

1

Dissemination of health information within social networks

Charanpal Dhanjal¹, Sandrine Blanchemanche², Stephan Clemençon³,
Akos Rona-Tas⁴ and Fabrice Rossi⁵

1.1 Introduction

State agencies responsible for managing various risks in social life issue advisories to the public to prevent and mitigate various hazards. In this Chapter we will investigate, how information about a common food born health hazard, known as *Campylobacter*, spreads once it was delivered to a random sample of individuals in France. The *Campylobacter* is most commonly found in chicken meat and causes diarrhoea, abdominal pain and fever. The illness normally lasts a week but in rare cases patients can develop an auto-immune disorder, called Guillain-Barré syndrome, that leads to paralysis and can be deadly. *Campylobacter* together with *Salmonella* is responsible for more than eighty percent of food born illnesses in France and strikes over 20,000 people each year. People can take simple steps to avoid infection by cleaning their hands, knives, cutting boards and other food items touched by raw chicken meat and by cooking the meat thoroughly.

In this Chapter we build two different network models to see how the information about *Campylobacter* diffuses in society, by mapping onto various network structures the data we gathered with three waves of surveys. In these models the spread of information depends on two sets of factors. First, each person has a set of individual properties that influences their

¹ Institut Telecom – Telecom Paristech, 46 rue Barrault, 75634 Paris cedex 13, France – email: charanpal@gmail.com

² INRA-Met@risk, 16 rue Claude Bernard, 75005 Paris, France – email: sandrine.blanchemanche@paris.inra.fr

³ Institut Telecom – Telecom Paristech, 46 rue Barrault, 75634 Paris cedex 13, France – email: lise.stephan.clemencon@telecom-paristech.fr

⁴ University of California, San Diego, 9500 Gilman Dr., La Jolla, USA – email: aronatas@dssmail.ucsd.edu

⁵ Institut Telecom – Telecom Paristech, 46 rue Barrault, 75634 Paris cedex 13, France – email: fabrice.rossi@telecom-paristech.fr

propensity to transmit the information to or to receive the information from someone they know. Second, each person is connected to others in ways that also affects transmission. There are three aspects of these social ties that matter. As the information travels through existing ties, the quantity of connections should have an influence as people with more ties should have more opportunity to disseminate the information. The quality of ties should matter as well, because certain types of ties may be more conducive to information transmission than others. Finally, the overall structure of the entire network, *i.e.* how ego-centric networks are linked into a larger whole, should also play a role. Our surveys provide data for the individual characteristics, as well as the quality and quantity of the social ties. In the diffusion model, two different overall network structures, the Erdős and Rényi random, and the Small World (SW) model, are introduced through modeling assumptions.

The central question of this Chapter is how individual characteristics and the various aspects of social network influence the spread of information. A key claim of our Chapter is that information diffusion processes occur in a patterned network of social ties of heterogeneous actors. Our percolation models show that the characteristics of the recipients of the information matter as much if not more than the characteristics of the sender of the information in deciding whether the information will be transmitted through a particular tie. We also found that at least for this particular advisory, it is not the perceived need of the recipients for the information that matters but their general interest in the topic.

As for the diffusion of information, we found that the two network structures behave differently in some ways. If the proportion of the population that receives the information initially from the center (our survey) is lower, the random graph model diffuses the information to a larger segment of the population than the SW model, as the advisory travels farther in random networks. However, as the initial exposure increases above a certain level, the two models deliver the advisory by word-of-mouth to very similar proportions of the population. For the SW models we find that as the size of the group that is initially exposed grows, diffusion first increases then decreases and there is an optimal proportion of the population that the message initially has to reach to result in the maximum word-of-mouth diffusion. Since the initial deployment of the information costs money and the subsequent diffusion is costless for the center, finding this optimal size is of practical importance.

Finally, we offer a visual presentation of the transmission process and the way the composition of those receiving the message shifts. Our analysis

finds that distribution of particular characteristics of message recipients changes most in the first round and in later rounds it gravitates towards the distribution of the characteristics in the population.

1.2 Theoretical overview

1.2.1 Diffusion of health information and interpersonal communication

Major public health campaigns are aimed at modifying individual conduct, by either trying to decrease risky behavior, such as alcohol abuse and smoking or seeking to increase health promoting activities such as exercise and following a healthy diet (Compas et al., 1998). The approaches to the dissemination of advisories in the field of public health have varied across time, countries and objectives of the campaigns but they are systematically based on specific views of society and on certain assumptions about social processes.

Broadcast approach

The dominant approach in disseminating health advisories has been the broadcast method where the information is released centrally and it is targeted at members of the public individually. The main assumption behind the broadcast approach is that the message travels directly from the center, usually the health agency, to each citizen via various channels of mass communication. The success of the broadcast approach depends on reaching as many people as possible, and on delivering the message in ways that the desired effects are created in each member of the audience. The broadcast model is a hub-and-spokes system with the agency in the middle and each unconnected member of the public at the end of one spoke. Communicators using this model, therefore, concentrate on two aspects of the process: reach and stickiness. Reach is the matter of getting the message to the largest possible part of the target audience and involves careful planning of the dissemination of the information. Communicators must decide what media are best suited to their purposes, where they should deploy the information, when and how many times. Stickiness is a matter of creating the proper effect once the information is delivered. Here the main objective is getting people to pay attention to, understand and act on the information in question. To improve individual reception of their messages, broadcasters have spent a lot of effort trying to understand the psychology of message reception and perception in their quest to craft the most effective messages.

1.2.2 The two-step model

Mass communication

The undisputable strength of the broadcast model is that it provides the maximum control for the center over the message as each recipient obtains the information from the central source. Doubts about the broadcast model, nevertheless, have been raised early on. One of its first critics was the sociologist, Paul F. Lazarsfeld who in a series of studies on political communication in the 1940s found that the direct effect of broadcasted political messages is negligible and that most people acquire their political opinions not from television, radio or newspapers, but from other people he named ‘opinion leaders.’ He argued that political messages produce their effects – if at all – in a two step process: the original message is picked up by the opinion leaders who then pass it on to the rest of the community (Lazarsfeld et al., 1948; Katz and Lazarsfeld, 1955). Lazarsfeld’s main insight was that most people receive political messages not from the media, from a central point of emission, but by word-of-mouth from others in their community and, therefore understanding the structure of the community is crucial for making communication effective (Katz, 1996; Burt, 1999).

Marketing

A similar idea surfaced in market research where advertisers encountered the same difficulty in getting their messages across. Researchers found that just as citizens in political discourse, consumers often obtain their information through social ties (Katona and Mueller, 1955) and not from the advertisements they are bombarded with. Following this observation, the field of market research has distinguished three types of customers who influence others: the early adopter, the opinion leader and the market maven. Early adopters of new products influence others by buying the product. Their purchases inform other people that the product is available and worth acquiring. Opinion leaders have special expertise about a particular piece of merchandise and dispense it to those interested in the product. Market mavens, on the other hand, are individuals who research and plan their purchases and pay a lot of attention to getting the best deal, and as a result, “have information about many kinds of products, places to shop, and other facets of markets, and initiate discussions with consumers and respond to requests from consumers for market information” (Feick and Price, 1987; Clark and Goldsmith, 2005). Marketing people are eager to find early adopters, opinion leaders and market mavens. They focus on the special characteristics of these senders of information. If they can sell them

their product (or in the latter two cases sometimes only the idea of the product) they can count on an interpersonal multiplier effect.

Public health

The limitations of the broadcast approach and the research on opinion leaders turned the attention of public health officials to new models of disseminating health information by trying to exploit interpersonal influence. During the 1980s, health promotion programs were increasingly built on community-based intervention that tried to differentiate populations with respect to their health related behavior (Shea and Basch, 1990). This approach drew upon the theoretical perspective of diffusion (Rogers, 2003) and social learning theory (Bandura, 1977) arguing that people acquire new behavior from people in their environment through observational learning. One well-known example of this approach is the North Karelia Project on smoking. Launched in 1972 in Finland to combat the country's record high mortality of cardiovascular diseases (Puska and Uutela, 2000), this project was one of the first major community-based projects for cardiovascular diseases prevention. It was built in partnership with WHO and targeted communities with health information through various channels (television, newspapers, personal health training, seminars, etc.) in an attempt to reduce the number of smokers. Based on the two-step theory of the diffusion of innovations, a network of local opinion leaders was identified in each community often through relevant local organizations. Opinion leaders then were trained to spread the advisory in a credible and effective manner. Unlike broadcasting, community intervention (as its name suggest) aims at a much smaller audience and assumes a simple, two-step connectedness among audience members.

1.2.3 Multi-step model

The concepts of the opinion leader, early adopter and market maven and the strategic actor of the community intervention approach drew attention to the fact that information cannot be thought of as a one step process and that the public is not atomized but linked through social ties. Yet all of them, with the exception of the early adopter, lead only to a two step model. Society or the community is neatly divided between leaders and followers, everyone is either the former or the latter but not both. Studies of early adopters, on the other hand, opened up the possibility of a multi-step diffusion model. As a few, very adventurous early adopters are followed by less adventurous early adopters, and then by mainstream adopters, and fi-

nally laggards, the process leads to a chain of diffusion where people in the middle are both leaders and followers. Empirical studies of diffusion began in the late 1930s and were mostly concerned with the spread of innovation such as the adoption of ham radios (Bowers, 1937), progressivist policies (McVoy, 1940), hybrid corn (Ryan and Gross, 1943), and the prescription of new drugs by doctors (Coleman et al., 1957) through imitation. Most studies were interested in measuring the accumulation of adoptions over time.

In a broadcast model, the curve plotting the cumulative number or proportion of adopters as the broadcast is repeated again and again is close to a logarithmic function truncated at or before the point where no person is left in the population to adopt. The marginal return to broadcasting is highest at the earliest part of the process. The first broadcasts have the highest effect and subsequent ones produce a declining yield. Once we move away from the broadcast model and allow for imitation, the typical diffusion model becomes an ogive in the shape of an S, as the diffusion process takes off slowly then accelerates when adopters achieve a critical mass or tipping point (Schelling, 1978; Galdwell, 2000). The curve would slow down as the pool of potential adopters begins to shrink completing the figure. There are several functions that can describe the empirical curve and much of diffusion research consist of correlating curves with the diffusion mechanisms that are thought to have produced them (Mahajan and Peterson, 1985). Until recently, most diffusion models have been macro models predicting only the total number of adopters. As they rarely observe the transmissions, diffusion researchers mostly deduce how they happen from the aggregate outcome by assuming simple transmission rules. The simplest models assume a homogenous population where individuals have different propensities to adopt. Models of the adoption chain start from the assumption that there are personal characteristics that make some people early adopters. Personal attributes determine at which stage of the diffusion process people enter, *i.e.* how many adopters they must perceive to make the move. In information diffusion processes, the most curious and cognitively astute people would be the first to find out about an advisory, others would not pick up the information unless they heard it from a friend, yet others would need to hear it from many to understand and believe it, etc. Here the spread of information would depend simply on how many people you have of each cognitive type. If you have too few people at the beginning of the chain, they will transmit the information to too few people in the second group, not enough to make anyone care in the third group and the diffusion fizzles out early. If there are more people in

the first group, but very few in the second group, the outcome could be similar, etc. The thing to notice is that in these models, propensities inherent to each individual that matter (Granovetter, 1978; Kuran, 1987) after that only the aggregate number – or proportion – of people who already adopted makes a difference. A different approach assumes that all people have the same propensity but they are located differently in a social structure because they are connected differently to others. Whether to adopt or transmit then depends on their relative position – their spatial (Ryan and Gross, 1943; Schelling, 1978) or social proximity (Coleman et al., 1957) to those who already adopted.

In recent years, researchers have moved further away from the broadcast model that assumes that the world can be sharply divided into a source and atomized targets, beyond the two-step model of center, leaders and followers, and past the simplistic, multistep diffusion models and entered a world where all actors are 1) both potential sources and targets, 2) linked in a patterned network of social ties that makes them more or less likely to transmit information and that influences the overall travel of the information in question, and 3) heterogeneous in terms of the properties that makes them more or less likely to send or receive information. In this Chapter, we will use percolation models that combine graph theory that captures the structural characteristics of networks with dynamic processes.

1.2.4 Structural characteristics of graphs

In order to describe a graph, we distinguish vertices (in our case people) and edges (their ties). To characterize a graph, it is useful to compute certain properties and we present several useful ones for social networks (for a survey, see Goldenberg et al., 2001). Perhaps the most well known property of a social network is that they are often ‘small worlds,’ which means that the shortest path length between two people is often small (Milgram, 1967). A formalisation of this measure is the mean geodesic distance for a connected graph, which is the average of the shortest paths between all possible pairs of vertices. Another useful measure for social networks is the clustering coefficient of a graph, which is the chance that a friend of your friend is also your friend.

1.2.5 Random graphs

Random graphs are formed by generating a set of edges for a graph in a random fashion. Many random graph models have been extensively stud-

ied both theoretically, and in relation to real networks, and later we apply them in our simulated graph processes.

A classic random graph model is that of Erdős and Rényi (Erdős and Rényi, 1959, 1960), which was independently discovered by Solomonoff and Rapoport (Solomonoff and Rapoport, 1951). In the model, for a graph of size n , each pair of vertices has an edge between them with probability p_e . One of the properties of the Erdős-Rényi random graph in the context of social networks is that the clustering coefficient is often quite small. The SW model tends to have a high clustering coefficient and small geodesic distance. It can be constructed by taking a one-dimensional lattice of n vertices in a ring, and joining each vertex to its neighbors k spaces away on the lattice. There are therefore kn edges in this lattice. The edges are re-wired by going through each one in turn and with probability p_s moving one end of the edge to a new vertex chosen uniformly at random. No self edges (edges from a vertex to itself) or double edges (pairs of edges between the same pair of vertices) are created.

An interesting question about random graphs is how they compare to real-world ones. Dekker (2007) finds that real social networks have a low average network distance, a moderate clustering coefficient and an approximate power-law distribution of node degrees. In Newman et al. (2002) the authors observe that real social networks have highly skewed degree distributions which can vary according to the property being measured. A random graph model is proposed which can be fitted to an arbitrary degree distribution. Such a model is useful since the theoretical computations of clustering and path length over the model often, but not always, bear strong similarities to those found on real data.

1.2.6 Modeling dynamic processes

Disease spread: epidemiological models

In epidemiology, there are two deterministic models used most often to study the spread of infectious disease from person to person. These models are based on a simple mathematical formulation that does not take into account network properties (Hethcote, 2000). The SIR model (people are either Susceptible, Infected or Recovered) is an appropriate approximation for diseases that infect a significant part of the population in a short outbreak (such as influenza). This model considers people who recovered from the disease to have acquired permanent immunity.

In the SIS (Susceptible/Infected/Susceptible) model, people do not acquire permanent immunity and return to the state of susceptibility when

they recover from a disease (e.g., tuberculosis). The SIS model is appropriate for endemic diseases which persist in a population for long years.

Even though the models now being applied to specific diseases are more complex and refined (Hethcote, 1989) – for instance some models use periodic contact rates to take into account the prevalence of many diseases which varies because of seasonal changes in daily contact rates – they are quite limited. One of the limits of these basic models is that they make the unrealistic assumption that the population is homogeneously or “fully mixed.” The homogeneity assumption, that states that everyone is equally susceptible before and infectious after acquiring the sickness, can be relaxed somewhat assuming that the population belongs to a small set of categories (men, women, or adults/children) with heterogeneous characteristics.

These models also assume that the population is equally and randomly connected (Watts, 2003; Brauer, 2005). Different network structures are addressed by positing multiple levels of mixing, e.g., people may belong to two levels: to a household and to the world, and connected more at one level (with household members) than at another (everyone else, see Ball et al., 1997), yet those solutions still miss many dimensions and configurations of social ties.

Information spread

Most empirical studies of information spread investigate diffusion of information through the Internet. For instance, information propagation in Weblogs or “blogs” is analyzed using a corpus containing 401,012 posts in (Gruhl et al., 2004). One of the dimensions of analysis is the topics of posts. They characterise topics as: “just spike” which are inactive, then active, and then inactive again; “spikey chatter” which have a significant chatter level and are sensitive to real world event and hence have spikes; and “mostly chatter” which have moderate levels of discussion.

The authors also model topic propagation among individuals. An information propagation model is derived based on the Independent Cascade model (Goldenberg et al., 2001). In this extended model, each vertex is a person and each directed edge has a probability of information being copied from one vertex to another in the next time quantum. The model is extended with an additional edge parameter which is the probability that a person reads another person's blog. Edge probabilities for the transmission graph are learnt using an EM-like algorithm (Dempster et al., 1977). One of the observations made from the learnt transmission graph is that most people transmit on average to less than one additional person whereas some

users transmit to many others, providing a boost to certain topics. An important difference between this work and ours is that individuals can reach many others through their blog posts, and in our case, the number of possible transmissions is limited by the number of regular social contacts.

Kempe et al. (2003) consider the problem of selecting the most influential nodes in a social network in the context of information propagation. One application of this work is in the analysis of the “word-of-mouth” effect in the promotion of products. The challenge is to discover which individuals should be targeted with information in order to trigger a cascade of further adoption. This problem is NP-Hard, however efficient greedy algorithms are shown to find a solution within sixty-three percent of the optimal for the Independent Cascade and Linear Threshold models. Our diffusion model is essentially a deterministic variant of the information cascade model, and hence this result can also be applied to our work.

Percolation process

A percolation process is one in which vertices (sites) or edges (bonds) on a graph are randomly designated either “occupied” or “unoccupied” and one asks about various properties of the resulting patterns of vertices (Newman, 2003). Percolation theory is mainly developed in physics but, as Newman (2003) reminds us, one of the initial motivation of its development in the 1950s was the modeling of the spread of disease and it is still used in epidemiological studies (Sander et al., 2002). Newman and Watts (1999) used site percolation on SW graphs as a model of the spread of information or a disease in social networks, and Allard et al. (2009) deployed a bond percolation model taking into account heterogeneity in the edge occupation probability through a multitype networks approach (see also Cohen et al., 2000a). Callaway et al. apply a more general approach (Callaway et al., 2000) that consider the probability of the occupation of a vertex given its degree k .

