Network Models of the Diffusion of Innovations

Article *in* Computational and Mathematical Organization Theory · January 1995

1 author:

Thomas Valente
University of Southern California

Network Models and Methods for Studying the Diffusion of Innovations

Thomas W. Valente
University of Southern California

6.1 Introduction

Diffusion of innovations theory attempts to explain how new ideas and practices spread within and between communities. The theory has its roots in anthropology, economics, geography, sociology, and marketing, among other disciplines (Hägerstrand 1967; Robertson 1971; Brown 1981; Rogers 2003), and has in some ways been adapted from epidemiology (e.g., Bailey 1975; Morris 1993). The premise, confirmed by empirical research, is that new ideas and practices spread through interpersonal contacts largely consisting of interpersonal communication (Ryan and Gross 1943; Beal and Bohlen 1955; Katz, Levine, and Hamilton 1963; Rogers 1995; Valente 1995; Valente and Rogers 1995).

In their pioneering study, Ryan and Gross (1943) laid the groundwork for the diffusion paradigm by showing that, among other things, social factors rather than economic ones were important influences on adoption (Valente and Rogers 1995). Hundreds of diffusion studies were conducted in the 1950s and early 1960s to examine the diffusion process in more detail across a variety of settings (Rogers 2003). Many studies sought to understand how information created in government or otherwise sponsored programs could be disseminated more effectively. Diffusion research peaked in the early 1960s, but has been reinvigorated more recently with the advent of more sophisticated network models and technology making it possible to study the diffusion process more explicitly.

Most diffusion studies focus on trying to understand the factors that lead some members of a population to adopt a new idea and others do not. Further, studies try to understand why some people adopt the behavior early, whereas others wait a substantial amount of time before accepting the new practice. For example, Ryan and Gross (1943) wanted to know why some farmers purchased hybrid seed corn almost immediately upon its availability, whereas others waited until almost all the farmers in the area purchased it before they were willing to do so. Similarly, Coleman, Katz, and Menzel (1966) wanted to know why some physicians began prescribing tetracycline as soon as it was available, whereas others waited until most physicians prescribed it before they were willing to do so.

This chapter describes a variety of mathematical and network models used to study the diffusion of these and other innovations. The Coleman and others (1966) study provided a conceptual leap from other diffusion studies by explicitly measuring who talked to whom within the community about the innovation. (NB: Rogers also collected such data in his dissertation on the diffusion weed spray in Iowa.) Burt (1987) unearthed the Coleman and other (1966) data and made it available to the network community so scholars could debate various models used to describe the network diffusion process. Although having data has been useful for clarifying diffusion models, the limitations of these data and this study make it a poor choice for studying adoption behavior. Rather, scholars should have focused on collecting better data or reanalyzing diffusion network data in which contagion are more likely.

This chapter chronicles the development of network diffusion models and indicates where such progress is being made. I first present macro models used to estimate the speed of diffusion and, with the Bass (1969) model, to estimate rates of innovation and imitation. Next, spatial autocorrelation is presented that is used to estimate the degree to which contiguous nodes adopt innovations. Spatial autocorrelation led to the network autocorrelation model, which is presented statically (cross-sectional data only) and then with one time lag. I then discuss event history analysis applications of network autocorrelation and its extension by including time-based network interaction terms. Throughout the chapter, I attempt to provide a review of more recent research conducted in a variety of domains, but mostly drawn from the public health field.

(A) *Macro Models*

One consistent finding of diffusion research has been that the cumulative pattern of diffusion follows a growth pattern approximated by a simple one-parameter logistic function such as:

$$y_t = b_0 + \frac{1}{1 + e^{-b_1 t}}, \tag{6.1}$$

where y is the proportion of adopters, b_0 the y intercept, t is time, and b_1 the rate parameter to be estimated. This simple model can be used to compare growth rates for various innovations, but is extremely limited in its applicability. A considerable improvement was advanced by Bass (1969) and many others (see Hamblin, Jacobsen, and Miller 1973; Mahajan and Peterson 1985; Valente 1993) by creating a two-parameter model:

$$y_t = b_0 + (b_1 - b_0)Y_{t-1} - b_1(Y_{t-1})^2, \tag{6.2}$$

where y is the proportion of adopters, b_0 a rate parameter for innovation, and b_1 a rate parameter for imitation (the degree of adoption due to prior adopters). The Bass model incorporates the percentage adopters at each time point and thus makes a better estimate of the growth attributable to personal network persuasion. The mathematical model in (6.2) can be used to (1) forecast expected levels of diffusion (Mahajan and Peterson 1985), (2) estimate the rate of diffusion attributed to different theoretical aspects of the diffusion processes, b_0, external influence or innovativeness, and b_1, internal influence or interpersonal persuasion (Bass 1969; Hamblin et al. 1973; Valente 1993). This model can be used to estimate rate of disease spread from a central source such as contaminated food or from infections spread through interpersonal contact. In

Table 6.1.1. *Diffusion Rate Parameter Estimates and Moran's* I *Estimates for Two Data Sets*

	Medical Innovation	Cameroon Tontine 1 Simulation
One-parameter model		
Coefficient (95% CI)[a]	0.23 (−.053–0.51)	0.06 (.01–0.12)
N	17	50
R^2	0.76	0.71
Two-parameter (Bass) model		
Innovation coefficient (95% CI)	−0.43(−0.83–0.03)	−0.20 (−0.30–0.09)
Imitation coefficient (95% CI)	4.09 (3.05–5.12)	2.96 (2.58–3.34)
N	16	49
R^2	0.89	0.89
Moran's I	−.13	−.08
z-Score	−6.73	−7.80

[a] CI, Confidence Interval.

the social realm, one can use the model to estimate rate of adoption from a mass media advertisement or from interpersonal influence. Rate parameter estimates from both models for two diffusion data sets are provided in Table 6.1.1. Interpretation of these estimates is highly dependent on the time scale used to measure diffusion.

