Press, Oxford.