Our Chapter proposes a model with varied susceptibility to infection and shows under strong heterogeneity in susceptibility there are patches of uninfected but susceptible people. The model uses both bond and site percolation in which vertices either have a specific piece of information or do not. The question in this Chapter relates to how individuals select among their network of people those interested in the information, and how this information is then diffused to the broader network.

1.3 Data and method

There are two approaches to network based research. The first looks at ego networks and takes a sample of individuals and through a series of questions tries to map out their social ties (Marsden, 2005). These are then analyzed together with other individual attributes in statistical models assuming independent and random observations. The advantage of ego network research is that it allows for large and representative samples and consequently provides generalizable findings. It also permits the collection of a large amount of information about individuals (egos). This tradition generates data that are strong on vertices, weak and incomplete on edges. The second approach takes entire populations and maps the relationships among all vertices. The relationships or edges are usually complete and they are observed as opposed to reported, but the data about the vertices tend to be limited. The choice of populations – and therefore the topics – is opportunistic.

Our approach belongs to the first line of investigation but with a unique design. We began with an ego network method then we followed the network path through which the information was dispersed collecting data on dyads in some cases from both the sender and the recipient of the information at the two ends of the tie.

1.3.1 Data

To recruit a sample representative for gender, age and socio-economic status of the French population, we first broadly sent an invitation to 24,000 people to answer a questionnaire on “Food Habits and Food Risks”. They answered socio-demographic questions which allowed us to select those who fit the quotas. Although the sample is broadly representative for the French population, people in the sample have more frequent connections to the Internet than the population as a whole. In the sample, all people use an Internet connection whereas 39% of the French never use one (Cardona and Lacroix, 2008) and 84.7% of people in the sample use an internet connection at least once a day while this number is 41% nationwide.

The survey was conducted in three waves. The first wave, that took place in December, 2008 and January, 2009 interviewed 6,346 individuals, we call Egos. Egos are those who receive the information from us. In our models, Egos can only be senders of the information as they are at the root of the diffusion process. All interviews were conducted through the web using self-administered web surveys. In the first wave, we asked Egos a series

of questions about their knowledge of food risks, their food habits, social networks and socio-demographic characteristics. We also exposed them to the information about *Campylobacter*, which explained the health hazard and provided advice on how to avoid it. We did not tell them that we are interested in information diffusion, but we informed them that we would revisit them for a follow up.

We returned to the sample of Egos three weeks later. In the second wave, we were able to interview 4,496 of the first wave sample, an attrition of twenty-nine percent. The analysis shows that those who dropped out tended to be somewhat less educated. This time we asked people what they remembered of the information about *Campylobacter*, how they changed their behavior, and if they told about *Campylobacter* to anyone else. We asked a series of questions about each person they reported to have talked to about *Campylobacter* and we requested that by writing in their contact's e-mail address send an email requesting that they fill out a questionnaire for the study. Egos contacted and described 7,655 contacts we call Alters. On the average, Egos transmitted the information to 1.7 Alters (Figure 1.1). The most common way to convey the information was face-to-face followed by phone conversations. Only less than 6% of Egos told others about *Campylobacter* via the internet (Table 1.1). None of the socio-demographic variables were strongly correlated to transmission but young people, people working in small companies and those who live with a partner and have children, and those with more than elementary education were a little more likely to transmit the information.

In the third wave, we interviewed the Alters willing to respond to Ego's request. The questionnaire for these Alters was similar to the first and second questionnaires administered to Egos adjusting for differences in context.⁶ We obtained 451 responses. Only 301 had any recollection of the encounter with Ego. We used only these Alters in our analysis. The responding Alters reported to have contacted yet another 138 people with the information.

There is no generic model for the spread of information independent of the nature of the information to be diffused. In our case, the advisory about the *Campylobacter* has certain peculiarities. Advisory about this sickness is not knowledge typically sought by its recipients, such as information about chronic diseases or jobs, but rather, it is "pushed" by its sender who wants to benefit others (or simply wants to pass the time with small talk). The

⁶ For instance, we could not ask any questions from Alters that would gauge their knowledge about *Campylobacter* before they received the information from Ego. We did ask such questions from the Egos in the first wave before we presented the advisory.

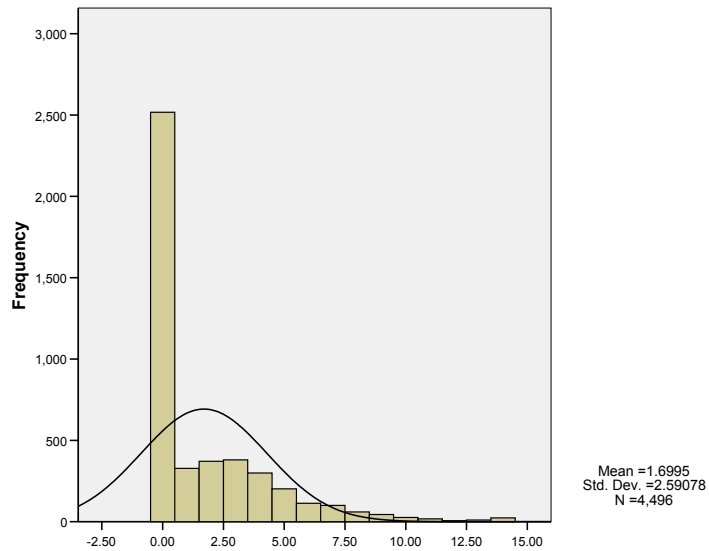


Figure 1.1 The distribution of the number of transmissions reported by Egos.

Type of contact	Percent
Face-to-face	78.95
Phone	15.37
Internet (e-mail, Skype, etc.)	5.68
Total	100 (N=7,655)

Table 1.1 *The type of contact through which the information was transmitted.*

motivation for transmitting the information is rarely self-interest because the sender typically does not directly benefit from sending the information unless he transmits it to a household member who cooks for him, in which case there is a motivation of self-protection against the disease.

Furthermore, the transmission is dyadic and does not involve a critical mass. In this respect, it is similar to epidemics where a single contact with an infectious person makes one as sick as multiple contacts with many ill people. Certain types of information do spread better when the sender received it from multiple sources. This is the case when the credibility of the information depends primarily on how widely it is held. Belief in the upward or downward trajectory of the stock market belongs in that category. In our model, the *Campylobacter* advisory is transmitted dyadically partly because of its content is found to be credible by most people (the average

score for both Egos and Alters was 6.1 on a seven-point scale) and there is no need for outside reinforcement.

Moreover, health advisories, like most diseases, are most infectious soon after they are received and as time passes they become less likely to be transmitted. In our data, 91% of Egos who could recall the timing transmitted the information during the first week and almost half of those on the day they received the information from our survey. Only less than a tenth of the transmissions happened in the second and third weeks. During the time window of our study we observed most of the transmissions that was going to happen and we are unlikely to have missed an outburst of diffusion after we completed our study.

1.4 Computational simulation

There are several limitations of collecting real data, influenced by factors such as cost and privacy. To alleviate these limitations we extract patterns from the survey data and then extrapolate them by simulating a social network. The main aim of this simulation is to measure the average information spread from each Ego on different types of random network. Average information spread is computed using a measure called *average hop length* which is defined as the the total number of transmissions of information divided by the total number of original senders.

All computational code is implemented using Python and the NumPy, SciPy, Matplotlib, and Mlpy (Albanese et al., 2009) libraries. The Support Vector Machine (SVM) code is provided using LIBSVM (Chang and Lin, 2001)⁷.

1.4.1 Data preparation

In the simulation, vertices represent people and edges are relationships between them. For example, an undirected edge exists between vertices if they know each other as friends, family members, colleagues or acquaintances. Vertices are $\mathbf{v} \in \mathbb{R}^d$, $d = 62$, and store values such as the age, gender, education level of each person. Appendix A shows the complete list of fields used in the simulation. Note that missing values are often replaced with the mode. Furthermore Q7, Q43M1-10, Q47CM1-5 and Q47EM1-5

⁷ The complete source code for the simulation experiments is available online at <http://sourceforge.net/projects/apythongraphlib/>.

are categorical variables and hence represented using binary indicator features. As an example, Q7 represents the respondents profession and has 8 categories, hence is represented by 8 binary variables indicating the presence or absence of each category.

The full details of the data preparation stage are given in Appendix B. Essentially, missing characteristics in the data are completed based on the knowledge aquired about the survey population. The total size of the dataset after the completion step is 86,755 pairs of people (examples), in witch 82,485 were negative (no transmission) and 4,270 were positive (transmission) examples. We denote by S the set of triples composed of pairs of people and an indicator label for transmission occurrence, $S = \{(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, y) : \mathbf{v} \in (E \cup A), y \in \{-1, +1\}\}$, where $y = +1$ indicates information transmission from $\mathbf{v}^{(1)}$ to $\mathbf{v}^{(2)}$.

1.4.2 Learning transmissions

Given S one must first find a function such that $f'(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}) \approx y$, and we use an SVM (Boser et al., 1992) to find this function. To simplify notation let $\mathbf{x} = (\mathbf{v}^{(1)'} \mathbf{v}^{(2)'})' \in \mathbb{R}^{114}$, and the function mapping $\mathbf{x}$ to y be f . We will refer to $\mathbf{x}$ as an *example* and y as the corresponding *label*. An SVM finds a hyperplane which separates the set of examples into those which are positively labelled and those which are negatively labelled. It does this with maximal margin, which often ensures good generalisation to unseen data. A further advantage of SVMs is that they can operate in a kernel defined feature space and hence model non-linear functions without explicit computation of the new feature space. See Shawe-Taylor and Cristianini (2004) for an overview of kernel methods.

As a first step to learning transmissions, the complete dataset is standardised so that the examples have zero mean and unit standard deviation. They are then randomly sampled into a subset of size 15,000. This subset is used for choosing the SVM penalty parameter C , and the kernel parameters. For this model selection stage, we use a linear kernel with parameter values selected as $C \in \{2^2, 2^3, \dots, 2^{11}, 2^{12}\}$, and also the Radial Basis Function (RBF) kernel with $C \in \{2^2, 2^4, \dots, 2^{10}, 2^{12}\}$. The RBF kernel is given by $\kappa(\mathbf{x}, \mathbf{z}) = \exp(-\|\mathbf{x} - \mathbf{z}\|^2 / 2\sigma^2)$ with kernel width $\sigma \in \{2^{-4}, 2^{-3}, \dots, 2^1\}$ in this case. As the dataset has many more negative examples than positive ones, the penalty on the errors on positive examples is weighted according to $C^- = C\gamma$ where $\gamma \in \{2^2, 2^3, 2^4\}$.

In order to choose a set of parameters we use k -fold Cross Validation (CV) to evaluate prediction error. In this procedure, the dataset is split into

k equal sets. One set is kept back for evaluating error, and the remaining are used for training. This is repeated a total of k times with a different test set used each time, and this whole process is repeated for each unique set of SVM parameters. In our case $k = 3$. The output at the model selection phase is the set of parameters for the SVM which result in the lowest error.

Following model selection, one would like to obtain an unbiased estimate of the error obtained on an unseen set of examples. The model selection phase resulted in the selection of the RBF kernel, $C = 2,048$, $\sigma = 1$ and $\gamma = 32$. We use the set of 71,755 examples disjoint from that used during model selection, and a 5-fold cross validation procedure to obtain an average error with the SVM using this set of parameters. The resulting balanced error is 0.092 (0.000)⁸ which compares favourably to the error obtained on predicting no transmissions, which is 0.5. The error on the positive examples is 0.111 (0.000) and the error on the negative ones is 0.073 (0.000). With the linear kernel, $C = 128.0$ and $\gamma = 16$, the best balanced error rate is 0.235 (0.000) with errors of 0.276 (0.000) and 0.195 (0.000) on positive and negative examples.

The model weights describe the net influence of the various factors determining whether the information is transmitted from Ego/Sender to Alter/Recipient (Table 1.2). These weights are derived from a model that predicts the pattern of transmission with error. As smaller magnitude weights can be due to prediction error, we concentrate on the large weights and cut out the smaller ones where the relative size of the noise is larger. One of our findings is that in our model, the characteristics of Alters matter more than the characteristics of Egos, as seventeen of the twenty variables with the highest weights belong to Alters. This is surprising because it is the Ego/Sender who decides whether to relay the information. There are two explanations for this finding, one is technical the other is substantive. The technical explanation starts from the recognition that we have considerably fewer fully observed Alters than Egos. The characteristics Alters are predicted and not actually observed. Because we generated the characteristics of recipient and non-recipient Alters differently, and non-transmitting ties always involve on one end a non-recipient Alter and transmitting ties a recipient one, this could have amplified the influence of Alter characteristics on whether or not a tie transmits the advisory.

The substantive explanation, on the other hand, rests on the assumption that the Egos decide to relay the information on the basis of the characteristics of Alters, although, as we will see not by estimating the recipients need

⁸ The value in parentheses is the standard deviation of the error.

Variable	Meaning	Finding Transmission is more likely if...	Weight
Ego's characteristics			
Perception			
1 Q17bisA#X	Risk perception of Campylo (after info)	Ego considers Campylobacter infection risky	76.002
Knowledge			
2 Q16#X	Seeking of info (after info)	Ego looked for extra information	-56.268
Social Network variables			
3 Q44BX	Number of weekly direct contacts with colleagues	Ego has fewer direct contacts with colleagues	-64.349
Alter's Characteristics			
Demographic variables			
4 Q185\$X	Age	Alter is younger	-51.404
5 Q184\$	Gender	Alter is female	-108.127
6 Q186\$	Education	Alter has less education	-60.097
7 Q187\$_2	Profession: enterprise head, artisan, merchant	Alter is not enterprise head, artisan, merchant	-46.412
Social Network variables			
8 Q178C\$	Frequency of contact with colleagues	Alter spends more time with colleagues	-61.636
9 Q178B\$	Frequency of contact with neighbors	Alter spends more time with neighbors	-57.886
10 Q179M\$_3	Group membership parental assoc	Alter is not a member of parental associations	-51.585
11 Q180C\$X	Number of weekly direct contacts with family members	Alter contacts many family members	46.635
12 Q183EM\$_2	Tie with discussion about Food Risk	Alter talks to household member about food risk	75.639
13 Q196\$X	Frequency of Internet connection	Alter uses the Internet less frequently	125.049
Experience			
14 Q4A\$	Personally cook	Alter does not cook often	-93.724
15 Q24\$X	How often eats chicken (fowl)	Alter eats poultry less often	48.795
Knowledge			
16 Q1A\$	General knowledge about food risk	Alter thinks s/he knows more about food risk	94.683
17 Q21\$	Previous knowledge about Campylobacter	Alter had no previous knowledge about Campylobacter	94.127
18 Q32\$X	Seeking of info (after info)	Alter did look for more info on Campylobacter	-63.833
Perception			
19 Q33A\$X	Risk perception of Campylobacter (after info)	Alter considers less risky Campylobacter	-86.062
20 Q20A\$	Finds info credible	Alter finds info credible	49.843

Table 1.2 Variables with the highest weights for predicting transmission.

for the message, which they may not know, but on the basis of their perception of the recipients general interest in food risk. Another mechanism that can explain the importance of Alters characteristics is that Recipients elicit the advisory. For instance, women raised to be more attentive to others will be more likely to receive the information by the way they communicate with Egos. Because the technical reasons do play some part in the elevated importance of Recipient characteristics the extent of which we cannot tell, we make our substantive claim with caution.

The factors can be sorted into five categories: demographics, social networks, experience, knowledge and perception. To interpret the effect of social networks on transmission, we have to keep in mind that our unit of analysis is the social tie and not the individual. The three variables with high weights that describe Ego belong to the perception, knowledge and network categories and our model does not point to any demographic or experiential variables on the Ego side. Two of the three factors that describe transmitting Egos are not surprising. Egos who perceive *Campylobacter* more risky and those who looked up additional information are more likely to send the information to Alters. The third one, that the frequency of direct contacts with colleagues has a negative effect, will be explained below.

A tie is more likely to transmit the information about *Campylobacter* if the Alter is female, young, has less formal education and if she is not a self-employed entrepreneur by profession.

Of the social network characteristics of Alters, spending more time with neighbors, family members and colleagues, increases the chances of a tie to transmit. Interestingly, frequent contacts with colleagues by Ego have the opposite effect. Ties are more likely to transmit if Ego has fewer contacts with, while Alter spends more time with colleagues. This, however, is less of paradox than it seems.

We observe that this apparent paradox involves two different aspects of collegial ties: quantity and intensity. The second part of this seeming contradiction, that intensity of ties has a positive effect on transmission, is not unexpected. People who spend more time together with colleagues will be more likely to hear about topics unrelated to work. We find the same relationship for neighbors.

The second part requires more explanation that involves our unit of analysis: ties and not people. Egos who are well-connected to colleagues in our data set will contribute many collegial ties. This means that if Ego has many collegial ties but only a few will carry the information, most ties with collegially well-connected Egos will not carry the information. When we consider ties, rather than people what matters is not whether the number of

transmission rises with the number of ties but whether it rises at a higher rate than ties do. If the non-transmitting ties rise faster than transmitting ones, the relationship for ties between transmission and the number of ties will be positive for people (the more ties the more transmission) but we will observe a negative relationship for ties (the more ties the less likely a tie will transmit). Collegial ties thus have a diminishing marginal return: the first few colleagues will raise the chances of Egos telling some of them about the advisory but working with ten as opposed to twenty colleagues makes little difference.

Why is it then that the number of family contacts increases and not decreases the probability of transmission? The answer is that people have fewer familial than collegial contacts. It seems that the general relationship between number of ties and transmission is such that for the first few ties, not just the number but also the rate of transmission rises. Ego must have a certain number of ties to find one Recipient who is interested, the very first or second tie may not make much of a difference, the third and fourth does. Therefore, what we learn from our model is that higher intensity of collegial and neighborly ties of the recipient will make it more likely that she or he receives the information. The number of ties will yield an increasing return in the beginning but a decreasing one after a certain point.

We also included measures of formal affiliation to organizations as a measure of social networks. We learn that transmission is less likely to be targeted at people who are members of parent associations and other formal associations do not seem to make a difference.

An indirect measure of a type of connectedness is the frequency with which people use the Internet. Our survey was conducted through the web, thus it seems strange that people who spend more time on the web are not more but less likely to receive the advisory. Yet this result is consistent with our finding that most of the transmission happens face to face (see Table 1.1).