These rate parameter estimates can be used as outcomes to study factors associated with diffusion at the macrolevel by comparing rates between groups and/or populations. For example, parameter estimates for different countries can be compared in order to study factors associated with the spread of behaviors in different countries. Modeling at this macrolevel, however, is imprecise at best because it assumes perfect social mixing, everyone interacting with everyone else (Granovetter 1978; Van den Bulte and Lillien 1997). These macro models do not measure whether people who are connected to one another engage in the same behaviors. Geographers have devoted considerable attention to trying to determine whether innovations spread between contiguous areas.

(B) *Spatial Autocorrelation*

Rather than just estimate rate of diffusion, spatial models measure whether artifacts, diseases, farming practices, and other behaviors spread between contiguous areas (Hägerstrand 1967; Cliff and Ord 1981; Griffith et al. 1999). Proximity data are easy to obtain and are relatively unambiguous, thus providing a network of connections based on distance. Moran's I (1956) was an early model developed to test for spatial association, geographic clustering of adoption:

$$I = \frac{N \sum_i^N \sum_j^N D_{ij}(y_i - \bar{y})(y_j - \bar{y})}{S \sum_i^N (y_i - \bar{y})^2}, \tag{6.3}$$

where N is the sample size, D a distance matrix (as proximities), y indicates adoption, and S the sum of the distances in the distance matrix. Moran's I measures the degree to which nodes that are connected to one another deviate from the average behavior in the network similarly or differently. Moran's I is high when connected nodes (positive elements of D) are either positively or negatively different from the average score. The statistical significance of Moran's I can be calculated in two ways – via permutation methods or analytically.

To use a permutation method to calculate the significance of Moran's I, assume adoption (y_i) is randomly distributed and calculate I repeatedly to get a sample of estimates based on D and the number of adopters. If Moran's I calculated is significantly different than the random sample generated, Moran's I is considered significant (z-scores can be obtained). The logic then is to calculate the degree to which neighbors (however defined) have similar adoption behavior compared with that expected if adoption were distributed randomly. Variance estimators for Moran's I can be found in spatial statistics textbooks (Cliff and Ord 1981; Bailey and Gatrell 1995) and used to calculate exact significance tests. Moran's I is useful and has been extended considerably (Nyblom, Borgatti, Roslakka, and Salo 2003), yet this approach often assumes that geographic proximity equates with communication and influence, which may not be true.

The spatial autocorrelation methodology was seen as a useful approach to measuring network autocorrelation, the bias inherent in a regression model when y appears as both the dependent and independent variable. Erbing and Young (1979) wrote an influential paper on measuring network effects and using network autocorrelation methods. Dow (1986) demonstrated the effects of network autocorrelation on estimate errors, and Doriean, Teuter, and Wang (1984) found considerable bias in the point estimates and their standard errors. Exactly how network autocorrelation applied to diffusion of innovations was not clear because spatial autocorrelation measured diffusion at the macrolevel, but did not show whether specific individuals were more or less likely to adopt based on their network position. Further, spatial autocorrelation did not show how network structure influenced diffusion. To do so, we turn to network models.

(C) *Network Models*

Figure 6.1.1 displays two networks from a study conducted in Cameroon among women in voluntary organizations (Valente et al. 1997). Women were asked to name their friends in the organization in an attempt to determine if friendship ties were associated with contraceptive choices (they were). The diffusion network model posits that initial contraceptive choices would be made by some women based on their innovativeness and exposure to outside sources of influence such as their cosmopoliteness, media use, or greater need for the innovation. The new idea, and its practice, then spreads through the network as users persuade nonusers to adopt either by exhortation, entreaty, enticement, or example.

Network influences are captured by an exposure or contagion model (Figure 6.1.2), and each individual's likelihood of adoption increases as the proportion (or number) of users in his or her personal network increases. Personal network exposure is the proportion or number of adopters in each person's network that provide information

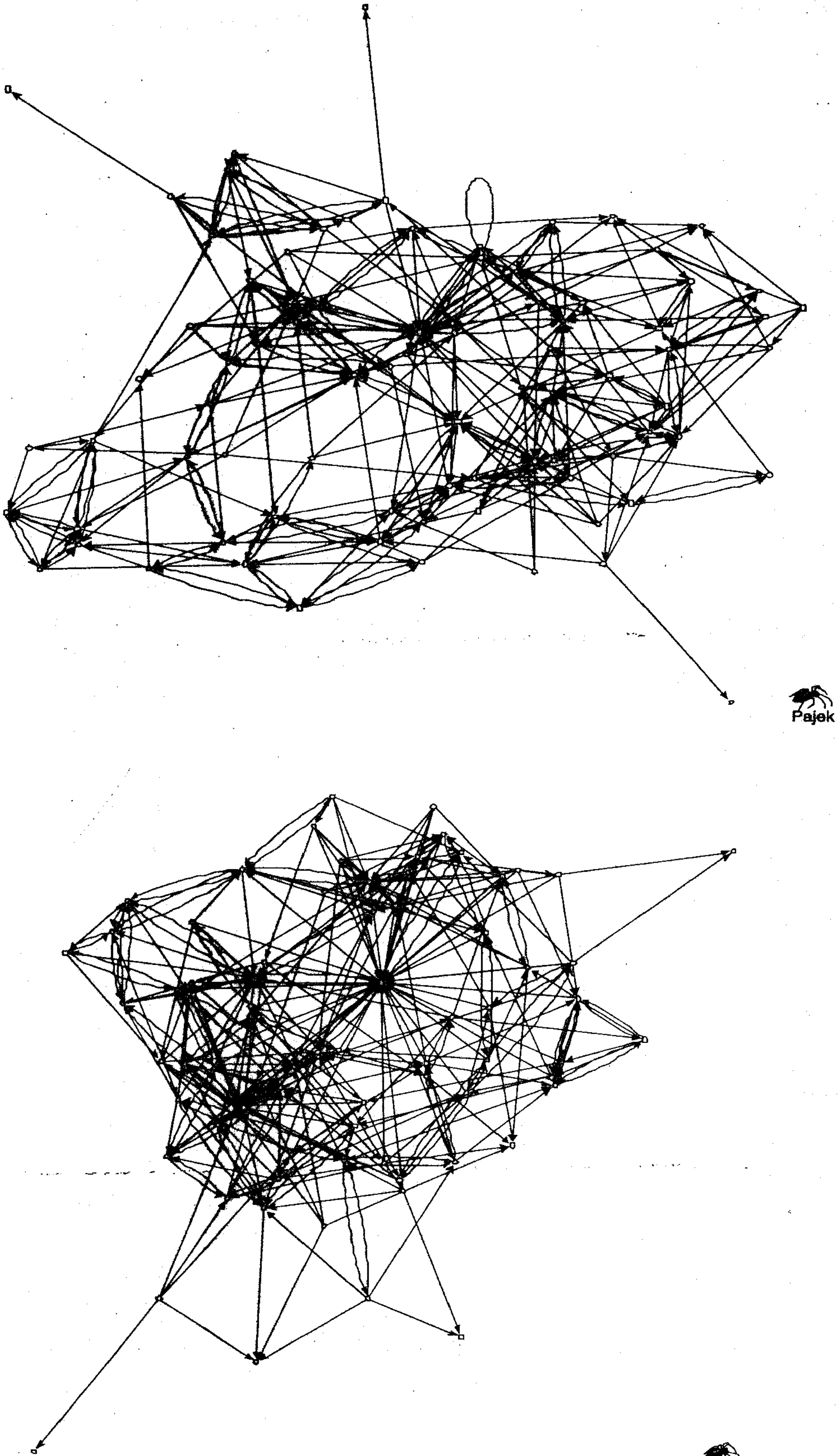

Figure 6.1.1. Networks 1 and 2 from the Cameroon Voluntary Association Study.

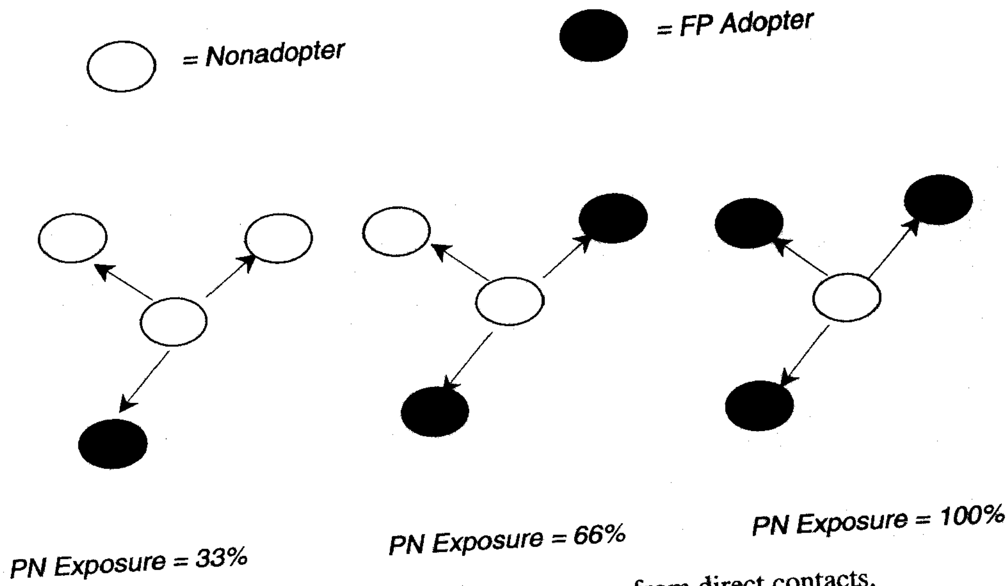

Figure 6.1.2. Personal network exposure from direct contacts.

and influence with regard to some behavior. The equation for nonrandom mixing, or personal network exposure, is:

$$E_i = \frac{\sum w_{ij} y_j}{\sum w_i}, \tag{6.4}$$

where w is the social network weight matrix and y is vector of adoptions. For an individual who reported three contacts, network exposure (E_i) is the proportion of those contacts that have adopted (Figure 6.1.2). When network exposure is measured on direct contacts, it captures social influence conveyed through overt transmission of information, persuasion, or direct pressure. Alternatively, exposure can be calculated by transforming the social network, W, to reflect other social influence processes. For example, W can be transformed to represent the degree of structural equivalence (similarity in network position) among people in the network. Exposure calculated on this network captures social influence conveyed via comparison to equivalent others by social comparison or competition (Burt 1987). Exposure can also be weighted by network properties such as centrality to reflect social influence by opinion leaders.

These three social influence processes are modeled with three different classes of network weight matrices (relational, positional, and central) constructed from the same social network data (Table 6.1.2). All three can be justified theoretically as sources of influence on adoption behavior, and all three can be calculated various ways (there are at least ten centrality measures). It is possible that all three operate for different people or at different times during the diffusion process.

In addition to the social influence process, a second dimension to these influence mechanisms is the weight attached to each based on social distance. For example, in relational influence models, different weights can be assigned to direct ties, ties of ties, and even the ties of ties of ties; in positional equivalence models, different weights can be assigned to those that are more equivalent than others (Valente 1995). A potential line of diffusion network research then is to compare different network weighting mechanisms in order to model and compare different social influence processes.

Table 6.1.2. *Social Network Influence Weightings*

Relational	Positional	Central
1. Direct ties	1. Percent positive matches (tie overlap)	1. Degree
2. Indirect ties	2. Euclidean distance	2. Closeness
3. Joint participation in groups or events	3. Regular equivalence	3. Betweenness
		4. Flow
		5. Integration/radiality
		6. Information
		7. Power

Diffusion was simulated through the two Cameroon networks in Figure 6.1.1 to illustrate how network exposure and network structure influence diffusion. At each time period, adoption occurred for the nonadopter with the most nominations received, then network exposure was calculated, all nodes with exposure of 50% or higher were categorized as adopters, and the process repeated. We compared diffusion in this network with that simulated in a network of the same size and density, but with links allocated randomly. Both conditions were averaged across 1,000 runs. Figure 6.1.3 shows that, in network 1, initially the diffusion trajectories are similar, but at about time 10, diffusion in the actual network accelerated.