Experience and practices of Alters seem to indicate that transmission is not driven by the need of the recipient to know about *Campylobacter*. Those who cook less and eat chicken more often are less likely to receive the information about disease. One possible explanation is that Ego assumes that people who cook a lot and eat lots of chicken are more likely to already be aware of the advisory. What seems to attract the information is if Alter has a general interest in food risk. Ties to recipients who report that they have more general knowledge and then seek additional information on *Campylobacter* once they received the advisory, *i.e.*, ties to people

who are alert and curious about the topic are the most likely to carry the transmission.

As for Alters' perceptions, information seems to decrease their fears as the advisory communicates the measures that can be taken to avoid the infection. This creates a break on the transmission: if perceiving riskiness increases the urge to send the advisory but receiving the information reduces the sense of risk, at the next step, there will be fewer people relaying the information.

1.4.3 Modelling diffusion

The SVM model which predicts information transmission is applied to a set of random graphs in this stage. We artificially created a set of vertices use these vertices as a basis of a graph, and select a random subset of the vertices to have information about *Campylobacter*. Using the SVM model, and various initial setups, we observe how information is diffused within the graph (see Algorithm 1 of Appendix C).

At the end of the algorithm, several measurements made during the simulation are output. The quantities $v_1, \dots, v_m$ are the proportions of people with information at each iteration. Another quantity of interest is ξ which is the total number of receivers/total number of unique senders. This is a measure of the “fanout” of information, *i.e.* how many vertices receive information directly by each sender on average (see Figure 1.1).

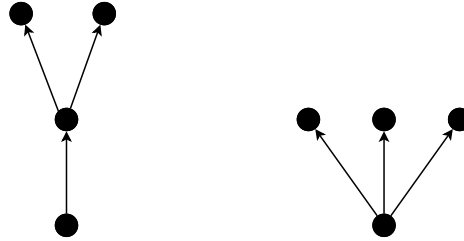


Figure 1.2 An illustration of the difference between μ_h and ξ measures, in which edges represent information transmissions between vertices. For the graph on the left there is one original sender, two senders and three transmissions and receivers, hence $\mu_h = 3$ and $\xi = 3/2$. For the one on the right there is only one sender and $\mu_h = 3$ and $\xi = 3$.

The simulation is run using 10,000 vertices, with an SVM trained using 20,000 random examples from the generated transmissions. We use the SW model to generate vertices with neighbors $k \in \{10, 15\}$, re-wiring probability

$p_s \in \{0.01, 0.05, 0.1, 0.2\}$ and initial information probability $a \in \{0.1, 0.2, 0.5\}$. Algorithm 1 is run five times for each set of parameters and the results are shown in Table 1.3. The same test is repeated with the Erdős-Rényi random graph with $p_e \in \{0.001, 0.002, 0.003, 0.004\}$.

p_s	k	a	μ_h	ξ	$v_1 - v_0$	$v_2 - v_1$	$v_3 - v_2$
.01	10	.1	3.544 (.124)	1.630 (.065)	.106 (.005)	.040 (.006)	.013 (.006)
.01	10	.2	2.677 (.069)	1.196 (.030)	.135 (.010)	.032 (.014)	.006 (.015)
.01	10	.5	1.835 (.031)	0.528 (.017)	.124 (.016)	.008 (.016)	.000 (.015)
.01	15	.1	4.499 (.075)	1.771 (.050)	.127 (.009)	.060 (.013)	.015 (.012)
.01	15	.2	3.363 (.060)	1.249 (.038)	.164 (.007)	.041 (.008)	.006 (.009)
.01	15	.5	2.229 (.041)	0.493 (.016)	.136 (.006)	.011 (.007)	.001 (.006)
.05	10	.1	3.609 (.097)	1.615 (.039)	.104 (.009)	.043 (.011)	.012 (.012)
.05	10	.2	2.676 (.036)	1.188 (.026)	.136 (.004)	.033 (.006)	.007 (.006)
.05	10	.5	1.815 (.034)	0.526 (.027)	.119 (.005)	.010 (.006)	.001 (.006)
.05	15	.1	4.782 (.115)	1.809 (.031)	.131 (.008)	.063 (.012)	.017 (.014)
.05	15	.2	3.343 (.053)	1.222 (.019)	.157 (.004)	.043 (.006)	.007 (.009)
.05	15	.5	2.242 (.031)	0.497 (.011)	.137 (.003)	.011 (.005)	.000 (.006)
.1	10	.1	3.681 (.033)	1.577 (.021)	.100 (.006)	.044 (.008)	.016 (.007)
.1	10	.2	2.713 (.088)	1.196 (.039)	.136 (.007)	.036 (.007)	.007 (.008)
.1	10	.5	1.831 (.016)	0.528 (.010)	.121 (.009)	.011 (.011)	.001 (.010)
.1	15	.1	4.812 (.128)	1.761 (.031)	.130 (.009)	.061 (.015)	.020 (.014)
.1	15	.2	3.373 (.037)	1.226 (.031)	.158 (.007)	.043 (.008)	.007 (.008)
.1	15	.5	2.244 (.057)	0.495 (.016)	.139 (.005)	.011 (.005)	.001 (.005)

Table 1.3 *Results from the information diffusion simulation using the SW model. Standard deviations shown in parentheses. The re-wiring probability is p_s and k is the initial number of neighbors. The initial information probability is a , μ_h is the average hop distance and ξ is total number of receivers/total number of unique senders. The proportion of vertices with information is recorded as $v_0, \dots, v_3$.*

p_e	a	μ_h	ξ	$v_1 - v_0$	$v_2 - v_1$	$v_3 - v_2$
.001	.1	2.900 (.098)	1.394 (.015)	.063 (.011)	.031 (.015)	.012 (.015)
.001	.2	2.122 (.081)	1.141 (.027)	.090 (.011)	.028 (.014)	.006 (.015)
.001	.5	1.460 (.013)	0.624 (.026)	.092 (.008)	.008 (.007)	.001 (.006)
.002	.1	4.112 (.056)	1.560 (.016)	.099 (.005)	.056 (.010)	.022 (.012)
.002	.2	2.790 (.065)	1.148 (.032)	.130 (.010)	.040 (.015)	.009 (.015)
.002	.5	1.829 (.028)	0.539 (.010)	.120 (.005)	.010 (.006)	.001 (.006)
.003	.1	5.130 (.167)	1.678 (.036)	.129 (.010)	.073 (.018)	.022 (.020)
.003	.2	3.455 (.102)	1.210 (.029)	.159 (.009)	.044 (.012)	.009 (.013)
.003	.5	2.225 (.027)	0.490 (.011)	.135 (.009)	.011 (.007)	.001 (.006)
.004	.1	6.029 (.179)	1.765 (.034)	.141 (.009)	.078 (.016)	.022 (.018)
.004	.2	4.123 (.033)	1.264 (.021)	.178 (.007)	.049 (.008)	.008 (.010)
.004	.5	2.721 (.022)	0.488 (.010)	.152 (.011)	.013 (.011)	.001 (.010)

Table 1.4 *Results from the information diffusion simulation using the Erdős-Rényi model. The probability of an edge is p_e .*

The results with the Erdős-Rényi graphs are shown in Table 1.4. In general we observe that the number of new vertices receiving information decreases at each iteration, and one would expect this decrease as the probability that a vertices neighbors also have information increases with i . Since p_e is probability of an edge the mean number of edges is expected to be np_e , and μ_h should increase with p_e . Table 1.4 shows that this is the case, however a doubling in the probability of an edge from 0.1 to 0.2 results in less than double the average hop distance. Note that the more neighbors a vertex has, the more likely that a greater number of those neighbors without information are suitable candidates for transmission (as learnt by the SVM). However, it is clear that as each vertex has more neighbors, the chance of an information-containing vertex coming across a neighbor which also has information increases. In a similar way, an increase in initial information probability corresponds with smaller average hop distances, though a doubling of a results in μ_h which is greater than half of the original value.

The values of ξ capture a different aspect of the information diffusion. One would expect a higher value of a to imply more neighbors with information for each vertex and hence a lower ξ value, and in general this is the case. When $a = 0.5$, an increase in the value of p_e results in a decrease in ξ possibly since there are more people who are able to send information and fewer who can receive. Notice also that the values of $v_{i+1} - v_i$ have an interesting trend: for $i = 0$ the increase is greatest in most cases when $a = 0.2$ compared to when it is either 0.1 or 0.5

The SW results are given in Table 1.3. Note that the Erdős-Rényi graph is similar to a SW model with a re-wiring probability of 1. Hence, a useful comparison is between the Erdős-Rényi graphs with $p_e = 0.003$ and the SW graphs with $k = 15$ and $p_s = 0.1$. The interesting differences in this case are those with the values of ξ . The values of ξ are smaller in the Erdős-Rényi graphs, and since surprisingly the total number of receivers are approximately the same, it implies that there are more unique senders in these graphs. Recall that with high clustering the chance of a friend of a friend being a friend is high and hence fewer senders are able to receive the same number of people as compared with the random connectivity of the Erdős-Rényi graphs.

Another interesting finding is that at the lowest level of $a(a=.1)$ the Erdős-Rényi graph results in more total number of Alter recipients than the SW network, but as a increases, the difference disappears. This implies that if the initial broadcasting of the message reaches only a few people, SW networks are less efficient but if the number of Egos gets above a certain pro-

portion of the population there is no difference between the two network types in terms how many Alters the diffusion process delivers.

The overall number of receivers does not vary significantly with changes in p_s , however for fixed p_s and k the total number of receivers is highest when $a = 0.2$. This latter trend is not generally observed with the Erdős-Rényi graphs. Clearly, the total number of transmission rises as more edges can accommodate transmissions and falls as the network becomes saturated. The peak is generally before 0.1 with Erdős-Rényi and approximately 0.2 with SW graphs.

Several of the other trends present in Table 1.3 are trivial. An increase in k from 10 to 15 always results in higher transmission since each vertex has more edges and hence more chance of passing information. The length of information paths is short for high values of a , implying that the person who receives information passes it onto many others but those receivers rarely pass it on. Furthermore, when $a = 0.5$ the information often only travels along a path of length 2.

In summary this relates to information propagation in the following ways: network structure does not seem to influence the total number of people receiving information. What is important is the number of connections and the probability of having information in the first place. Clearly, not every pair of people will facilitate a transmission and hence if too few people are provided with information, then it may stop before reaching everyone interested in it. Similarly, a saturated network does not permit a lot of transmissions. In the SW model the maximum total number of receivers occurred when 20% of the population was provided with the information. Since $v_{i+1} - v_i$ always decreases with i , this information is clearly in contrast with “viral” information spread in all of the scenarios presented.

Transmission visualisation

The measures recorded in Tables 1.3 and 1.4 give a good idea of the information diffusion processes occurring in the generated graphs. We additionally consider the visualisation of graph transmissions. We start with the SW model with $n = 10,000$, $k = 15$, $p_s = 0.1$, and a value of initial information probability of $a = 0.1$. In this particular instance, the total number of recipients (including the initial ones) is 2,897. Disregarding orientation of information transmission, 2,346 persons belong to an unique connected components, while the remaining 551 persons fall into much smaller components (the largest having thirty-five members). However, despite constant progress in graph visualisation (Di Battista et al., 1999; Herman et al.,

2000), representing the main connected component in a legible remains impossible. We rely therefore on two simplifying assumptions.

The first one consists in building a clustering of the nodes of the main connected component (Schaeffer, 2007; Fortunato, 2010): we find groups of individuals which are more likely to transmit or receive the information inside their group than to members of other groups. Then, rather than displaying the original large graph, we draw a graph of the clusters: each node corresponds to a group of people from the original graph (the surface of the node is proportional to the number of persons in the cluster). The edges between nodes indicate information transmission between members of the corresponding clusters. Concretely, we use maximal modularity clustering (Newman and Girvan, 2004) with the algorithm described in (Noack, 2007; Noack and Rotta, 2009) using the implementation provided by Andreas Noack⁹. The visualisation of the clustered graph is done with Fruchterman-Reingold algorithm (Fruchterman and Reingold, 1991) as implemented in the Igraph R package (Csárdi and Nepusz, 2006; R Development Core Team, 2010).

The clustering process finds fifty clusters: less than 3.7 percent of information transmission happens between clusters while the rest takes place inside clusters. This provides a validity index for the clustering: ignoring information propagation outside of a cluster will not introduce major distortions in the analysis. The resulting display is given by Figure 1.3, upper left panel. The figure shows also information propagation in the graph. Apart from the upper left panel, each graph of this figure shows the number of persons that have received the information at each time step of the propagation (the surface of each node encodes the number of persons). The upper right graph corresponds to the initial receivers (611 Egos), while arrows show information propagation from one cluster to another. The left graph of the second line shows the number of receivers after one step of propagation (*i.e.*, Alters 1), etc. The fact that most of the propagation happens between clusters during the first two steps is easily explained by two factors. Firstly, the initial growth is the largest one and corresponds the largest numbers of transmissions: it should generate the largest part of the external transmissions. Moreover, the clustering algorithm used is based on modularity maximisation. Modularity is a quality criterion for graph clustering that rewards putting connected nodes in the same cluster. However, the reward is inversely proportional to the degree of the nodes. Therefore, the obtained clustering tends to put high degree nodes in different clusters.

⁹ Available at <http://code.google.com/p/linloglayout>.

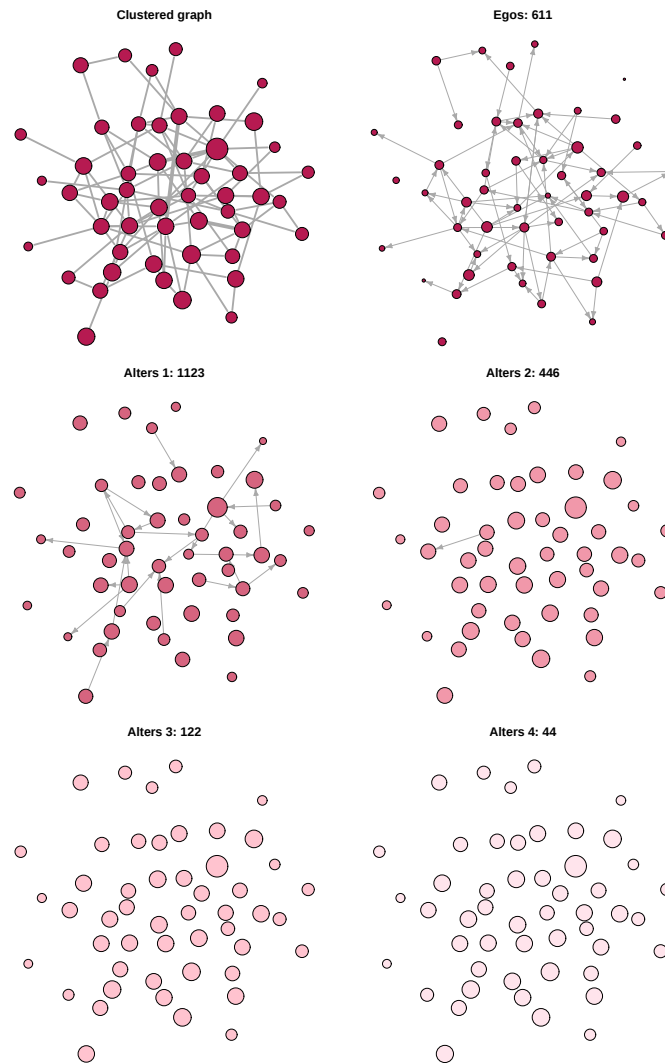


Figure 1.3 A clustered representation of the largest connected component of propagation graph used in Section 1.4.3 (upper left panel) and a clustered representation of information propagation, see main text for details.

As the first information transmission is the one in which persons tend to pass knowledge to the largest number of alters, those egos are more likely to be assigned to distinct clusters than the transmitters of the following steps.

As most of information propagation happens inside clusters, focusing

on one cluster provides a good idea of the general transmission. We display information propagation in the largest cluster on Figure 1.4. To avoid missing information propagation, the cluster was extended in the following way: when information flows *from* a person *in* the cluster under consideration *to* a person in *another* cluster, the recipient is added to the cluster. In the present case, the cluster grows from an initial size of 96 persons to 112 persons. Then, the result presented in Figure 1.4 is exactly the propagation that would have happened even if no other persons apart from those in the cluster had received initially the information.

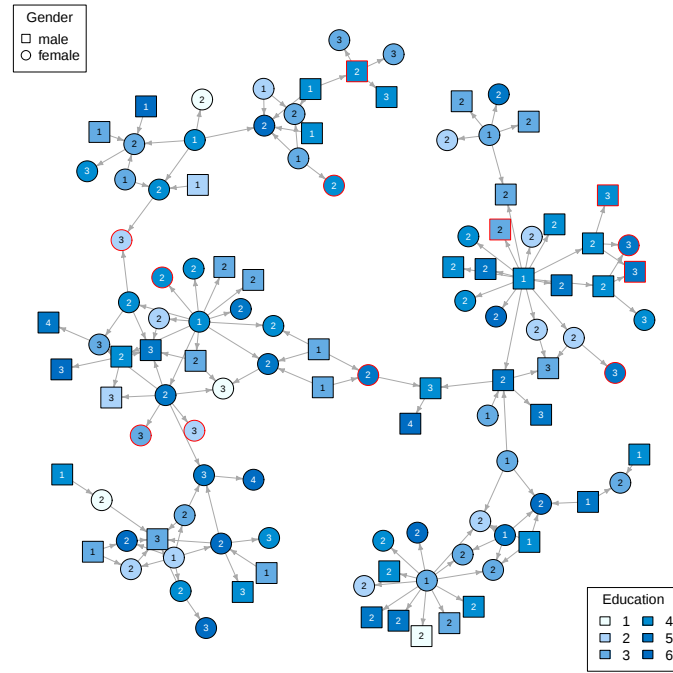


Figure 1.4 Information propagation in the largest cluster, extended as explained in the main text. Education levels are encoded by colours and genders by shape. The iteration number at which the information reached the person is written in the corresponding node. A node with a red border received also the information from a person in a distinct cluster.