The network accelerated diffusion because it is somewhat centralized (in-degree 21.7%), and once diffusion reaches the center of the network it can propagate rapidly. Notice that at about time 20 diffusion slowed, accelerated again at about 25, and then slowed from about time 30 to 40. These "fits and starts" are a product of the network structure: diffusion reaches pockets of interconnectivity and spreads rapidly within these dense pockets, but slows between groups. Network 2 (Figure 6.1.3) had more rapid and sustained diffusion because it was even more centralized (in-degree, 47.2%). Note that, in the spatial autocorrelation model, adoptions were randomized to measure statistical significance, and in this simulation, the network structure was randomized to illustrate its influence on the rate of diffusion.

Simulation assumptions regarding influences on adoption could easily be changed to achieve different outcomes. For example, when adoptions were assigned randomly, diffusion was constant in network 1 (and saturation lower) and similar to the random network in network 2. The validity of these diffusion models rests partly on determining whether network exposure influences adoption. To that end, a number of empirical studies have been conducted to measure the degree to which social network exposure is associated with adoption.

6.2 Empirical Studies

Empirical support for an association between one's own behavior and that of one's peers can be found throughout the behavioral sciences literature. Although many scholars assume adoption is associated with network exposure, few studies have traced an innovation through a network of social contacts to empirically validate this proposition.

Network 1

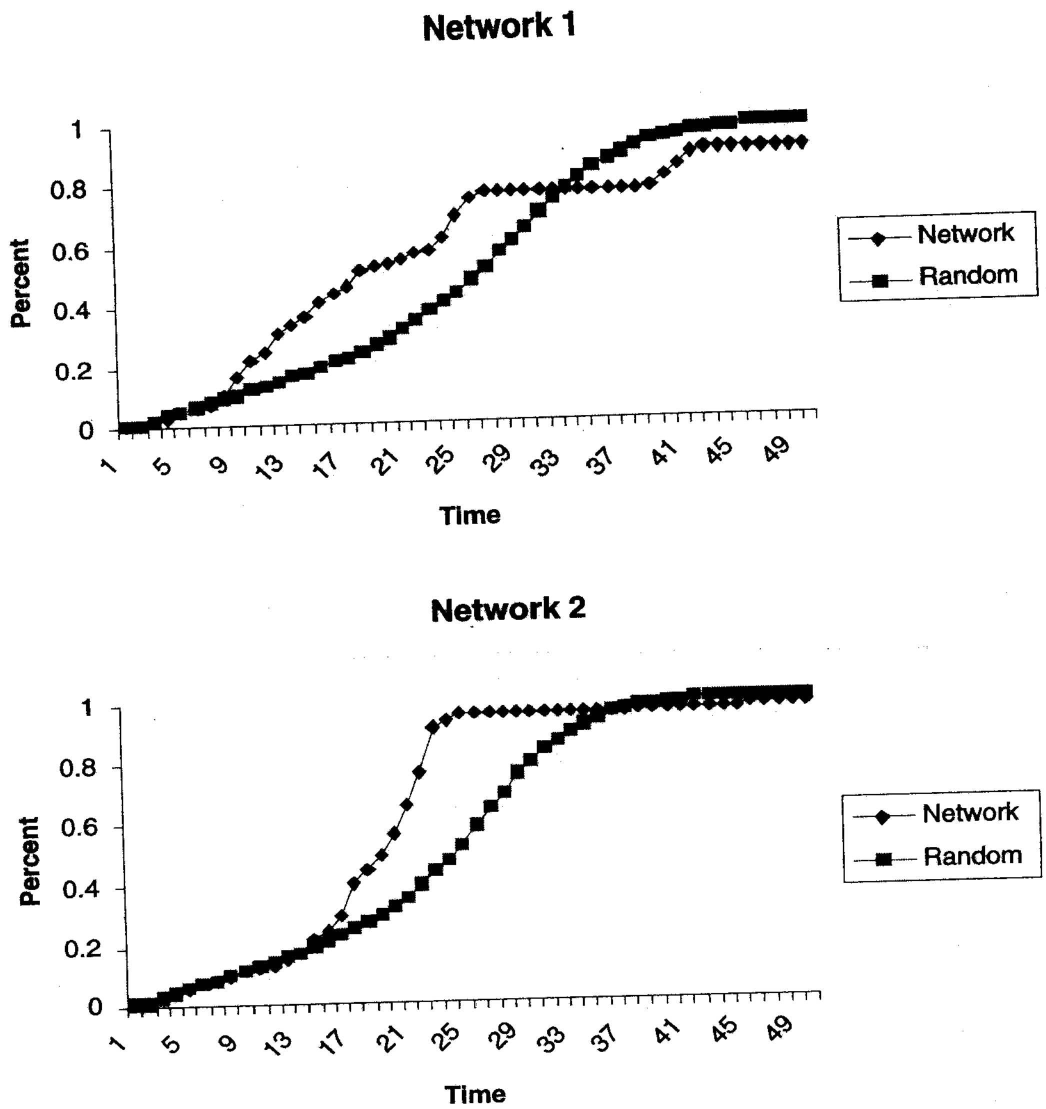

Figure 6.1.3. Simulated diffusion in two networks, each compared with random networks of the same size and density. Network 2 is more centralized.

The lack of data on diffusion within an entire network stems largely from the difficulty of trying to collect data over a time period long enough for diffusion to occur. Consequently, most studies have relied on retrospective data that introduces some but not much bias (Coughenour 1965; Nischan et al. 1993). It has also meant that several scholars have reanalyzed two studies that collected network and adoption data: (1) medical innovation study (Coleman et al. 1966), reanalyzed by Burt (1987), Marsden and Podolny (1990), Strang and Tuma (1993), Valente (1995, 1996), and Van den Bulte and Lillien (2001); and (2) Korean family planning study (Rogers and Kincaid 1981), reanalyzed by Dozier (1977), Montgomery and Chung (1999), Kohler (1997), and Valente (1995, 1996). More recent studies in the fields of reproductive health (Casterline 2001) and substance abuse (Neaigus et al. 2001) have provided new data, but these classics remain classic.

Because collecting complete network data is difficult, most empirical research has been egocentric (Marsden 1987, 1990), based on respondent reports of their behavior and that of their network peers who are not necessarily connected to one another and not interviewed. Social influence is often based on respondent reports of perceptions of peer behavior or perceptions of peer influence (Valente and Saba 1998, 2001; Valente and Vlahov 2001). Comparison of exposure scores based on respondent perceptions with alters' reports in one study found that perceptions were more strongly related to behavior than exposure based on alter reports (Valente et al. 1997; also see Urberg, Degirmencioglu, and Pilgrim 1997; Montgomery and Chung 1999).

Sociometric studies interview members of a bounded community and attempt to gather information from everyone in the community (typically conducted in schools, organizations, and small communities) and record their time of adoption (Coleman et al. 1966; Becker 1970; Rogers and Kincaid 1981; Wasserman and Faust 1994; Scott 2000). Sociometric studies are useful for understanding how an innovation flows within the community and how certain network structural variables influence the diffusion process. Sociometric data capture network influences by the alters' reports because they were also interviewed. For example, sociometric studies can determine whether structural positions such as centrality are associated with adoption and/or whether centralization is associated with more rapid diffusion (Valente 1995).

A number of more recent diffusion network studies have been cross-sectional and, in many cases, retrospective involving only one time point. For example, a study in Thailand by Entwisle and others (1996) found that contraceptive choices made by early adopters contributed significantly to the contraceptive choices made by later adopters (also see Rogers and Kincaid 1981). Valente and others (1997) collected sociometric data on contraceptive use among women in voluntary associations in Cameroon and showed that perceptions of these friends' behavior, and in particular, perceptions that these friends encouraged contraceptive use, were significantly associated with behavior. In general, these statistical analyses use the following model:

$$\log \frac{\Pr(y_t = 1)}{(1 - \Pr(y_t = 1))} = \alpha + \sum B_k X_k + B_{(k+1)} \omega y_t, \tag{6.5}$$

where y is a binary vector of adoption behavior, α is the intercept, β_k are parameter estimates for vectors of K sociodemographic characteristics (Xs), and ω represents the social network matrix. The ωy_t term represents the calculation of contemporaneous network exposure, and this vector is usually divided by a count of the number of nominations sent (alternatively, the number of nominations can be entered into the regression separately).

Significant estimates for β_{k+1} indicate contagion effects by showing that network exposure is associated with adoption. The variances for these estimates, however, are usually biased because the observations are not independent; hence, the errors in prediction are not independent. One partial solution is to obtain robust estimates by controlling for clustering. Clustering is the degree that elements from the same cluster are similar compared with those of different clusters. For example, two individuals chosen at random from the same organization are more likely to be similar than two chosen at random from different organizations. Table 6.2.3 reports regression results of the Cameroon data with and without correction for clustering. Without correction, network exposure is strongly and significantly associated with adoption, but with the correction it is only marginally statistically significant ($p = .04$). Controlling for clustering is

Table 6.2.3. *Logistic Regression on the Likelihood of Contraceptive Behavior on Controls and Network Exposure with and without Correction for Clustering* (N = 555; *Groups* = 9)

	Contraceptive Method Use			
	Without Correction		With Correction	
	Adjusted Odds Ratios	*P*-Value	Adjusted Odds Ratios	*P*-Value
Age	0.97	0.001	0.97	0.010
Education	0.91	0.247	0.91	0.184
Possessions	1.39	0.000	1.39	0.000
Network exposure	1.14	0.005	1.14	0.047

particularly important in network exposure models because network choices are often restricted to the cluster.

Even with clustering controlled, social influence as measured through social networks seems to be strongly associated with behavior. For example, a school-based sociometric study was conducted by Alexander and colleagues (2001) using adolescent health data (Bearman, Jones, and Udry 2000) to show that students with a majority of network ties who were smokers were almost two times as likely to smoke themselves, with an additional two times greater likelihood of smoking for those with best friends who smoke. Intra-school clustering was controlled and the multilevel model accurately captured microlevel effects within the context of macrolevel influences. The study measured the influence of peers on smoking, while conditioning on the smoking rate within the school (Alexander et al. 2001).

Estimating the network exposure (autocorrelation) term with a multilevel model can provide contagion estimates across settings and estimate the degree it varies between settings (i.e., communities, schools, organizations, etc.). The models are incomplete, however, because there may be factors that influence both adoption and choice of social network contacts. For example, the decision to smoke and to nominate friends who smoke may both be a function of delinquency or rebellion. Hence, an association between behavior and peer behavior can be spurious. Testing social influence with network methods then requires longitudinal data involving at least two time points. Boulay and Valente (in press) collected data among women in three villages of Nepal and found that having discussion partners who used contraception influenced information-seeking behavior and contraceptive choice. Having data from two time points allows testing of a simple dynamic model on adoption:

$$\log \frac{\Pr(y_t = 1)}{(1 - \Pr(y_t = 1))} = \alpha + \sum B_k X_k + B_{(k+1)}\omega_t y_t + B_{(k+2)}\omega_{(t-1)} y_{(t-1)}, \tag{6.6}$$

where y is a binary indicator of behavior, α is the intercept, β_k are parameter estimates for vectors of K sociodemographic characteristics (Xs), and ω represents the social network matrix. A positive and significant β_{k+2} indicates that respondents with high network exposure at baseline were more likely to adopt at time two. A positive and significant β_{k+1} indicates that change in network exposure is associated with change in behavior. This may indicate contagion, but still may be a product of some omitted factor. Panel data collected at two time periods are adequate for most research needs and can provide evidence of network influences on behavior. However, because there is often a considerable time between the two measures, many factors may account for simultaneous change in behavior and network exposure. To cope with this threat, data can be collected on time of adoption, expanding the microlevel dynamic analysis by using event history analysis (Tuma and Hannan 1984).