1.5 Distribution of individual (vertex) characteristics

In this test we run a simulation using the SW model with parameters $p_s = 0.1$ and $k = 15$, and observe how the distribution of various characteris-

tics such as gender, age, etc. vary at each iteration. Transmissions are learnt using a sample of 20,000 examples, and an SVM is trained using the parameters found in Section 1.4.2. The simulation is run with 20,000 vertices for 3 iterations and repeated a total of 5 times with different random seeds. We observe that there are 1,495.2 Egos, and 2,689.6, 1,245.2, 369.4 new Alters iterations 1, 2 and 3, respectively.

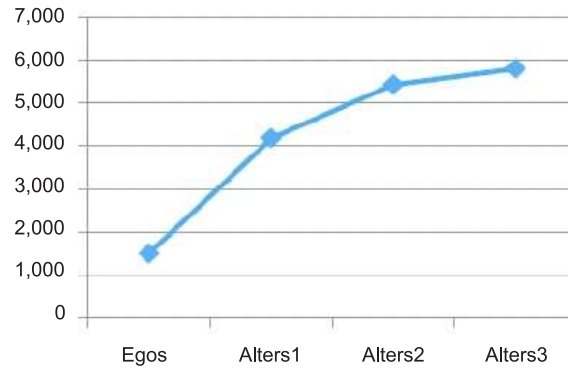


Figure 1.5 Cumulative number in our population who received the advisory in three iterations.

The simulation shows that after the initial jump the number of receivers taper down. The gender composition changes with each round of iteration. As we have seen earlier women are more likely to be the recipients of the information than men. This explains that there is an overall increase in the proportion of women. This increase mostly levels off after the first round of transmission. This is due not just to the fact that there are increasingly fewer new recipients of the advisory but also that subsequent transmissions are more gender balanced. In the second round the proportion of women actually decreases compared to the first one and levels off in the third, and final round (Table 1.5).

	Female	Male
All	0.509	0.491
Egos	0.535	0.465
Alters 1	0.667	0.333
Alters 2	0.602	0.398
Alters 3	0.598	0.402

Table 1.5 *Distribution of genders in simulated information diffusion.*

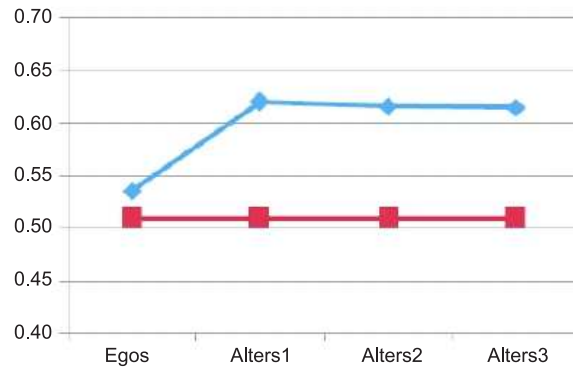


Figure 1.6 Blue: proportion of those who have the information. Red: over-all population proportion.

As the information spreads in our simulation we see a similar pattern for the average age of those who receive the advisory (Table 1.6). The blue graph in this chart is almost the mirror image of the previous one. There is an initial drop in age then it changes little.

	1	2	3	4	5	6	7	8	9	10	11	12
All	0.000	0.032	0.065	0.106	0.142	0.166	0.163	0.132	0.094	0.057	0.029	0.013
Egos	0.000	0.029	0.066	0.107	0.141	0.176	0.158	0.135	0.093	0.053	0.032	0.012
Alters 1	0.000	0.041	0.072	0.120	0.154	0.169	0.162	0.121	0.083	0.047	0.021	0.011
Alters 2	0.000	0.034	0.070	0.113	0.145	0.160	0.159	0.138	0.091	0.055	0.027	0.009
Alters 3	0.000	0.031	0.072	0.109	0.152	0.171	0.163	0.130	0.083	0.060	0.019	0.009

Table 1.6 *Distribution of ages.*

We have similar findings for education and general knowledge of food risk. Overall the pattern is that the decisive change is in the first round and then smaller changes occur for subsequent rounds which together with the ever decreasing number of new recipients results in a stable average.

	1	2	3	4	5	6
All	0.030	0.123	0.265	0.313	0.199	0.069
Egos	0.031	0.126	0.263	0.307	0.204	0.070
Alters 1	0.040	0.138	0.269	0.306	0.182	0.065
Alters 2	0.039	0.140	0.277	0.309	0.175	0.060
Alters 3	0.031	0.129	0.269	0.305	0.197	0.068

Table 1.7 *Distribution of education (1=Without degree/primary/BEPC, 2=CAP/BEP, 3=BAC, 4=BAC+2, 5=More than BAC+2, 6=Doctorate).*

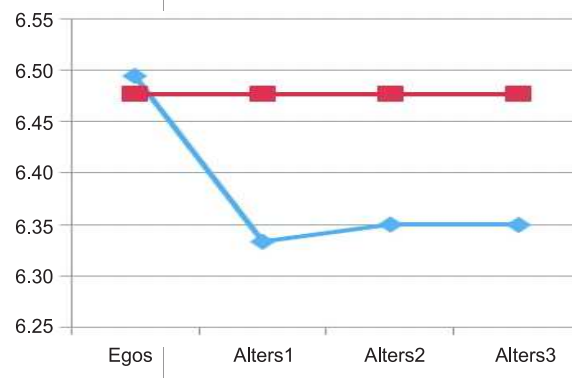


Figure 1.7 Blue: average age of those who have the information. Red: Average age of overall population.

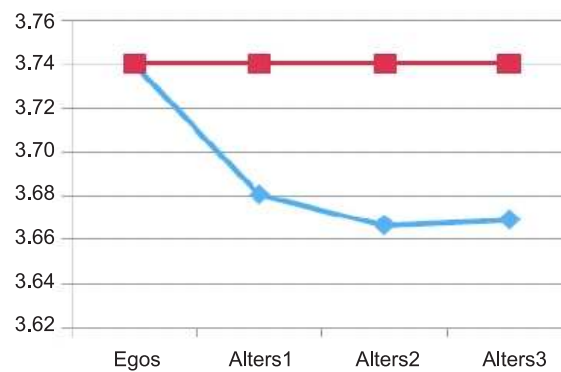


Figure 1.8 Blue: average education of those who have the information. Red: Average education of overall population.

The simulation shows that with the information spreading, the population it reaches is increasingly female, young, less educated and tend to be more knowledgeable about food risks in general, however, much of this shift takes place in the first round of the transmission. This is partly because the characteristics that make people more likely to receive the information are not making them more likely to send it further and it is also partly because of saturation as recipients with those characteristics not having the information in a person's social circle become more scarce.

In terms of occupational categories, the distribution shifts in a way that among the information recipients we find a smaller portion of retirees and people not working at the end of the third round than we had in the beginning (Table 1.7). The diffusion process reaches the economically active

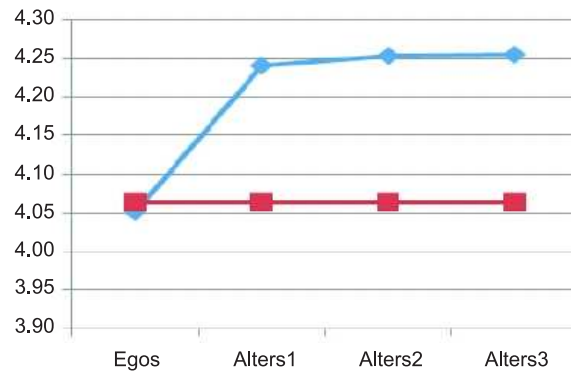


Figure 1.9 Blue: average food risk knowledge of those who have the information. Red: Average food risk knowledge of overall population.

people more successfully than the inactive ones, but, again, we see the same pattern: the first step is the largest one, and then movement is in the opposite direction but in much smaller steps. This pattern is in line with our earlier findings about the role of collegial ties, the type of ties only economically active people possess. Because collegial ties are usually within occupational categories, we can explain the seemingly contradictory findings that net of other factors Alters with less education are more likely to receive the information but the occupational groups with higher average education, such as professionals and employees, will grow faster than others among those who receive the information. It seems that professionals and employees get the advisory not because they are better educated but probably because the nature of the workplace interaction they have compared to workers. Within each group, however, it is the less educated who are more likely to be given the news about the *Campylobacter*.

When we rerun the diffusion using the comparable Erdős-Rényi model, the distribution of the characteristics are very similar and the differences are within the range of random error (Table 1.1).

1.6 Conclusion

The diffusion of the advisory about the *Campylobacter* shows that the broadcast model misses an important part of the spread of this type of information; the diffusion of the advisory from its initial audience to others. Unlike the two step process model that, in an attempt to identify opinion leaders, focuses on characteristics make people good senders of the information

	1	2	3	4	5	6	7	8
All	0.000	0.006	0.158	0.228	0.233	0.042	0.146	0.186
Egos	0.000	0.005	0.160	0.217	0.245	0.040	0.157	0.174
Alters 1	0.000	0.001	0.174	0.271	0.290	0.024	0.102	0.138
Alters 2	0.000	0.001	0.164	0.259	0.278	0.033	0.122	0.143
Alters 3	0.000	0.002	0.175	0.238	0.273	0.033	0.128	0.151

Table 1.8 *Distribution of professions.* (1=Agricultural workers, 2=Self-employed, 3=Cadres, 4=Professionals, 5=Employees, 6=Workers, 7=Retirees, 8= Other not working).

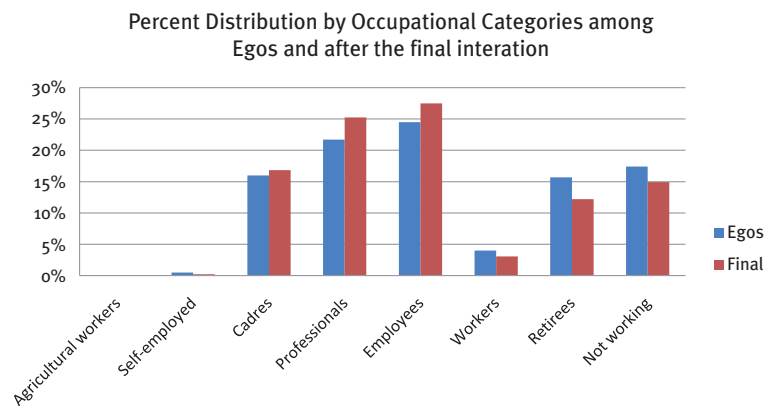


Figure 1.10 Percentage distribution of Occupational categories.

, we found that the characteristics that make people good recipients are equally, if not more important. Modeling the diffusion showed us that random networks and SW networks produce roughly the same results in terms of the proportion of the population that receives the information but the less clustered random networks achieve the result through longer hop distances and smaller ratio of receivers to unique senders. The composition of those who receive the information changes most in the first transmission and then it levels off not just because there are fewer and fewer recipients making it harder to change the cumulative total but also because new recipients in each round are less different from the overall population. This is partly because the characteristics that make someone more likely to receive the information are not those that make them more likely to pass it on. The mismatch between the two sets of characteristics will dampen transmission even in the absence of network barriers.

The diffusion of the *Campylobacter* advisory shows that word-of-mouth dispersion of the information, an aspect of the process the broadcast model ignores, is significant. It also demonstrates that while the first transmission from original Egos to Alters is the biggest part of this word-of-mouth dispersion, the Two-Step model is flawed because there are still substantial – although smaller – steps spreading the message it fails to recognize. The Two-Step model also focuses on the qualities of the sender (opinion leader) and our findings show that the recipients' qualities are important. Furthermore, our percolation model was able to build a complex dispersion process which allowed us to account for the heterogeneity of people and of their social networks. Our simulations also showed that not just the local but the global properties of networks matter. Replicating this study with other advisories will tell us the extent to which our empirical findings can be generalized and be of practical use to state agencies.

Appendix A List of features used in simulation

ID	Field Name
Q4	Gender
Q5	Age in categories
Q48	Education
Q7	Profession (8 categories)
Q51	Income
Q52	Single or married
Q53	Number of children
Q13	Size of village/town/city
Q20A	Personally cook
Q26	Frequency of eating poultry meat
Q28A	General knowledge about food risk
Q22	Personal experience with food risk
Q34	Previous knowledge about Campylobacter
Q16#	Seeking of info (after info)
Q55	Frequency of Internet connection
Q17bisA#	Risk perception of Campylobacter
Q37A, Q37B	Finds the information credible, understandable
Q37C, Q37D	Finds the information convincing to change behaviour, worrisome
Q42A-D	Contact frequency with friends, neighbours, colleagues, family
Q43M1-10	Organisation memberships (1 organisation per field, 10 options)
Q44A-D	Weekly contact with friends, colleagues, family, acquaintances
Q45	Friends closeness
Q47CM1-5	Tie with discussion about Health (6 categories)
Q47EM1-5	Tie with discussion about Food Risk (6 categories)
Q50	Number of people at physical place of work
Q46A-D	Homophily of contacts (age, gender, education, income)

Table A.1 *The list of features used for each vertex.*

Appendix B Data completion

There are $q = 4,496$ Egos, denoted by $E = \{\mathbf{c}_1, \dots, \mathbf{c}_q\}$ with $c_i \in \mathbb{R}^{62}$ for all i . The Egos collectively listed a total of 7,655 people with whom they discussed the Campylobacter information. Among them, there are $r = 301$ Alters who responded to the survey for whom we got the set of 62-dimensional vectors, $A = \{\mathbf{d}_1, \dots, \mathbf{d}_r\}$. The information recorded for the set of other receivers (7,204 Alters) is limited to their gender, age, profession and education¹⁰. To make a prediction for when information transmission occurs between two people, one needs the complete, immediate, social network for Egos, and the 62-dimensional vector for Alters. Hence, we completed

¹⁰ An additional 150 Alters failed to remember the information provided to them

the data by generating a set of non-receivers and completing the set of receivers (Alters) for each Ego.

To compute non-receivers, we consider 2 characteristics: one is the number of times Egos have direct contact with friends, colleagues, family members out of the household and acquaintances in a typical week (Q44A-D). The total number of non-receivers is computed using Q44A-D. The other is whether most of the Egos' contacts are homophilic (*i.e.*, they are similar in their characteristics to the Ego, Q46A-D). Homophily can be related to sex, age, education and income. The set of homophiles for the i th person $H_i \subseteq E$ is the set of people with identical features as indicated by Q46A-D. For example, if a person $\mathbf{c}_i$ states in Q46A-D that most of their contacts have the same age and gender, then H_i is the set $H_i = \{\mathbf{c}_j : \mathbf{c}_j \in E, \mathbf{c}_{ik} = \mathbf{c}_{jk}, \mathbf{c}_{i\ell} = \mathbf{c}_{j\ell}, i \neq j\}$ where $\mathbf{c}_{ik}$ is the k th element of $\mathbf{c}_i$, and k and ℓ are the indices corresponding to age and gender respectively. Similarly the set of non-homophiles is $H'_i = E \setminus (H_i \cup \{\mathbf{c}_i\})$. Given the total number of contacts N_i for the i th person and a probability h of being a homophile, a set of hN_i vectors is randomly sampled from H_i and $(1 - h)N_i$ vectors are sampled from H'_i . In our case $h = 0.7$.

As previously stated, the Egos record only the age, gender, profession and education of their Alters, and one needs to complete the data using A . For the i th person in E and the j th Alter recorded, we find a random element of the set of Alter homophiles $G_i = \{\mathbf{d}_j : \mathbf{d}_j \in A, \mathbf{d}_{ik} = \mathbf{d}_{jk}, \mathbf{d}_{i\ell} = \mathbf{d}_{j\ell}, \mathbf{d}_{im} = \mathbf{d}_{jm}\}$ where $\mathbf{d}_{ik}$ is the k th element of $\mathbf{d}_i$, and k, ℓ, m are the indices corresponding to age, gender and education respectively.

Appendix C Information diffusion algorithm

In Algorithm 1 the vertices of the graph are generated using a multivariate normal distribution, with the mean vector and covariance matrix computed using the Egos and Alters from the survey data. Edges in the graph are added according to the Erdős-Rényi or Small World graph models. At line 4, $a|V|$ vertices are selected at random and marked as having information. The following inner for loop iterates through all of the edges in the graph, and makes a prediction for whether information is transmitted along that edge using the characteristics of the vertices and the Support Vector Machine (SVM) model learnt at line 2. Note that the original survey data are not necessarily identically and independently distributed, but we assume so in order to apply the SVM.

Algorithm 1 Pseudo code for information diffusion simulation.