(A) *Event History Analysis*

Event history analysis techniques have been developed to analyze data with a substantive number of time points, estimating coefficients with maximum likelihood estimators (Bartholomew 1982; Allison 1984; Tuma and Hannan 1984; Strang and Tuma 1993; Teachman and Hayward 1993). There are two types of event history analysis, discrete time, in which the outcome is binary, and continuous time, in which the outcome is time to an event. Because diffusion occurs over time, there is an explicit time dimension in diffusion studies captured by both discrete and continuous time models. The time of adoption variable is the dependent variable and may be influenced by both time-varying and time-constant factors. Some individuals may not have adopted by the time of data collection giving rise to time-censored observations (right censoring occurs when the data are collected before the innovation has finished diffusing or does not diffuse to all members of the community or study).[1]

There are a variety of event history techniques, including hazard models developed in epidemiology, used to understand the hazard or risk to disease or injury over time. Hazard and/or event history analysis generally requires that the data are reshaped from simple observations to a case–time format, such that there is a case in the data for each individual at each time period of study up to and including that person's time of adoption (Table 6.2.4). The time-varying and time-constant independent variables are included in each case, as well as a binary indicator for whether the individual adopted the behavior (or got sick).

Maximum-likelihood estimation can determine whether the independent variables are associated with the dependent variable (adopt/not adopt) (Eliason 1993). A study of 100 people with an average adoption time of seven translates into 700 person–time cases. Each person–time case has a variable for the network exposure at that time period plus an indicator for whether the person adopted (plus additional time-constant and time-varying covariates as desired). The event history model is:

$$\log \frac{\Pr(y_t = 1)}{(1 - \Pr(y_t = 1))} = \alpha + \sum B_j X_j + \sum B_{kt} X_{kt} + \sum B_{(k+1)} \omega y_t, \qquad (6.7)$$

where y is a binary indicator of behavior, α is the intercept, β_j are parameter estimates for vectors of J sociodemographic characteristics (X_j), β_{kt} are parameter estimates for the matrix of time-varying sociodemographic characteristics (X_{kt}), ω represents the social network weight matrix, and t a time indicator. Note here we have assumed a

	Medical Innovation ($N = 947$)	Brazilian Farmers ($N = 10{,}092$)	Korean Family Planning ($N = 7{,}103$)
	0.21	0.72	0.31

Table 6.2.4. *Event History Analysis of Factors Associated with Adoption for the Three Diffusion Network Datasets (Coefficients Are Adjusted Odds Ratios for Likelihood of Adoption).*

	Medical Innovation ($N = 947$)	Brazilian Farmers ($N = 10{,}092$)	Korean Family Planning ($N = 7{,}103$)
Time (recoded as proportion)	0.21	0.72	0.31
Time logistically transformed	0.68	1.94	0.67
Infection	11.8*	10.9**	9.26**
Susceptibility	2.31	2.24*	2.44*
Number sent	0.91	0.90†	0.96
Number received	1.06†	1.02†	1.06**
Exposure via direct contacts	0.64	1.07	1.19
Exposure via structural equivalence	0.93	2.47**	1.12
Attitude toward science	0.65**		
Journals	1.84*		
Income		1.17**	
Visits		1.00	
No. of children			1.25**
Campaign exposure			1.03†

† $p < .10$; * $p < .01$; ** $p < .001$.

static (constant) network. Standard statistical packages allow testing of event history or survival data in a relatively straightforward manner, once the data are reformatted. Event history analysis requires the construction of exposure matrices for each time period, which can be a formidable task, particularly if one uses more than one network weight matrix (Valente 1995).

Marsden and Podolny (1990) used event history analysis and tested network exposure's association with adoption in the medical innovation data. Results showed that exposure was not associated with adoption in that study. Strang and Tuma (1993) revisited the issue with the same data by postulating time variance in network influence (i.e., how much lag time, if any, is there in the influence). Strang and Tuma (1993) found evidence of contagion. Van den Bulte and Lillien (2001) supplemented the medical innovation data with archival data on media promotion by pharmaceutical firms at the time of the original study and showed that network contagion effects disappear once these data are added. Their analysis demonstrates the importance of omitted variables when studying diffusion through networks. The rapid diffusion measured in the medical innovation study indicates that contagion was probably not the primary factor driving diffusion.

To illustrate, I conducted event history analysis of three classic diffusion network data sets. The analysis controlled for within village and within person covariation, as well as terms for time and a logistic transformation of time, were included to control for macrolevel effects. Terms for infection and susceptibility (Strang and Tuma 1993; Myers 2000) were included to measure whether adoption by central individuals (high in-degree) influenced subsequent adoption – infection – and whether centrality (out-degree) influenced a person's likelihood to adopt as diffusion occurred – susceptibility.

In- and out-degree were also included in the model. Two network exposure terms were computed – direct ties and structural equivalence.[2] For this analysis, network exposure was calculated using contemporaneous measures because two of the data sets recorded adoption in 1-year intervals. Two control variables representing individual characteristics were included. Analysis was conducted only on those who adopted. The following model was estimated:

$$\log(\Pr(y_t = 1)) = \alpha + \sum \beta_{lm} X_{lm} + \sum \beta_{lmt} V_{lmt} + \sum \beta_{lmt} \omega_s y_t + \lambda_{lm1} C_D(y_+) + \lambda_{lm2} C_D(y_+), \qquad (6.8)$$

where y is a binary indicator of behavior; α is the intercept; Xs are vectors or time-constant sociodemographic and network characteristics; V represents vectors of time-varying terms, in this case, time and its transformation; ω_s represents the social network matrices; and λ estimates the effects of centrality degree variables multiplied by the time-varying proportion of adopters in the network (infection and susceptibility). Results are mixed, but seem to indicate that both infection and susceptibility effects are present. In all three data sets, infection is positively associated with adoption indicating that, as those with high in-degree adopt, it increases the likelihood others in the network will adopt. In two studies, Brazilian farmers and Korean women, susceptibility is associated with adoption, indicating that those with a high number of nominations sent are more likely to adopt as the innovation diffuses. Ties sent and received are marginally associated with adoption, and only for the Brazilian data is exposure, through structural equivalence, associated with adoption. These results, however, change dramatically when nonadopters are included or when a term for the average exposure at each time period is included such that infection and susceptibility effects disappear.