- 1: **Input:** Graph size n , iterations m , initial proportion with information a
 - 2: Learn SVM model of information transmission
 - 3: Create $G = (V, E)$ with $|V| = n$ randomly generated vertices
 - 4: Randomly select a set I_0 of $a|V|$ vertices to mark with information
 - 5: **for** $i = 1$ to m **do**
 - 6: **for** $j = 1$ to $|E|$ **do**
 - 7: Make information transmission prediction along edge j
 - 8: **end for**
 - 9: I_i is the set of vertices with information, let $v_i = |I_i|/n$
 - 10: **end for**
 - 11: Set μ_h as total number of transmissions/total number of original senders
 - 12: Set ξ as total number of receivers/total number of unique senders
 - 13: **Output:** Proportions $v_1, \dots, v_m$ of people with information, ξ and μ_h .
-

Bibliography

- Abbott, A. 1995. Things of Boundaries. *Social Research*, **62**, 857–882.
- Adams, J. D., Black, G. C., Clemmons, J. R., and Stephan, P. E. 2005. Scientific teams and institutional collaborations: evidence from U.S. universities. *Research Policy*, **34**, 259–285.
- Afgan, N. H., Carvalho, M. G., Pilavachi, P. A., and Martins, N. 2007. Evaluation of natural gas supply options for south east and central Europe. Part 1: indicator definitions and single indicator analysis. *Energy Conversion and Management*, **48**, 2517–2524.
- Agency for Toxic Substances and Disease Registry. 2007. *Case studies in environmental medicine (CSEM) nitrate/nitrite toxicity*. www.atsdr.cdc.gov/csem/nitrate/nitrate.html.
- Aksnes, D. W. 2006. Citation rates and perceptions of scientific contribution. *Journal of the American Society for Information Science and Technology*, **57**, 169–185.
- Albanese, D., Merler, S., Jurman, G., Visintainer, R., and Furlanello, C. 2009. *MLPY Machine Learning PY*. <http://mloss.org/software/view/66/>.
- Albert, R., and Barabási, A.-L. 2002. Statistical mechanics of complex networks. *Reviews of Modern Physics*, **74**, 47–97.
- Albert, R., Jeong, H., and Barabási, A.-L. 2000. Error and attack tolerance of complex networks. *Nature*, **406**, 378–381.
- Alcácer, J., and Gittleman, M. 2006. Patent citations as a measure of knowledge flows: the influence of examiner citations. *The Review of Economics and Statistics*, **88**, 774–779.
- Alexa, N., Bárdos, R., Szántó, Z., and Tóth, I. J. 2008. Corruption risks in the business sector. Integrity system country study (part two). *Transparency International*.
- Ali, S. N., Young, H. C., and Ali, N. M. 1996. Determining the quality of publications and research for tenure or promotion decisions: a preliminary checklist to assist. *Library Review*, **45**, 39–53.
- Allard, A., Noël, P.-A., L., Dube, and Pourbohloul, B. 2009. Heterogeneous bond percolation on multitype networks with an application to epidemic dynamics. *Physical Review E*, **79**.
- Allesina, S., and Bodini, A. 2004. Who dominates whom in the ecosystem? Energy flow bottlenecks and cascading extinctions. *Journal of Theoretical Biology*, **230**, 351–358.
- Allesina, S., Bodini, A., and Bondavalli, C. 2006. Secondary extinctions in ecological networks: bottlenecks unveiled. *Ecological Modelling*, **194**, 150–161.
- Allesina, S., Bodini, A., and Pascual, M. 2009. Functional links and robustness in food webs. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, **364**, 1701–1709.
- Amaral, L. A. N. 2005. Novel collaborations within experienced teams lead to best research outcomes. *Annals of Vascular Surgery*, **19**, 753–754.
- Amaral, L. A. N., and Ottino, J. 2004. Complex networks: Augmenting the framework for the study of complex systems. *European Physical Journal B*, **38**, 147–162.
- Amaral, L. A. N., and Uzzi, B. 2007. Complex systems—a new paradigm for the integrative study of management, physical, and technological systems. *Management Science*, **53**, 1033–1035.

- Amaral, L. A. N., Scala, A., Barthélemy, M., and Stanley, H. E. 2000. Classes of small-world networks. *Proceedings of the National Academy of Sciences U.S.A.*, **97**, 11149–11152.
- Amenta, E., Carruthers, B., and Zylan, Y. 1992. A hero for the aged? The Townsend Movement, the political mediation model, and U.S. old-age policy, 1934–1950. *American Journal of Sociology*, **98**, 308–339.
- Andrews, K.T., and Edwards, B. 2005. The organizational structure of local environmentalism. *Mobilization*, **10**, 213–234.
- Ansell, C. 2003. Community embeddedness and collaborative governance in the San Francisco Bay Area Environmental Movement. Pages 123–144 of: *Social Movements and Networks*. Oxford University Press.
- Armstrong, E.A. 2005. From struggle to settlement: the crystallization of a field of lesbian/gay organizations in San Francisco, 1969–1973. Pages 161–187 of: *Social Movements and Organizations*. Cambridge University Press.
- Aronowitz, S., and Bratsis, P. 2002. *Paradigm Lost: State Theory Reconsidered*. University of Minnesota Press.
- Arrighi, G., and Silver, B. J. 1999a. Conclusion. Pages 271–289 of: *Chaos and Governance in the Modern World System*. University of Minnesota Press.
- Arrighi, G., and Silver, B. J. 1999b. Introduction. Pages 1–36 of: *Chaos and Governance in the Modern World System*. University of Minnesota Press.
- Arrighi, G., and Silver, B. J. 2001. Capitalism and world (dis)order. *Review of International Studies*, **27**, 257–279.
- Ashby, W. R. 1960. *Design for a Brain: The Origin of Adaptive Behavior*. Chapman & Hall.
- Babchuk, N., Bruce, K., and George, P. 1999. Collaboration in sociology and other scientific disciplines: a comparative trend analysis of scholarship in the social, physical, and mathematical sciences. *The American Sociologist*, **30**, 5–21.
- Baker, W. E. 1992. The network organization in theory and practice. Pages 397–429 of: *Networks and Organizations*. Harvard Business School Press.
- Balirwa, J. S., Chapman, C. A., Chapman, L. J., Cowx, I. G., Geheb, K., Kaufman, L., Lowe-McConnell, R. H., Seehausen, O., Wanink, J. H., Welcomme, R. E., and Witte, F. 53. Biodiversity and fishery sustainability in the Lake Victoria Basin: an unexpected marriage? *BioScience*, **2003**, 703–715.
- Ball, F., Mollison, D., and Scalia-Tomba, G. 1997. Epidemics with two levels of mixing. *The Annals of Applied Probability*, **7**, 46–89.
- Balthazard, P., Potter, R. E., and Warren, J. 2004. Expertise, extraversion and group interaction styles as performance indicators in virtual teams: how do perceptions of IT's performance get formed? *ACM SIGMIS Database*, **35**, 41–64.
- Bandura, A. 1977. *Social Learning Theory*. General Learning Press.
- Barabási, A.-L. 2003. *Linked: How Everything is Connected to Everything Else and What it Means for Business and Everyday Life*. Cambridge: Plume.
- Barabási, A.-L., and Albert, R. 1999. Emergence of scaling in random networks. *Science*, **286**, 509–512.
- Barabási, A.-L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., and Vicsek, T. 2002. Evolution of the social network of scientific collaborations. *Physica A*, **311**, 590–614.
- Barringer, B. R., and Harrison, J. S. 2000. Walking a tightrope: creating value through interorganizational relationships. *Journal of Management*, **26**, 367–403.

- Barsade, S. G., Ward, A. J., Turner, J. D. F., and Sonnenfeld, J. A. 2000. To your heart's content: a model of affective diversity in top management teams. *Administrative Science Quarterly*, **45**, 802–836.
- Bascompte, J. 2007. Networks in ecology. *Basic and Applied Ecology*, **8**, 485–490.
- Baum, J. A. C., and Oliver, C. 1991. Institutional linkages and organizational mortality. *Administrative Science Quarterly*, **36**, 187–218.
- Baum, J. A. C., McEvily, B., and Rowley, T. 2007. Better with age: the longevity and the performance implications of bridging and closure. *Rotman School of Management, University of Toronto, Working paper n. 1032282*.
- Becker, G. S. 1968. Crime and punishment: an economic approach. *Journal of Political Economy*, **76**, 169–217.
- Bennani-Chraïbi, M., and Fillieule, O. 2003. *Résistances et Protestations dans les Sociétés Musulmanes*. Presses de Sciences Po.
- Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C., and West, G. B. 1995. Explaining the level of bridewealth. *Current Anthropology*, **35**, 311–316.
- Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C., and West, G. B. 2007. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences U.S.A.*
- Bodini, A. 1998. Representing ecosystem structure through signed digraphs. Model reconstruction, qualitative predictions and management: the case of a freshwater ecosystem. *Oikos*, **83**, 93–106.
- Bodini, A. 2000. Reconstructing trophic interactions as a tool for understanding and managing ecosystems: application to a shallow eutrophic lake. *Canadian Journal of Fisheries and Aquatic Sciences*, **57**, 1999–2009.
- Bodini, A., and Bondavalli, C. 2002. Towards a sustainable use of water resources: a whole-ecosystem approach using network analysis. *International Journal of Environment and Pollution*, **18**, 463–485.
- Bodini, A., Bellingeri, M., Allesina, S., and Bondavalli, C. 2009. Using food web dominator trees to catch secondary extinctions in action. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, **364**, 1725–1731.
- Bondavalli, C., Bodini, A., Rossetti, G., and Allesina, S. 2006. Detecting stress at the whole ecosystem level. The case of a mountain lake: Lake Santo (Italy). *Ecosystems*, **9**, 768–787.
- Borgatti, S. P., and Foster, P. 2003. The network paradigm in organizational research: a review and typology. *Journal of Management*, **29**, 991–1013.
- Borgatti, S. P., Everett, M. G., and Freeman, L. C. 2002. *Ucinet for Windows: Software for Social Network Analysis*. Analytic Technologies, Harvard.
- Borner, K., Maru, J. T., and Goldstone, R. L. 2004. The simultaneous evolution of author and paper networks. *Proceedings of the National Academy of Sciences U.S.A.*, **101**, 5266–5273.
- Boser, B. E., Guyon, I. M., and Vapnik, V. N. 1992. A training algorithm for optimal margin classifiers. Pages 144–152 of: *Proceedings of the 5th Annual Conference on Computational Learning Theory*. New York, NY, USA: ACM Press.
- Bowers, R. V. 1937. The direction of intra-societal diffusion. *American Sociological Review*, **2**, 826–836.
- Bowler, P. J., and Morus, I. R. 2005. *Making Modern Science: A Historical Survey*. Chicago: University of Chicago Press.
- BP. 2008. *BP Statistical Review of World Energy 2008*. Tech. rept. British Petroleum.

- Brass, D. J., Galaskiewicz, J., Greve, H. R., and Tsai, W. 2004. Taking stock of networks and organizations: a multilevel perspective. *Academy of Management Journal*, **47**, 795–817.
- Braudel, F. 1973. *Capitalism and Material Life, 1400-1800*. Harper and Row.
- Brauer, F. 2005. The Kermack McKendrick epidemic model revisited. *Mathematical Biosciences*, **198**, 119–131.
- Breiger, R.L. 1974. The duality of persons and groups. *Social Forces*, **53**, 181–190.
- Brown, J. R. 2000. Privatizing the university—the new tragedy of the commons. *Science*, **290**, 1701–1702.
- Burstein, P., and Linton, A. 2002. The impact of political parties, interest groups, and social movement organizations on public policy. *Social Forces*, **75**, 135–169.
- Burt, R. S. 1987. Social contagion and innovation: cohesion versus structural equivalence. *American Journal of Sociology*, **92**, 1287–1335.
- Burt, R. S. 1992. *Structural Holes: The Social Structure of Competition*. Harvard University Press.
- Burt, R. S. 1999. The social capital of opinion leaders. *Annals of the American Academy of Political and Social Science*, **566**, 37–54.
- Burt, R. S. 2005. *Brokerage and Closure: An Introduction to Social Capital*. Oxford University Press.
- Burt, R. S. 2008. Information and Structural Holes: Comment on Reagans and Zuckerman. *Industrial and Corporate Change*, **17**, 953–969.
- 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.
- Camerano, L. 1880. Dell'equilibrio dei viventi mercé la reciproca distruzione. *Atti della Reale Accademia della Scienze di Torino*, **15**, 393–414.
- Camerer, C., and Knez, M. 1996. Coordination, organizational boundaries and fads in business practices. *Industrial and Corporate Change*, **5**, 89–112.
- Campbell, J. L. 2005. Where do we stand? Common mechanisms in organizations and social movement research. Pages 41–68 of: *Social Movements and Organizations*. Cambridge University Press.
- Cardona, J., and Lacroix, C. 2008. *Chiffres Clés: Statistiques de la culture*. Paris: La Documentation française.
- Carroll, W. K., and Ratner, R. S. 1996. Master framing and cross-movement networking in contemporary social movements. *Sociological Quarterly*, **37**, 601–625.
- Carruthers, B. G., and Halliday, T. 1998. *Rescuing Business: The Making of Corporate Bankruptcy Law in England and the United States*. Oxford University Press.
- Castells, M. 1996. *The Rise of the Network Society*. Blackwell.
- Çetin, T., and Oguz, F. 2007. The reform in the Turkish natural gas market: A critical evaluation. *Energy Policy*, **35**, 3856–3867.
- Ceballos, G., and Ehrlich, P. R. 2002. Mammal population losses and the extinction crisis. *Science*, **296**, 904–439.
- Chandler, T. 1987. *Four Thousand Years of Urban Growth*. St. David's University Press.
- Chang, C.-C., and Lin, C.-J. 2001. *LIBSVM: a library for support vector machines*. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- Cheit, R. E., and Gersen, J. E. 2000. When businesses sue each other: An empirical study of state court litigation. *Law & Social Inquiry*, **25**, 789–816.

- Christakis, N. A., and Fowler, J. H. 2009. *Connected: The Surprising Power of our Social Networks and how they Shape our Lives*. Little Brown and Co.
- CIA. 2008. *The World Factbook*. Potomac Books.
- Clark, H. H., and Wilkes-Gibbs, D. 1986. Referring as a collaborative process. *Cognition*, **22**, 1–39.
- Clark, R. A., and Goldsmith, R. E. 2005. Market mavens: psychological influences. *Psychology and Marketing*, **22**, 289–312.
- Cohen, J. E. 1994. Lorenzo Camerano's contribution to early food web theory. Pages 351–359 of: *Frontiers in Mathematical Biology, Part V: Frontiers in Community and Ecosystem Ecology*. Springer-Verlag.
- Cohen, J. E., Briand, F., and Newman, C. M. 1990. *Community Food Webs: Data and Theory*. Springer-Verlag.
- Cohen, M. D., and Bacdayan, E. 1994. Organizational routines are stored as procedural memory: evidence from a laboratory study. *Organization Science*, **5**, 554–568.
- Cohen, R., Erez, K., Ben-Avraham, D., and Havlin, S. 2000a. Resilience of the Internet to Random Breakdowns. *Physical Review Letters*, **85**, 4626–4628.
- Cohen, S., Brissette, I., Skoner, D., and Doyle, W. 2000b. Social integration and health: the case of the common cold. *Journal of Social Structure*, **1**, 1–7.
- Coleman, J., Katz, E., and Menzel, H. 1957. The diffusion of an innovation among physicians. *Sociometry*, **20**, 253–270.
- Collins, R. 1998. *The Sociology of Philosophies: A Global Theory of Intellectual Change*. Cambridge: Harvard University Press.
- Compas, B. E., David, A., Haaga, F., Keefe, F. J., Leitenberg, H., and Williams, D. A. 1998. Sampling of empirically supported psychological treatments from health psychology: smoking, chronic pain, cancer, and bulimia nervosa. *Journal of Consulting and Clinical Psychology*, **66**, 89–112.
- Contractor, N. 1994. New approaches to organizational communication. Pages 39–65 of: Kovacic, B. (ed), *New Approaches to Organizational Communication*. SUNY Press.
- Corden, W. M., and Neary, J. P. 1982. Booming sector and de-industrialisation in a small open economy. *Economic Journal*, **92**, 825–848.
- Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R. V., Paruelo, J., Raskin, R. G., Sutton, P., and van den Belt, M. 1997. The value of the world's ecosystem services and natural capital. *Nature*, **387**, 253–260.
- Coyle, D. 2011. *The Economics of Enough: How to Run the Economy as If the Future Matters*. Princeton University Press.
- Cronin, B., Shaw, D. D., and La Barre, K. 2003. A cast of thousands: Coauthorship and subauthorship collaboration in the 20th century as manifested in the scholarly journal literature of psychology and philosophy. *Journal of American Society for Information Technology*, **54**, 855–871.
- Csárdi, G., and Nepusz, T. 2006. The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695. <http://igraph.sf.net>.
- Csermely, P. 2008. Creative elements: network-based predictions of active centres in proteins, cellular and social networks. *Trends in Biochemical Sciences*, **33**, 569–576.
- Csermely, P. 2009. *Weak Links: The Universal Key to the Stability of Networks and Complex Systems*. Springer Verlag.

- Cummings, J. N., and Kiesler, S. 2007. Who works with whom? Collaborative tie strength in distributed interdisciplinary projects. In: *Third International e-Social Science Conference*.
- D., Obstfeld. 2005. Social networks, the tertius iungens orientation, and involvement in innovation. *Administrative Science Quarterly*, **50**, 100–130.
- Dalton, R. 1994. *The Green Rainbow: Environmental Groups in Western Europe*. Yale University Press.
- Dalton, R. 2008. *Citizen Politics*. 5th edn. CQ Press.
- de Solla Price, D. J. 1963. *Little Science, Big Science-and Beyond*. New York: Columbia University Press.
- Degenne, A., and Forsé, M. 1999. *Introducing Social Networks*. Sage.
- Dekker, A. H. 2007. Realistic social networks for simulation using network rewiring. Pages 677–683 of: *International Congress on Modelling and Simulation*.
- Della Porta, D. 1995. *Social Movements, Political Violence and the State*. Cambridge University Press.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, **39**, 1–38.
- Derényi, I., Farkas, I., Palla, G., and Vicsek, T. 2004. Topological phase transitions of random networks. *Physica A*, **334**, 583–590.
- Devezas, T. 2001. The biological determinants of long wave behavior in socioeconomic growth and development. *Technological Forecasting and Social Change*, **68**, 1–57.
- Devezas, T., and Modelski, G. T. 2008. The Portuguese as system-builders: technological innovation in early globalization. Pages 30–57 of: *Globalization as Evolutionary Process: Modeling, Simulating, and Forecasting Global Change*. Routledge.
- Dewey, J. 1998. The pattern of inquiry. Pages 169–179 of: *The Essential Dewey, Volume 2: Ethics, Logic, Psychology*. Indiana University Press.
- Di Battista, G., Eades, P., Tamassia, R., and Tollis, I. G. 1999. *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice Hall.
- Diani, M. 1995. *Green Networks. A Structural Analysis of the Italian Environmental Movement*. Edinburgh University Press.
- Diani, M. 2002. Network analysis. Pages 173–200 of: *Methods of Social Movement Research*. University of Minnesota Press.
- Diani, M., and Bison, I. 2004. Organizations, coalitions, and movements. *Theory and Society*, **33**, 281–309.
- DiMaggio, P. J., and Powell, W. W. 1983. The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, **48**, 147–160.
- Dunbar, R. I. M. 2005. Why are good writers so rare? An evolutionary perspective on literature. *Journal of Cultural and Evolutionary Psychology*, **3**, 7–22.
- Dunne, J. A., Williams, R. J., and Martinez, N. D. 2002a. Food-web structure and network theory: The role of connectance and size. *Proceedings of the National Academy of Sciences U.S.A.*, **99**, 12917–12922.
- Dunne, J. A., Williams, R. J., and Martinez, N. D. 2002b. Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, **5**, 558–567.