The event history analysis approached has also been used by Montgomery and others (2001) using egocentric data to study network exposure's influence on contraceptive use in Ghana. Current analysis of four rounds of data over 2 years has shown that contraceptive use is strongly associated with use by social network peers. The Ghana field study provides some of the most conclusive evidence of the magnitude of social influence on behavior change by showing that, as the number of social network contacts who use contraceptives increases, the likelihood of contraceptives use by ego also increases. Of all variables, the network exposure variables were the most significant influences on contraceptive adoption. Another longitudinal field study in Kenya found similar results, again based on egocentric network data (Behrman, Kohler, and Watkins 2003).

Montgomery and others (2001) also reported preliminary analysis of network influences weighted by tie characteristics, such as the frequency of communication. They found that adding these weights did not change the strength of peer influence. Similar results have been reported in Valente (1995) and Valente and Saba (1998: p. 109). Consequently, it seems that the influence of social networks on behavior (contraceptive use in these cases) seems broad in nature and is not conditioned on specific factors such as the frequency of communication between dyads or their sociodemographic similarity. These factors may play a strong and even pervasive role in determining who is connected to whom (White and Watkins 2000), but they do not seem to determine the degree of influence social contacts provide.

Network exposure and adoption may not always be strongly correlated for a number of reasons. First, exposure may not be associated with adoption for everyone, but may be most influential during the middle stages of diffusion, when awareness and uncertainty about its relative advantages are both high (Carley 2001). Exposure may have less of an effect early in the process when there are few adopters and obvious advantages to waiting, and late in the process when most people have a majority of adopters in their personal network anyway. Second, individuals may have varying thresholds to adoption, such that some are innovative and others are not (Granovetter 1978). Valente (1995, 1996) posited a social network threshold model in which contagion (majority rule) is a special case. Most simulation models assume majority influence on adoption decisions as was done in the beginning of this chapter. It is reasonable, however, to expect that individuals vary in the amount of network exposure needed to adopt an innovation. Disproving thresholds may not be possible, but construct validity for the concept has been demonstrated (Valente 1996). Valente and Saba (1998) replicated the threshold model using egocentric data and showed that people with a minority of network members using contraception had higher campaign recall, indicating that the media campaign could substitute for interpersonal sources of influences. If thresholds vary, network exposure is needed for people to reach those thresholds; if they do not and the special case of contagion exists, network exposure will determine when individuals adopt.

In spite of the impressive list of studies showing some support for an association between one's own behavior and network exposure, and the theoretical simulations of network structure and thresholds, significant work remains to be done. Most scholars and lay people would agree that social networks influence behavior. The barriers to demonstrating this effect, however, have been challenges of data collection and agreement on appropriate statistical methodology. The most commonly analyzed data set, Medical Innovation, is 50 years old, consists of only 125 respondents, and arguably is not a diffusion study at all. Further, and perhaps most damaging, is that we have probably approached the problem wrong all along.

Although the distinction between dimensions of social influence (Table 6.1.2) represents a rich sociological map of influences on behavior, empirical investigations to date have found little variation in their role in the adoption process. Most network studies of diffusion are small bounded communities and hence do not differentiate much between cohesive and structurally equivalent alters. Therefore, measuring the influence of direct ties on adoption is probably sufficient for most studies, although future research comparing social network influence mechanisms could still be quite interesting, particularly in business settings where positional equivalence is likely to be a stronger influence, the majority of attention will still be paid to understand how direct contact influences adoption decisions.

6.3 Network-Based Interventions

Debate concerning network tie selection and the difficulty of specifying the time order of adoption will be hard to resolve. In addition, many behavioral scientists will argue over the meaning of any associations between network exposure and adoption. Specifying the direction of causal influences is likely to be difficult, no matter how complete the data. It may be that the best use of network data and the best way to demonstrate network influences on adoption is to design behavior change interventions. If these network-based interventions are successful, the value in understanding network models of diffusion will be apparent.

Several studies (Lomas et al. 1991; Latkin 1998; Soumerai et al. 1998; Kincaid 2000; Sikkemma et al. 2000) have identified opinion leaders using network data and had these leaders implement successful behavior change programs. Valente and Davis (1999) further suggested that leaders could be matched to others in the network based on minimum distances, and a randomized trial using this technique for preventing smoking among middle school students found it to be successful (Valente et al. 2003). Broadhead et al. (1998) and Latkin (1998) demonstrated the utility of networks for recruitment into behavioral change programs. Given the challenges inherent in collecting full diffusion network data, using networks as intervention points may present the best opportunity for understanding how networks influence behavior change.

A second network-based intervention is to target promotional programs to subgroups defined by social network affiliations. The subgroup becomes a source of social support and behavioral reinforcement not available if behavior change is spread out among people in the larger group. Critical mass is more likely to be achieved in the subgroup than in the community as a whole. A third approach would be to locate network bridges – linking agents – between organizations and subgroups, and to provide the support they need to transport new ideas and procedures between groups. A fourth approach is to locate isolates who may be at risk of not receiving information through the network or who may feel "left out" of activities. Finally, promotional programs might try to match structurally equivalent individuals and groups so messages and programs are appropriately tailored. In sum, most marketing programs have segmented audiences on demographic characteristics, and some on psychographic ones, but a new era of sociometric segmentation is now possible.

6.4 Conclusion

Much progress has been made since 1943 when Ryan and Gross first laid the foundation for diffusion of innovations theory. Rogers (2003) chronicled the many studies conducted since then and helped create a general diffusion model with wide applicability now being renewed and reinvigorated with fresh theory and analytic models. Overall, results indicate that social network influences on behavior are important and have consequences for the health and well-being of populations and individuals. These new insights have shed light on important aspects of how new ideas and practices spread within and between communities.