- Dunne, J. A., Williams, R. J., and Martinez, N. D. 2004. Network structure and robustness of marine food webs. *Marine Ecology Progress Series*, **273**, 291–302.
- Easley, D., and Kleinberg, J. 2010. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press.
- Ebenman, B., and Jonsson, T. 2005. Using community viability analysis to identify fragile systems and keystone species. *Trends in Ecology and Evolution*, **20**, 568–575.
- Edmondson, A. 1999. Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, **44**, 350–383.
- EEGA. 2008. *Major Gas Pipelines of the Former Soviet Union and Capacity of Export Pipelines*. Tech. rept. East European Gas Analysis. www.eegas.com/fsu.htm.
- EIA. 2008. *Country Analysis Briefs*. Tech. rept. Energy Information Administration. www.eia.doe.gov/emeu/cabs.
- English, J. F. 2005. *The Economy of Prestige: Prizes, Awards, and the Circulation of Cultural Value*. Cambridge: Harvard University Press.
- Erdős, P., and Rényi, A. 1959. On random graphs. *Publicationes Mathematicae*, **6**, 290–297.
- Erdős, P., and Rényi, A. 1960. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, **5**, 17–61.
- Erikson, E., and Bearman, P. 2006. Foundations for global trade: the structure of English trade in the East Indies. *American Journal of Sociology*, **112**, 195–230.
- Estes, J. A., and Palmisano, J. F. 1974. Sea otters: their role in structuring nearshore communities. *Science*, **185**, 1058–1060.
- Etzkowitz, H., Kemelgor, C., Neuschatz, M., Uzzi, B., and Alonzo, J. 1994. The paradox of critical mass for women in science. *Science*, **266**, 51–54.
- EU. 2002. *EU Green Paper: Towards a European Strategy for the Security of Energy Supply*. Tech. rept. European Commission.
- Eurostat. 2008. *Energy - Yearly Statistics 2006, Edition 2008. Theme: Environment and energy Collection: Statistical books*. Tech. rept. European Union.
- Evans, J. A. 2008. Electronic publication and the narrowing of science and scholarship. *Science*, **321**, 395–399.
- Fafchamps, M., Goyal, S., and van der Leij, M. J. 2010. Matching and network effects. *Journal of the European Economic Association*, **8**, 203–231.
- Falconi, A. M., Guenfoud, K., Lazega, E., Lemercier, C., and Mounier, L. 2005. Le Contrôle social du monde des affaires: une étude institutionnelle. *L'Année Sociologique*, **55**, 451–484.
- Feick, L. F., and Price, L. L. 1987. The market maven: a diffuser of marketplace information. *The Journal of Marketing*, **51**, 83–97.
- Fernandez-Mateo, I. 2007. Who pays the price of brokerage? Transferring constraint through price setting in the staffing sector. *American Sociological Review*, **72**, 291–317.
- Festinger, L., and Schachter, S. 1950. *Social Pressures in Informal Groups: A Study of Human Factors in Housing*. Harper.
- Figyelő. 2002. Kétszázak klubja 2001. In: *Figyelő Top 200*.
- Fiore, S. M. 2008. Interdisciplinarity as teamwork how the science of teams can inform team science. *Small Group Research*, **39**, 251–277.
- Fitzgerald, F. S. 1993. The Crack Up. Page 122 of: Wilson, E. (ed), *The Crack Up*. New York: New Directions.

- Flanagin, A. J., Stohl, C., and Bimber, B. 2006. Modeling the structure of collective action. *Communication Monographs*, **73**, 29–54.
- Flemming, R. B. 1998. Contested terrains and regime politics: thinking about America's trial courts and institutional change. *Law & Social Inquiry*, **23**, 941–965.
- Fortunato, S. 2010. Community detection in graphs. *Physics Reports*, **486**, 75–174.
- Freeman, L. C., Borgatti, S. P., and White, D. R. 1991. Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks*, **13**, 141–154.
- Friedkin, N. E. 2004. Social cohesion. *Annual Review of Sociology*, **30**, 409–425.
- Fruchterman, T. M., and Reingold, E. M. 1991. Graph drawing by force-directed placement. *Software - Practice and Experience*, **21**, 1129–1164.
- Fulk, J., Flanagin, A., Kalman, M., Monge, P. R., and Ryan, T. 1996. Connective and communal public goods in interactive communication systems. *Communication Theory*, **61**, 60–87.
- Fulk, J., Heino, R., Flanagin, A., Monge, P., and Bar, F. 2004. A test of the individual action model for organizational information commons. *Organization Science*, **15**, 569–585.
- Galdwell, M. 2000. *The Tipping Point*. New York: Little Brown and Company.
- Gamson, W. 1961. A theory of coalition formation. *American Sociological Review*, **26**, 373–382.
- Gardiner, J. A. 2009. Defining corruption. Pages 25–40 of: Heidenheimer, A. J., and Johnston, M. (eds), *Political Corruption: Concepts and Contexts*, 5th edn. Transaction Publishers, New Brunswick, New Jersey.
- Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., Botstein, D., and Brown, P. O. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Molecular Biology of the Cell*, **11**, 4241–4257.
- George, E., Chattopadhyay, P., Sitkin, S. B., and Barden, J. 2006. Cognitive underpinnings of institutional persistence and change: a framing perspective. *Academy of Management Review*, **31**, 347–365.
- Gergely, G., Derényi, I., Farkas, I., and Vicsek, T. 2005. Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, **435**, 814–818.
- Gerlach, L. 1971. Movements of revolutionary change. Some structural characteristics. *American Behavioral Scientist*, **43**, 813–836.
- Gerlach, L. 2001. The structure of social movements: environmental activism and its opponents. Pages 289–310 of: *Networks and Netwars: The Future of Terror, Crime, and Militancy*. Rand.
- Gerstner, L. V. 2002. *Who Says Elephants Can't Dance?* Harper Business, New York.
- Giddens, A. 1994. *The Constitution of Society*. Polity Press.
- GIE. 2008a. *Storage Map - Information by point*. Tech. rept. Gas Infrastructure Europe. http://www.gie.eu/maps_data/availablecapacities.asp.
- GIE. 2008b. *The European Natural Gas Network (Capacities at cross-border points on the primary market)*. Tech. rept. Gas Infrastructure Europe. http://www.gie.eu/maps_data/capacity.asp.
- Girvan, M., and Newman, M. E. J. 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences U.S.A.*, **99**, 7821–7826.

- Glenn, J. K. 1999. Competing challengers and contested outcomes to state breakdown: the velvet revolution in Czechoslovakia. *Social Forces*, **78**, 187–211.
- Gohen, W., and Levinthal, D. 1994. A fortune favors the prepared firm. *Management Science*, **40**, 227–251.
- Goldenberg, J., Libai, B., and Muller, E. 2001. A complex systems look at the underlying process of word-of-mouth. *Marketing Letters*, **12**, 211–223.
- Gould, R. V. 1991. Multiple networks and mobilization in the Paris Commune. *American Sociological Review*, **56**, 716–729.
- Gould, R. V. 2003. *Collision of Wills: How Ambiguity about Social Rank Breeds Conflict*. University of Chicago Press.
- Granovetter, M. 1973. The strength of weak ties. *American Journal of Sociology*, **78**, 1360–1380.
- Granovetter, M. 1978. Threshold models of collective behavior. *American Journal of Sociology*, **83**, 1420–1443.
- Granovetter, M. 2005. Business groups and social organization. Pages 429–450 of: *Handbook of Economic Sociology*. Princeton University Press.
- Granovetter, M. 2007. The social construction of corruption. In: Nee, V., and Swedberg, R. (eds), *On Capitalism*. Stanford University Press.
- Green, S., and Green, K. 1996. *Broadway Musicals Show by Show*. 5th edn. Milwaukee: Hal Leonard.
- Gruhl, D., Liben-Nowell, D., Guha, R., and Tomkins, A. 2004. Information diffusion through blogspace. *SIGKDD Explorations Newsletter*, **6**, 43–52.
- Guimerá, R., Uzzi, B., Spiro, J., and Amaral, L. A. N. 2005. Team assembly mechanisms determine collaboration network structure and team performance. *Science*, **308**, 697–702.
- Gulati, R., and Gargiulo, M. 1999. Where do interorganizational networks come from? *American Journal of Sociology*, **104**, 1439–1493.
- Hackman, J. R., and Katz, N. 2010. Group behavior and performance. Pages 1208–1251 of: Fiske, S. T., Gilbert, D. T., and Lindzey, G. (eds), *Handbook of Social Psychology*, 5th edn. New York: Wiley.
- Halbeisen, R. E., and Gerber, A. P. 2009. Stress-dependent coordination of transcriptome and translatome in yeast. *PLoS Biology*, **7**, e105.
- Hall, B. H., Jaffe, A., and Trajtenberg, M. 2001. The NBER patent citations data file: lessons, insights and methodological tools. *National Bureau of Economic Research, Working Paper No. 8498*.
- Hall, B. H., Jaffe, A. B., and Trajtenberg, M. 2005. Market value and patent citations. *RAND Journal of Economics*, **36**, 16–38.
- Hannan, M. T., and Freeman, J. H. 1977. The population ecology of organizations. *American Journal of Sociology*, **82**, 929–964.
- Hargreaves Heap, S., Hollis, M., Lyons, B., Sugden, R., and Weale, A. 1992. *The Theory of Choice. A Critical Guide*. Blackwell.
- Harrison, D. A., Price, K. H., and Bell, M. P. 1998. Beyond relational demography: time and the effects of surface- and deep-level diversity on work group cohesion. *Academy of Management Journal*, **41**, 96–107.
- Hartwell, L. H., Hopfield, J. J., Leibler, S., and Murray, A. W. 1999. From molecular to modular cell biology. *Nature*, **402**, 47–52.
- Hayek, F. A. 1945. The use of knowledge in society. *American Economic Review*, **35**, 519–530.

- Heckman, J. J. 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement*, **5**, 475–492.
- Heckman, J. J. 1979. Sample selection bias as a specification error. *Econometrica*, **47**, 153–161.
- Hellman, J., and Kaufmann, D. 2001. Confronting the challenge of state capture in transition economies. *Finance & Development*, **38**, 31–35.
- Hellman, J., Jones, G., and Kaufmann, D. 2000. Seize the state, seize the day: state capture, corruption and influence in transition. *World Bank Policy Research, Working Paper No. 2444*.
- Herman, I., Melançon, G., and Scott Marshall, M. 2000. Graph visualization and navigation in information visualisation. *IEEE Transactions on Visualization and Computer Graphics*, **6**, 24–43.
- Hethcote, H. W. 1989. Three basics epidemiological. *Biomathematics*, **18**, 119–144.
- Hethcote, Herbert W. 2000. The Mathematics of Infectious Diseases. *SIAM Review*, **42**(4), 599–653.
- Heydebrand, W., and Seron, C. 1990. *Rationalizing justice: The Political Economy of Federal District Courts*. SUNY Press.
- Higashi, M., Burns, T. P., and Patten, B. C. 1989. Food network unfolding: an extension of trophic dynamics for application to natural ecosystems. *Journal of Theoretical Biology*, **140**, 243–261.
- Hinds, P. J., Carley, K. M., Krackhardt, D., and Wholey, D. 2000. Choosing work group members: balancing similarity, competence, and familiarity. *Organizational Behavior and Human Decision Processes*, **81**, 226–251.
- Holling, C. S. 1986. The resilience of terrestrial ecosystems: local surprise and global change. Pages 292–317 of: *Sustainable Development of the Biosphere*. Cambridge University Press.
- Hollingshead, A. B. 1997. Retrieval processes in transactive memory systems. *Personality and Social Psychology*, **74**, 659–671.
- Hollingshead, A. B. 1998. Communication, learning, and retrieval in transactive memory systems. *Journal of Experimental Social Psychology*, **34**, 423–442.
- Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell*, **95**, 717–728.
- Huang, M., Huang, Y., Ognyanova, K., Margolin, D., Shen, C., and Contractor, N. S. 2010. The effects of diversity and repeat collaboration on team performance in distributed nanoscientist teams. In: *Academy of Management Annual Conference*.
- Huffaker, D. 2010. Dimensions of leadership and social influence in online communities. *Human Communication Research*, **36**, 593–617.
- Ibolya, T. 2010. A hűtlen kezelés bizonyítása. *Rendészeti Szemle*, **10**, 82–94.
- IEA. 1998. *Caspian Oil and Gas*. Tech. rept. International Energy Agency, Paris, France.
- IEA. 2005. *Natural Gas Information 2005*. Tech. rept. IEA/OECD Paris.
- IEA. 2007. *World Energy Outlook 2007*. Tech. rept. OECD/IEA, Paris.
- IEA. International Energy Agency, Paris, France. *Energy balances of OECD countries 1960–2005*. Tech. rept. 2007.
- Jackson, B. A. 2006. Groups, networks, or movements: a command-and-control-driven approach to classifying terrorist organizations and its application to Al Qaeda. *Studies in Conflict and Terrorism*, **29**, 241–262.

- Jacobs, J. 1969. *The Economy of Cities*. Random House.
- Jacobs, J. 1983. The Economy of Regions. In: *Third Annual E. F. Schumacher Lecture*, Mount Holyoke College, South Hadley.
- Jacobs, J. 1992. *Systems of Survival: A Dialogue on the Moral Foundations of Commerce and Politics*. Random House.
- Jehn, K. A., Northcraft, G. B., and Neale, M. A. 1999. Why differences make a difference: a field study of diversity, conflict, and performance in workgroups. *Administrative Science Quarterly*, **44**, 741–763.
- Jenkins, M. 2003. Prospects for biodiversity. *Science*, **302**, 1175–1177.
- Jeppesen, E., Kristensen, P., Jensen, J. P., Søndergaard, M., Mortensen, E., and Lauridsen, T. 1991. Recovery resilience following a reduction in external phosphorous loading of shallow, eutrophic Danish lakes: duration, regulating factors and methods for overcoming resilience. *Memorie Istituto Italiano Idrobiologia*, **48**, 127–148.
- Johnston, H., and Snow, D. A. 1998. Subcultures and the emergence of the Estonian nationalist opposition 1945–1990. *Sociological Perspectives*, **41**, 473–497.
- Jones, B. F. 2005. The burden of knowledge and the ‘death of the renaissance man’: is innovation getting harder? *National Bureau of Economic Research, Working Paper No. 11360*.
- Jones, C., Hesterly, W. S., and Borgatti, S. P. 1997. A general theory of network governance: exchange conditions and social mechanisms. *Academy of Management Review*, **22**, 911–945.
- Jordán, F. 2009. Keystone species and food webs. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, **364**, 1733–1741.
- Jordán, F., Liu, W., and Wyatt, T. 2005. Topological constraints on the dynamics of wasp-waist ecosystems. *Journal of Marine Systems*, **57**, 250–263.
- Kadushin, C., Ryan, D., Brodsky, A., and Saxe, L. 2005. Why it is so difficult to form effective community coalitions. *City and Community*, **4**, 255–275.
- Karl, T. L. 1997. *The Paradox of Plenty: Oil Booms and Petro-States*. University of California Press.
- Katona, G., and Mueller, E. 1955. A study of purchase decisions. Pages 30–87 of: Clark, L. H. (ed), *Consumer Behavior: The Dynamics of Consumer Reaction*. New York University Press.
- Katz, B. 2010. City centered. *Time*, **21**, 1–4.
- Katz, E. 1996. Diffusion research at Columbia. Pages 61–70 of: Dennis, E. E., and Wartella, E. (eds), *American Communication Research: The Remembered History*. Lawrence Erlbaum Associates.
- Katz, E., and Lazarsfeld, P. F. 1955. *Personal Influence: The Part Played by People in the Flow of Mass Communications*. New York: Free Press.
- Katzenback, J. R., and Smith, D. K. 1993. *The Wisdom of Teams*. New York: Harper Business.
- Kaufmann, D., and Kraay, A. 2007. On measuring governance: framing issues for debate. In: *Roundtable on Measuring Governance Hosted by the World Bank Institute and the Development Economics Vice-Presidency of The World Bank*.
- Kaufmann, D., Kraay, A., and Mastruzzi, M. 2009. Governance matters VIII: aggregate and individual governance indicators, 1996–2008. *World Bank Policy Research, Working Paper No. 4978*.
- Kempe, D., Kleinberg, J., and Tardos, É. 2003. Maximizing the spread of influence through a social network. Pages 137–146 of: *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

- Kenis, P., and Knoke, D. 2002. How organizational field networks shape interorganizational tie-formation rates. *Academy of Management Review*, 27, 275–293.
- Khanna, T., and Rivkin, J. W. 2001. Estimating the Performance Effects of Business Groups in Emerging Markets. *Strategic Management Journal*, 22, 45–74.
- Klein, N. 2008. *The Shock Doctrine: The Rise of Disaster Capitalism*. Knopf.
- Klitgaard, R. 1991. *Controlling Corruption*. University of California Press.
- Knoke, D. 1990. *Political Networks*. Cambridge University Press.
- Kogut, B., and Zander, U. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science*, 3, 383–397.
- Kondoh, M. 2003. Foraging adaptation and the relationship between food-web complexity and stability. *Science*, 299, 1388–1391.
- Kontopoulos, K. 1993. *The Logics of Social Structure*. Cambridge University Press.
- Korcsmáros, T., Kovács, I. A., Szalay, M. S., and Csermely, P. 2007. Molecular chaperones: the modular evolution of cellular networks. *Journal of Biosciences*, 32, 441–446.
- Kovács, I. A., Palotai, R., Szalay, M. S., and Csermely, P. 2010. Community landscapes: a novel, integrative approach for the determination of overlapping network modules. *PLoS ONE*, 7, e12528.
- Krackhardt, D. 1987. Cognitive social structures. *Social Networks*, 9, 109–134.
- Krackhardt, D. 1990. Assessing the political landscape: structure, cognition, and power in organizations. *Administrative Science Quarterly*, 35, 342–369.
- Krempel, L., and Plumper, T. 2003. Exploring the dynamics of international trade by combining the comparative advantages of multivariate statistics and network visualization. *Journal of Social Structure*, 4, 1–22.
- Kruyt, B., van Vuuren, D. P., de Vries, H. J. M., and Groenenberg, H. 2009. Indicators for energy security. *Energy Policy*, 37, 2166–2181.
- Kuhn, T. S. 1970. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Kuran, T. 1987. Preference falsification, policy continuity and collective conservatism. *The Economic Journal*, 97, 642–665.
- Kynge, J. 2007. *China Shakes the World: a Titan's Rise and Troubled Future – and the Challenge for America*. New York: Houghton Mifflin.
- Lambsdorff, J. G. 2007. *The Institutional Economics of Corruption and Reform Theory, Evidence and Policy*. Cambridge University Press.
- Lammers, J. C., and Barbour, J. B. 2006. An institutional theory of organizational communication. *Communication Theory*, 16, 356–377.
- Larson, J. R., Christensen, C., Abbott, A. S., and Franz, T. M. 1996. Diagnosing groups: charting the flow of information in medical decision-making teams. *Journal of Personality and Social Psychology*, 71, 315–330.
- Lasswell, H. D. 1930. *Encyclopedia of Social Sciences*. Vol. II. New York, Macmillan.
- Laumann, E., and Knoke, D. 1987. *The Organizational State*. Wisconsin University Press.
- Lawless, W. F., Rifkin, S., Sofge, D., Hobbs, S. H., Angjellari-Dajci, F., and Chaudron, L. 2010. Conservation of information: reverse engineering dark social systems. *Structure and Dynamics: eJournal of Anthropological and Related Sciences*, 4.
- Lawrence, P., and Lorsch, J. 1967. Differentiation and integration in complex organizations. *Administrative Science Quarterly*, 12, 1–30.

- Lazarsfeld, P. F., Berelson, B., and Gaudet, H. 1948. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Election*. New York: Columbia University Press.
- Lazega, E. 1992. *Micropolitics of Knowledge: Communication and Indirect Control in Workgroups*. Aldine de Gruyter.
- Lazega, E. 2003. *Networks in Legal Organizations: On the Protection of Public Interest in Joint Regulation of Markets*. Wiarda Institute Publications.
- Lazega, E. 2009. Theory of cooperation among competitors: a neo-structural approach. *Sociologica*, **1**, 1–34.
- Lazega, E., and Mounier, L. 2002. Interdependent entrepreneurs and the social discipline of their cooperation: structural economic sociology for a society of organizations. In: *Conventions and Structures in Economic Organization: Markets, Networks, and Hierarchies*. Edward Elgar Publishers.
- Lazega, E., and Mounier, L. 2003. Interlocking judges: on joint external and self-governance of markets. Pages 267–296 of: *Research in the Sociology of Organizations*. Emerald Group Publishing Limited.
- Lazer, D., and Friedman, A. 2007. The network structure of exploration and exploitation. *Administrative Science Quarterly*, **52**, 667–694.
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., and Van Alstyne, M. 2009. Social science: computational social science. *Science*, **323**, 721–723.
- Lehner, J. 2010. A physicist solves the city. *NYT Magazine*.
- Lengauer, T., and Tarjan, R. E. 1979. A fast algorithm for finding dominators in a flowgraph. *ACM Transactions on Programming Languages and Systems*, **1**, 121–141.
- Lester, R. K., and Piore, M. J. 2004. *Innovation: The Missing Dimension*. Harvard University Press.
- Levins, R. 1974. The qualitative analysis of partially specified systems. *Annals of New York Academy of Science*, **231**, 123–138.
- Levins, R. 1995. Preparing for uncertainty. *Ecosystem Health*, **1**, 47–57.
- Levy, S. F., and Siegal, M. L. 2008. Network hubs buffer environmental variation in *Saccharomyces cerevisiae*. *PLoS Biology*, **6**, e264.
- Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E., and Åberg, Y. 2001. The web of human sexual contacts. *Nature*, **411**, 907–908.
- Lincoln, J. R., Gerlach, M. L., and Ahmadjian, C. L. 1996. Keiretsu networks and corporate performance in Japan. *American Sociological Review*, **61**, 67–88.
- Lochner, S., and Bothe, D. 2009. The development of natural gas supply costs to Europe, the United States and Japan in a globalizing gas market-Model-based analysis until 2030. *Energy Policy*, **37**, 1518–1528.
- Lofland, J. 1996. *Social Movement Organizations*. Aldine de Gruyter.
- Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J. P., Hector, A., Hooper, D. U., Huston, M. A., Raffaelli, D., Schmid, B., Tilman, D., and Wardle, D. A. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science*, **294**, 804–808.
- Luciani, G. 2004. *Security of Supply for Natural Gas Markets: What is it and what is it not?* Tech. rept. 119.04. FEEM. <http://www.feem.it/userfiles/attach/Publication/NDL2004/NDL2004-119.pdf>.
- MacArthur, R. H. 1955. Fluctuation of animal populations and a measure of community stability. *Ecology*, **36**, 533–536.

- Macaulay, S. 1963. Non-contractual relations in business: A preliminary study. *American Sociological Review*, **28**, 55–67.
- Mahajan, V., and Peterson, R. A. 1985. *Models for Innovation Diffusion (Quantitative Applications in the Social Sciences)*. Newbury Park, CA: Sage.
- March, J., and Simon, H. 1958. *Organizations*. Wiley.
- Marsden, P. V. 2005. Recent developments in network measurement. Pages 8–30 of: *Models and Methods in Social Network Analysis*. Cambridge University Press.
- Martinez, N. D., Hawkins, B. A., Dawah, H. A., and Feifarek, B. P. 1999. Effect of sampling effort on characterization of food-web structure. *Ecology*, **80**, 1044–1055.
- Marwell, G., Oliver, P. E., and Prahl, R. 1988. Social networks and collective action – a theory of the critical mass. *American Journal of Sociology*, **94**, 502–534.
- Mavrakakis, D., Thomaidis, F., and Ntroukas, I. 2006. An assessment of the natural gas supply potential of the south energy corridor from the Caspian Region to the EU. *Energy Policy*, **34**, 1671–1680.
- May, R. M. 1973. *Complexity and Stability in Model Ecosystems*. Princeton University Press.
- May, R. M. 2006. Network structure and the biology of populations. *Trends in Ecology and Evolution*, **21**, 394–399.
- Mayntz, R. 2003. New challenges to governance theory. Pages 27–40 of: *Governance as Social and Political Communication*. Manchester University Press.
- McAdam, D. 2003. Beyond structural analysis: toward a more dynamic understanding of social movements. Pages 281–298 of: *Social Movements and Networks*. Oxford University Press.
- McCann, K. S. 2000. The diversity-stability debate. *Nature*, **405**, 228–233.
- McDonald, K. 2002. From solidarity to fluidarity: social movements beyond ‘collective identity’. The case of globalization conflicts. *Social Movement Studies*, **1**, 109–128.
- McIntosh, W. V., and Cates, C. L. 1997. *Judicial Entrepreneurship: The Role of the Judge in the Marketplace of Ideas*. Greenwood Press.
- McPherson, M., and Smith-Lovin, L. 2002. Cohesion and membership duration: linking groups, relations and individuals in an ecology of affiliation. *Advances in Group Processes*, **19**, 1–36.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. 2001. Birds of a feather: homophily in social networks. *Annual Review of Sociology*, **27**, 415–444.
- McVoy, E. C. 1940. Patterns of diffusion in the United States. *American Sociological Review*, **5**, 219–227.
- Melián, C., and Bascompte, J. 2002. Complex networks: two ways to be robust? *Ecology Letters*, **5**, 705–708.
- Melucci, A. 1996. *Challenging Codes*. Cambridge University Press.
- Merton, R. K. 1968. The Matthew Effect in Science. *Science*, **159**, 56–63.
- Merton, R. K. 1973a. *The Sociology of Science*. Chicago: University of Chicago Press. Pages 545–550.
- Merton, R. K. 1973b. *The Sociology of Science*. Chicago: University of Chicago Press.
- Metcalfe, B., and Linstead, A. 2003. Gendering teamwork: re-writing the feminine. *Gender, Work & Organization*, **10**, 94–119.
- Meyer, D. S., and Corrigall-Brown, C. 2005. Coalitions and political context: U.S. movements against war in Iraq. *Mobilization*, **10**, 327–344.

- Meyer, J. W., and Rowan, B. 1991. Institutionalized organizations: formal structure as myth and ceremony. Pages 41–62 of: Powell, W. W., and DiMaggio, P. J. (eds), *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.
- Mihalik, Á., Palotai, R., and Csermely, P. 2008. The effect of stress on the modular structure of the yeast protein-protein interaction network (in Hungarian). *Biochemistry*, **32**, 67.
- Milgram, S. 1967. The small-world problem. *Psychology Today*, **2**, 60–67.
- Mizruchi, M. S., and Galaskiewicz, J. 1993. Networks of interorganizational relations. *Sociological Methods and Research*, **22**, 46–70.
- Mizruchi, M. S., and Stearns, L. B. 1988. A Longitudinal study of the formation of interlocking directorates. *Administrative Science Quarterly*, **33**, 194–210.
- Modelski, G. T. 2000a. What causes K-waves. *Technological Forecasting and Social Change*, **68**, 75–80.
- Modelski, G. T. 2000b. World system evolution. In: Denemark, R., Friedman, J., Gills, B., and G., Modelski (eds), *World-System Evolution: The Social Science of Long-Term Change*. Routledge.
- Modelski, G. T., and Thompson, W. R. 1996. *Leading Sectors and World Powers: The Coevolution of Global Economics and Politics*. Columbia University Press.
- Monge, P. R., and Contractor, N. S. 2003. *Theories of Communication Networks*. Oxford University Press.
- Montoya, J. M., and Solé, R. V. 2002. Small world patterns in food webs. *Journal of Theoretical Biology*, **214**, 405–491.
- Montoya, J. M., Pimm, S. L., and Solé, R. V. 2006. Ecological networks and their fragility. *Nature*, **214**, 259–493.
- Moody, J. A., and Douglas, R. W. 2003. Structural cohesion and embeddedness: a hierarchical concept of social groups. *American Sociological Review*, **68**, 103–127.
- Moreno, J., and Jennings, H. 1937. Statistics of social configurations. *Sociometry*, **1**, 342–374.
- Navarrete, S. A., and Menge, B. A. 1996. Keystone predation and interaction strength: interactive effects of predators on their main prey. *Ecological Monographs*, **66**, 409–429.
- Nelson, R. R., and Winter, S. G. 1983. *An Evolutionary Theory of Economic Change*. Boston: Bellknap.
- Newman, M. E. J. 2001. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences U.S.A.*, **98**, 404–409.
- Newman, M. E. J. 2003. The Structure and function of complex networks. *SIAM Review*, **45**, 167–256.
- Newman, M. E. J. 2004. Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences U.S.A.*, **101**, 5200–5205.
- Newman, M. E. J., and Girvan, M. 2004. Finding and evaluating community structure in networks. *Physical Review E*, **69**.
- Newman, M. E. J., and Watts, D. J. 1999. Scaling and percolation in the small-world network model. *Physical Review E*, **60**, 7332–7342.
- Newman, M. E. J., Watts, D. J., and Strogatz, S. H. 2002. Random Graph Models of Social Networks. *Proceedings of the National Academy of Sciences U.S.A.*, **99**, 2566–2572.
- Noack, A. 2007. Energy models for graph clustering. *Journal of Graph Algorithms and Applications*, **11**, 453–480.

- Noack, A., and Rotta, R. 2009. Multi-level algorithms for modularity clustering. Pages 257–268 of: *SEA '09: Proceedings of the 8th International Symposium on Experimental Algorithms*. Springer-Verlag.
- Nonaka, I., and Takeuchi, H. 1995. *The Knowledge-Creating Company*. Oxford University Press, New York.
- Norris, P. 2003. *Democratic Phoenix*. Cambridge University Press.
- Nye, J. S. 2008. Corruption and political development: a cost benefit analysis. Pages 281–300 of: Heidenheimer, A. J., and Johnston, M. (eds), *Political Corruption*. Transaction Publishers.
- Nyíró, A., and Szakadát, I. 1993. *Politika Interaktív CD-rom*. Budapest: Aula.
- Oka, R. C., and Chapurukha, M. K. 2008. Archaeology of trading systems, part 1: towards a new trade synthesis. *Journal of Archaeological Research*, **16**, 339–395.
- Olcott, M. B. 2004. *International Gas Trade in Central Asia: Turkmenistan, Iran, Russia and Afghanistan*. Tech. rept. Working Paper #28. Stanford Institute for International Studies. http://www.rice.edu/energy/publications/docs/GAS_InternationalGasTradeinCentralAsia.pdf.
- Ouchi, W. G. 1980. Markets, bureaucracies and clans. *Administrative Science Quarterly*, **25**, 129–141.
- Paine, R. T. 1969. A note on trophic complexity and community stability. *American Naturalist*, **103**, 91–501.
- Paine, R. T. 1980. Food webs, linkage interaction strength, and community infrastructure. *Journal of Animal Ecology*, **49**, 667–685.
- Palla, G., Derényi, I., Farkas, I., and Vicsek, T. 2005. Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, **435**, 814–818.
- Palla, P., Barabási, A.-L., and Vicsek, T. 2007. Quantifying social group evolution. *Nature*, **466**, 664–667.
- Palotai, R., Szalay, M. S., and Csermely, P. 2008. Chaperones as integrators of cellular networks: changes of cellular integrity in stress and diseases. *IUBMB Life*, **60**, 10–18.
- Pascual, M., and Dunne, J. A. 2006. *Ecological Networks: Linking Structure to Dynamics in Food Webs*. Oxford University Press.
- Pavan, E. in press. *Ties (In)Formation and Communication. Online and Offline Collaboration in the Internet Governance Forum Space*. Lanham: Rowman & Littlefield.
- Pavlopoulos, G. A., Wegener, A. L., and Schneider, R. 2008. A survey of visualization tools for biological network analysis. *BioData Mining*, **1**, 12.
- Pelletier, C., and Wortmann, J. C. 2009. A risk analysis for gas transport network planning expansion under regulatory uncertainty in Western Europe. *Energy Policy*, **37**, 721–732.
- Peuhkuri, T., and Jokinen, P. 1999. The role of knowledge and spatial contexts in biodiversity policies: a sociological perspective. *Biodiversity and Conservation*, **8**, 133–147.
- Pfeffer, J., and Salancik, J. 1978. *The External Control of Organizations*. New York: Harper and Row.
- Piattoni, S. 2010. *The Theory of Multi-level Governance*. Oxford University Press.
- Pimm, S. L., Lawton, J. H., and Cohen, J. E. 1991. Food web patterns and their consequences. *Nature*, **350**, 669–674.

- Pimm, S. L., Raven, R., Peterson, A., Ekericioglu, C. H., and Ehrlich, P. R. 2006. Human impacts on the rates of recent, present, and future bird extinctions. *Proceedings of the National Academy of Sciences U.S.A.*, **103**, 10941–10946.
- Pincus, S. C. A. 2005. *England's Glorious Revolution 1688-89: A Brief History with Documents*. St. Martin's Press.
- Pirani, S., Stern, J., and Yafimava, K. 2009. *The Russo-Ukrainian gas dispute of January 2009: a comprehensive assessment*. Tech. rept. NG 27. Oxford Institute for Energy Studies.
- Pizzorno, A. 1978. Political exchange and collective identity in industrial conflict. Pages 277–298 of: *The Resurgence of Class Conflict in Western Europe*. Holmes and Meier.
- Pizzorno, A. 2008. *Rationality and Recognition*. Cambridge University Press.
- Podolny, J. M., and Page, K. L. 1998. Network Forms of Organization. *Annual Review of Sociology*, **24**, 57–76.
- Poole, M. S., and DeSanctis, G. 1992. Microlevel structuration in computer-supported group decision-making. *Communication Research*, **19**, 5–49.
- Poole, M. S., McPhee, R. D., and Seibold, D. R. 1982. A comparison of normative and interactional explanations of group decision-making: social decision schemes versus valence distributions. *Communication Monographs*, **49**, 1–19.
- Poole, M. S., Seibold, D. R., and McPhee, R. D. 1996. The structuration of group decisions. Pages 114–146 of: Hirokawa, R., and Poole, M. (eds), *Communication and Group Decision Making*. Sage.
- Powell, W. 1990. Neither markets nor hierarchy: network forms of organization. *Research in Organizational Behavior*, **12**, 295–336.
- Powell, W. W., White, D. R., Koput, K. W., and Jason, O.-S. 2005. Network dynamics and field evolution: the growth of interorganizational collaboration in the life sciences. *The American Journal of Sociology*, **110**, 1132–1205.
- Proulx, S. R., Promislow, D. E., and Phillips, P. C. 2005. Network thinking in ecology and evolution. *Trends in Ecology and Evolution*, **20**, 345–353.
- Provan, K. G., Fish, A., and Sydow, J. 2007. Interorganizational networks at the network level: A review of the empirical literature on whole networks. *Journal of Management*, **33**, 479–516.
- Puccia, C. J., and Levins, R. 1985. *Qualitative Modelling of Complex Systems: An Introduction to Loop Analysis and Time Averaging*. Harvard University Press.
- Puska, P., and Uutela, A. 2000. Community intervention in cardiovascular health promotion: North Karelia, 1972-1999. Pages 73–96 of: Schneiderman, N., Speers, M. A., Silva, J. M., Tomes, H., and Gentry, J. H. (eds), *Integrating Behavioral and Social Sciences With Public Health*. Baltimore: United Book Press, American Psychological Association.
- R Development Core Team. 2010. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Ramasco, J. J., Dorogovtsev, S. N., and Pastor-Satorras, R. 2004. Self-organization of collaboration networks. *Physical Review E*, **70**, 036106.
- Rasmusen, E. 1989. *Games and Information: An Introduction to Game Theory*. Blackwell Publishing Ltd.
- Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., and Barabási, A.-L. 2002. Hierarchical organization of modularity in metabolic networks. *Science*, **297**, 1551–1555.

- Reagans, R., and Zuckerman, E. W. 2001. Networks, diversity, and productivity: the social capital of corporate R&D teams. *Organization Science*, **12**, 502–517.
- Reichardt, J., and White, D. R. 2007. Role models for complex networks. *European Physical Journal B*, **60**, 217–224.
- Reymond, M. 2007. European key issues concerning natural gas: Dependence and vulnerability. *Energy Policy*, **35**, 4169–4176.
- Robins, G., Pattison, P., and Wang, P. 2006. Closure, connectivity and degrees: new specifications for Exponential Random Graph (p^*) Models for directed social networks. *Unpublished manuscript*. University of Melbourne.
- Robins, G., Pattison, P., and Wang, P. 2007. An introduction to Exponential Random Graph (p^*) Models for social networks. *Social Networks*, **29**, 173–191.
- Rogers, E. 2003. *Diffusion of Innovations*. Free Press, New York.
- Rootes, C. 2003. *Environmental Protest in Western Europe*. Oxford University Press.
- Rose-Ackerman, S. 1978. *Corruption A Study in Political Economy*. Academic Press.
- Rose-Ackerman, S. 1999. *Corruption and Government Causes, Consequences, and Reform*. Cambridge University Press.
- Rose-Ackerman, S. 2006. *International Handbook on the Economics of Corruption*. Edward Elgar.
- Ruef, M., Aldrich, H. E., and Carter, N. M. 2003. The structure of founding teams: homophily, strong ties and isolation among U.S. entrepreneurs. *American Sociological Review*, **68**, 195–222.
- Rutherford, S. L., and Lindquist, S. 1998. Hsp90 as a capacitor for morphological evolution. *Nature*, **396**, 336–342.
- Ryan, B., and Gross, N. C. 1943. The diffusion of hybrid seed corn in two Iowa communities. *Rural Sociology*, **8**, 15–24.
- Saavedra, S., Reed-Tsochas, F., and Uzzi, B. 2008. Asymmetric disassembly and robustness in declining networks. *Proceedings of the National Academy of Sciences U.S.A.*, **105**, 16466–16471.
- Sampat, B. 2005. *Determinants of Patent Quality: An Empirical Analysis*. New York: Columbia University.
- Sander, L. M., Warren, C. P., Sokolov, I. M., Simon, C., and Koopman, J. 2002. Percolation on heterogeneous networks as a model for epidemics. *Mathematical Biosciences*, **180**, 293–305.
- Sartori, G. 1970. Concept misformation in comparative politics. *American Political Science Review*, **64**, 1033–1052.
- Sassen, S. 1991. *The Global City: New York, London, Tokyo*. Princeton University Press.
- Sassen, S. 2000. *Cities in a World Economy*. Pine Forge Press.
- Sassen, S. 2009. Bridging the ecologies of cities and of nature. In: *The New Urban Question – Urbanism Beyond Neo-Liberalism. The 4th International Conference of the International Forum on Urbanism (IFoU)*.
- Schaeffer, S. E. 2007. Graph Clustering. *Computer Science Review*, **1**, 27–64.
- Schelling, T. C. 1978. *Micromotives and Macrobehavior*. New York: W. W. Norton.
- Schilling, M. A., and Steensma, H. K. 2001. The use of modular organizational forms: an industry level analysis. *Academy of Management Journal*, **44**, 1149–1168.
- Schumpeter, J. 1942. *Capitalism, Socialism and Democracy*. Harper and Brothers.
- Schumpeter, J. A. 1934. *The Theory of Economic Development*. Cambridge, Mass: Harvard University Press.

- Schwab, A., and Miner, A. S. 2008. Learning in hybrid-project systems: The effects of project performance on repeated collaboration. *Academy of Management Journal*, **51**, 1117–1149.
- Schweitzer, F., Sornette, D., Vespignani, A., Fagiolo, G., Vega-Redondo, F., and White, D. R. 2009a. Economic networks: the new challenges. *Science*, **325**, 422–425.
- Schweitzer, F., Sornette, D., Fagiolo, G., Vega-Redondo, F., and White, D. R. 2009b. Economic networks: what do we know and what do we need to know? *Advances in Complex Systems*, **12**, 407–422.
- Scotti, M. 2008. Development Capacity. Pages 911–920 of: Jørgensen, S. E., and Fath, B. D. (eds), *Ecological Indicators. Vol. [2] of Encyclopedia of Ecology*. Elsevier, Oxford.
- Seibold, D. R., and Meyers, R. A. 2007. Group argument: a structuration perspective and research program. *Small Group Research*, **38**, 312–336.
- Sewell, W. H. J. 1992. A theory of structure: duality, agency and transformation. *American Journal of Sociology*, **98**, 1–29.
- Shalizi, C. 2011. Do city economies scale with population? *Carnegie Mellon University*, **12**.
- Shawe-Taylor, J., and Cristianini, N. 2004. *Kernel methods for pattern analysis*. Cambridge University Press.
- Shea, S., and Basch, C. E. 1990. A review of five major community-based cardiovascular disease prevention programs. Part I: rationale, design, and theoretical framework. *American Journal of Health Promotion*, **4**, 203–213.
- Shirkani, N. 2008. *Egyptian gas flows to Israel*. Tech. rept. Upstream Online. www.upstreamonline.com/live/article150348.ece.
- Sőti, C., Sreedhar, A. S., and Csermely, P. 2003. Apoptosis, necrosis and cellular senescence: chaperone occupancy as a potential switch. *Aging Cell*, **2**, 39–45.
- Simas, R. 1988. *The Musicals No One Came to See: A Guidebook to Four Decades of Musical-Comedy Casualties on Broadway, Off Broadway and in Out-Of-Town Tryout, 1943-1983*. New York: Garland.
- Simmel, G. 1898. The persistence of social groups. *American Journal of Sociology*, **3**, 662–698.
- Simmel, G. 1964. *Conflict and the Web of Group Affiliations*. New York: Free Press.
- Simon, H. A. 1959. Theories of decision-making in economics and behavioral science. *American Economic Review*, **3**, 253–283.
- Smith, D. A., and White, D. R. 1992. Structure and dynamics of the global economy: network analysis of international trade 1965-1980. *Social Forces*, **70**, 857–894.
- Smith, J. 2005a. Building bridges or building walls? Explaining regionalization among transnational social movement organizations. *Mobilization*, **10**, 251–270.
- Smith, J. 2005b. Globalization and transnational social movement organizations. Pages 226–248 of: *Social Movements and Organizations*. Cambridge University Press.
- Solomonoff, R., and Rapoport, A. 1951. Connectivity of random nets. *Bulletin of Mathematical Biology*, **13**, 107–117.
- Somers, M. R. 1994. The narrative constitution of identity: a relational and network approach. *Theory and Society*, **23**, 605–649.
- Sørensen, E., and Torfing, J. 2007. *Theories of Democratic Network Governance*. Edward Elgar.

- Sorenson, O., Rivkin, J. W., and Fleming, L. 2006. Complexity, networks and knowledge flow. *Research Policy*, **35**, 994–1017.
- Spicer, A., McDermott, G. A., and Kogut, B. 2000. Entrepreneurship and privatization in Central Europe: the tenuous balance between destruction and creation. *Academy of Management Review*, **25**, 630–649.
- Spufford, P. 2002. *Power and Profit: The Merchant in Medieval Europe*. Thames and Hudson.
- Spufford, P. 2006. From Antwerp to London: the decline of financial centres in Europe. *De Economist*, **154**, 143–175.
- Stark, C., Breitzkreutz, B. J., Chatr-Aryamontri, A., Boucher, L., Oughtred, R., Livstone, M. S., Nixon, J., van Auken, K., Wang, X., Shi, X., Reguly, T., Rust, J. M., Winter, A., Dolinski, K., and Tyers, M. 2011. The BioGRID Interaction Database: 2011 update. *Nucleic Acids Research*, **39**, D698–D704.
- Stark, D. 1996. Recombinant property in East European capitalism. *American Journal of Sociology*, **101**, 993–1027.
- Stark, D. 2009. *The Sense of Dissonance: Accounts of Worth in Economic Life*. Princeton University Press.
- Stark, D., and Bruszt, L. 1998. *Postsocialist Pathways: Transforming Politics and Property in East Central Europe*. Cambridge University Press.
- Stark, D., and Vedres, B. 2006. Social times of network spaces: network sequences and foreign investment in Hungary. *American Journal of Sociology*, **111**, 1367–1412.
- Stasser, G., Stewart, D. D., and Wittenbaum, G. M. 1995. Expert roles and information exchange during discussion: the importance of knowing who knows what. *Journal of Experimental Social Psychology*, **31**, 244–265.
- Stern, J. 2002. *Security of European natural gas supplies: the impact of import dependence and liberalization*. Tech. rept. 36 pp. Royal Institute of International Affairs, London. www.isn.ethz.ch/isn/Digital-Library/Publications/Detail/?id=99313&lng=en.
- Stern, J. 2006. Natural gas security problems in Europe: the Russian-Ukrainian crisis of 2006. *Asia-Pacific Review*, **13**, 32–59.
- Stokols, D., Hall, K. L., Taylor, B. K., and Moser, R. P. 2008. The science of team science: overview of the field and introduction to the supplement. *American Journal of Preventive Medicine*, **2**, S77–S89.
- Strang, D., and Meyer, J. W. 1993. Institutional conditions for diffusion. *Theory and society*, **22**, 487–511.
- Strang, D., and Soule, S. 1998. Diffusion in organizations and social movements: from hybrid corn to poison pills. *Annual Review of Sociology*, **24**, 265–290.
- Streeck, W., and Schmitter, P. C. 1985. Community, market, state – and associations? The prospective contribution of interest governance to social order. *European Sociological Review*, **1**, 119–138.
- Swedberg, R. 1993. Non-contractual relations in business: a preliminary study. *Journal of Institutional and Theoretical Economics*, **149**, 204–209.
- Sydow, J., Schreyögg, G., and Koch, J. 2009. Organizational path dependence: opening the black box. *Academy of Management Review*, **34**, 689–709.
- Szalay, M. S., Kovács, I. A., Korcsmáros, T., Böde, C., and Csermely, P. 2007. Stress-induced rearrangements of cellular networks: consequences for protection and drug design. *FEBS Letters*, **581**, 3675–3680.
- Szántó, Z. 2009. Kontraszelekció és erkölcsi kockázat a politikában. *Közgazdasági Szemle*, **6**, 563–571.

- Szántó, Z., Tóth, I. J., Varga, S., and Cserpes, T. 2011. A korrupció típusai és médiareprezentációja Magyarországon (2001-2009) [Types of corruption and their media representation in Hungary (2001-2009)]. *Rendészeti Szemle*.
- Thevenot, L. 2001. Organized complexity: conventions of coordination and the composition of economic arrangements. *European Journal of Social Theory*, 4, 405–425.
- Thomas, P. 2006. The communication rights in the information society (CRIS) campaign. Applying social movement theories to an analysis of global media reform. *Gazette*, 68, 291–312.
- Thye, S. R., Yoon, J., and Lawler, E. J. 2002. The theory of relational cohesion: review of a research program. *Advances in Group Processes*, 19, 139–166.
- Tilly, C. 1996. *The Contentious French: Four Centuries of Popular Struggle*. Cambridge: Harvard University Press.
- Tilly, C. 2005. *Identities, Boundaries and Social Ties*. Paradigm.
- Tilly, C., and Tarrow, S. 2007. *Contentious Politics*. Paradigm.
- Tondera, D., Grandemange, S., Jourdain, A., Karbowski, M., Mattenberger, Y., Herzig, S., Da Cruz, S., Clerc, P., Raschke, I., Merkwirth, C., Ehse, S., Krause, F., Chan, D. C., Alexander, C., Bauer, C., Youle, R., Langer, T., and Martinou, J. C. 2009. SLP-2 is required for stress-induced mitochondrial hyperfusion. *EMBO Journal*, 28, 1589–1600.
- Tóth, I. J. 2010. A szélérőművek engedélyeztetésének labirintusa [Obtaining wind power plant permits]. *Rendészeti Szemle*.
- Trajtenberg, M. 1990. A Penny for Your Quotes: Patent Citations and the Value of Innovations. *RAND Journal of Economics*, 21, 172–187.
- Turchin, P. 2003. *Historical Dynamics: Why States Rise and Fall*. Princeton University Press.
- Turchin, P. 2005. Dynamical feedbacks between population growth and sociopolitical instability in agrarian states. *Structure and Dynamics*, 1, 49–69.
- Turchin, P., and Nefedov, S. 2009. *Secular Cycles*. Princeton University Press.
- Ulanowicz, R. E. 1986. *Growth & Development: Ecosystems Phenomenology*. Springer Verlag, New York.
- Ulanowicz, R. E. 1997. *Ecology, the Ascendent Perspective*. Columbia University Press.
- Ulanowicz, R. E., Heymans, J. J., and Egnatovich, M. S. 2000. Network analysis of trophic dynamics in South Florida Ecosystems, FY 99: the Graminoid Ecosystem. *Annual Report to the United States Geological Service Biological Resources Division University of Miami Coral Gables, FL 33124*.
- Useem, M. 1980. Corporations and the corporate elite. *Annual Review of Sociology*, 6, 41–77.
- Uzzi, B. 1997. Social structure and competition in interfirm networks: the paradox of embeddedness. *Administrative Science Quarterly*, 42, 35–67.
- Uzzi, B., and Spiro, J. 2005. Collaboration and creativity: the small world problem. *American Journal of Sociology*, 111, 447–504.
- Van Der Vegt, G. S., Bunderson, J. S., and Oosterhof, A. 2006. Expertness diversity and interpersonal helping in teams: why those who need the most help end up getting the least. *Academy of Management Journal*, 49, 877–893.
- Van Dyke, N., and McCammon, H. 2010. *Social Movement Coalitions*. University of Minnesota Press.
- Walgrave, S., and Rucht, D. 2010. *The World Says No to War: Demonstrations Against the War in Iraq*. University of Minnesota Press.

- Wallerstein, I. 1974. *The Modern World-system. Vol. 1. Capitalist Agriculture and the Origin of the European World-economy in the Sixteenth Century*. Academic Press.
- Wang, P., Sharpe, K., Robins, G. L., and Pattison, P. E. 2009. Exponential Random Graph (p^*) Models for affiliation networks. *Social Networks*, **31**, 12–25.
- Wang, S., Szalay, M. S., Zhang, C., and Csermely, P. 2008. Learning and innovative elements of strategy update rules expand cooperative network topologies. *PLoS ONE*, **3**, e1917.
- Ward Jr., J. H. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, **58**, 236–244.
- Watts, D. J. 1999. Networks, dynamics and the small-world phenomenon. *American Journal of Sociology*, **105**, 493–527.
- Watts, D. J. 2003. *Six Degrees: The Science of a Connected Age*. New York: W. W. Norton.
- Watts, D. J., and Strogatz, S. H. 1998. Collective dynamics of ‘small-world’ networks. *Nature*, **393**, 440–442.
- WEC. 2007. *2007 Survey of Energy Resources*. Tech. rept. World Energy Council, London.
- Wegner, D. M. 1987. Transactive memory: a contemporary analysis of the group mind. Pages 185–208 of: *Theories of Group Behavior*. Springer-Verlag.
- Weick, K. E. 1992. *The Social Psychology of Organizing*. Addison–Wesley, Reading.
- Weisser, H. 2007. The security of gas supply – a critical issue for Europe? *Energy Policy*, **35**, 1–5.
- Weitzman, M. L. 1998. Recombinant growth. *Quarterly Journal of Economics*, **113**, 331–360.
- Wellman, B. 1979. The community question: the intimate networks of East Yorkers. *American Journal of Sociology*, **84**, 1201–1231.
- Wellman, B., and Gulia, M. 1999. Net surfers don’t ride alone: virtual community as community. Pages 331–367 of: *Networks in the Global Village*. Westview Press.
- Westman, W. E. 1990. Managing for biodiversity. *BioScience*, **30**, 26–33.
- Wheeler, S., Mann, K., and Sarat, A. 1988. *Sitting in judgment: the sentencing of white collar criminals*. Yale University Press.
- White, D. R. 2009a. *Dynamics of Human Behavior (Cohesion and Resistance)*. *Encyclopedia of Complexity and Systems Science*. Springer-Verlag.
- White, D. R. 2009b. *The evolution of the medieval world economic network, and the Chinese link*. Central European University, June Conference on “The Unexpected Link Using network science to tackle social problems”, organized by Balazs Vedres & Marco Scotti, Inauguration of the CEU Center for Network Sciences. Powerpoint: http://intersci.ss.uci.edu/wiki/ppt/economic_networksCEU09.ppt – Old world city systems and economic networks 950-1950: how the growth and decline of cities and the rise and fall of city-size hierarchies is related to the network structure of intercity connections.
- White, D. R., Tambayong, L., and Kejžar, N. 2008. Oscillatory dynamics of city-size distributions in world historical systems. Pages 190–225 of: *Globalization as Evolutionary Process: Modeling, Simulating, and Forecasting Global Change*. Routledge.
- White, D. R., Scott, D. W., Tolga, O., and Ren, F. 2011. *Exploratory causal analysis for networks of ethnographically well-studied populations*.

- White, H. 1998. *Advances in Econometric Theory: The Selected Works of Halbert White*. Edward Elgar.
- White, H. 2004. *New Perspectives in Econometric Theory: The Selected Works of Halbert White, Volume 2*. Edward Elgar.
- Williams, K. Y., and O'Reilly, C. A. 1998. Demography and diversity in organizations: a review of 40 years of research. Pages 77–140 of: *Research in Organizational Behavior*. JAI Press.
- Wuchty, S., Jones, B., and Uzzi, B. 2007. The increasing dominance of teams in production of knowledge. *Science*, **316**, 1036–1039.
- Yenikeyeff, S. M. 2008. *Kazakhstan's Gas: Export Markets and Export Routes*. Tech. rept. NG 25. Oxford Institute for Energy Studies. http://www.offnews.info/downloads/oxford_energy_kazakhstan_nov08.pdf.
- Yinger, M. J. 1960. Contraculture and subculture. *American Sociological Review*, **25**, 625–635.
- Zald, M. N., and McCarthy, J. D. 1980. Social movement industries: competition and cooperation among movement organizations. *Research in Social Movements, Conflict and Change*, **3**, 1–20.
- Ziman, J. M. 1994. *Prometheus Bound*. Cambridge: Cambridge University Press.
- Zucker, L. G. 1991. The role of institutionalism in cultural persistence. Pages 83–107 of: Powell, W. W., and DiMaggio, P. J. (eds), *The New Institutionalism in Organizational Analysis*. University of Chicago Press.