Along with new insights have come new questions and new perspectives to be addressed. It is clear that a lack of data on both time of adoption and network influences has hampered developments. Few diffusion or behavioral studies collect information on networks, and conversely few network studies record time of adoption. There are advantages to marrying these two ideas, however, and future research will hopefully try to collect both types of data.

It is also clear that our understanding of how diffusion occurs is still somewhat limited. The Medical Innovation data have often been used to demonstrate the importance of networks in adoption, yet analyses by Valente (1995) and Van den Bulte and Lillien (2001) have shown that contagion via social influence in this setting was unlikely. Given the number of confounding factors and some of the data requirements, it may be prohibitively difficult to substantiate the role of social networks in innovation adoption via survey methods alone. Purposively intervening on social networks, however, may prove to be a fruitful avenue of research. If network-based interventions can be used to accelerate innovation diffusion, then a stronger case can be made for the importance of social contagion in the diffusion process.

Nonetheless, it is clear that networks are important influences on behavior because most people acknowledge that they receive information and influence via their social networks and that they model the behavior of others. What is less clear is how to capture that influence in quantitative terms that mimic the theoretical progress made in the network field. Further, verbal accounts on how people make decisions and adopt behaviors usually reveal nonlinearities, chance circumstances, and whims that are not independent of networks, but not easily captured in social influence models.

The link between micro- and macrolevels of analysis represents an opportunity for study of diffusion processes. The opportunity lies in the fact that multilevel modeling techniques enable the separation of microlevel network exposure influences from macrolevel contextual factors. Yet both are social network influences and both represent elements of the diffusion paradigm. It is hoped that by controlling for contextual effects we do not "throw the baby out with the bathwater" by eliminating the microlevel influences that provide expressions for those contextual effects.

In spite of controls for macrolevel contextual effects, microlevel associations between peer network behavior and those of respondents are still sometimes strong. Debate remains about the meaning of these associations; is it peer influence, peer selection, or further contextual effects? More rigorous studies may eventually tease this out; in the interim, better study designs and interventions will need to be created. This review has attempted to point out some of the challenges diffusion scholars face and some of the promising new directions it may take. It is hoped that such organization will clear the way for promising new studies to be conducted. Social networks are fundamental influences on human behavior and conduits for the diffusion of ideas and practices, yet their roles are varied and complex and defy easy categorization.

Acknowledgments

My thanks to Christophe Van den Bulte, John Casterline, Chih-ping Chou, and Peter Marsden for helpful comments on earlier drafts. Support for this research was provided by NIDA grant no. DA10172 and DA16094.

Endnotes

1. Left censoring occurs when the data are incomplete at the beginning of the process. For example, adoption data for the period 1993 to 2000 may have some people who adopted in 1989 to 1992 classified as 1992 adopters.
2. Structural equivalence was computed as in Burt's (1987) measure, and Euclidian distance was raised to the sixteenth power.

References

Alexander, C., Piazza, M., Mekos, D., and Valente, T. W. (2001). Peer networks and adolescent cigarette smoking: An analysis of the national longitudinal study of adolescent health. *Journal of Adolescent Health, 29*, 22–30.

Allison, P. D. (1984). *Event History Analysis*. Newberry Park, CA: Sage.

Bailey, N. T. J. (1975). *The Mathematical Theory of Infectious Diseases and Its Applications*. London: Charles Griffen.

Bailey, T. C., and Gatrell, A. C. (1995). *Interactive Spatial Data Analysis*. Essex, UK: Longman.

Bartholomew, D. J. (1982). *Stochastic Models for Social Processes*. New York: John Wiley and Sons.

Bass, F. M. (1969). A new product growth model for consumer durables. *Management Science, 15*(5), 215–227.

Beal, G. M., and Bohlen, J. M. (1955). *How Farm People Accept New Ideas*. Cooperative Extension Service Report 15: Ames, IA: Iowa State University.

Bearman, P. S., Jones, J., and Udry, J. R. (2000). *The National Longitudinal Study of Adolescent Health: Research Design*. Available from: http://www.cpcp.unc.edu/projects/addhealth/design.html.

Becker, M. H. (1970). Sociometric location and innovativeness: Reformulation and extension of the diffusion model. *American Sociological Review, 35*, 267–282.

Behrman, J. R., Kohler, H.-P. and Watkins, S., "Social Networks, HIV/AIDS and Risk Perceptions" (February 18, 2003). PIER Working Paper No. 03-007. http://ssrn.com/abstract=382844.

Boulay, M., and Valente, T. W. (in press). Dynamic sources of information and dissonance in the discussion networks of women in rural Nepal. *Journal of Health Communication*.

Broadhead, R. S., Hechathorn, D. D., Weakliem, D. L, Anthony, D. L., Madray, H., Mills, R. J., and Hughes, J. (1998). Harnessing peer networks as an instrument fo raids prevention: Results from a peer-driven intervention. *Public Health Reports, 113*(S1), 42–57.

Brown, L. (1981). *Innovation Diffusion: A New Perspective*. New York: Methuen.

Burt, R. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology, 92*, 1287–1335.

Carley, K. M. (2001). Learning and using new ideas: A socio-cognitive approach. In J. Casterline (Ed.), *Diffusion Processes and Fertility Transition: Selected Perspectives*. Washington, DC: National Academy Press.

Casterline, J. (2001). (Ed.) *Diffusion Processes and Fertility Transition: Selected Perspectives*. Washington, DC: National Academy Press.

Cliff, A., and Ord, J. K. (1981). *Spatial Processes: Models and Applications*. London: Pion.

Coleman, J. S., Katz, E., and Menzel, H. (1966). *Medical Innovation: A Diffusion Study*. New York: Bobbs Merrill.

Coughenour, C. M. (1965). The problem of reliability of adoption data in survey research. *Rural Sociology, 30*, 184–203.

Doreian, P., Teuter, K., and Wang, C. (1984). Network autocorrelation models: Some Monte Carlo results. *Sociological Methods and Research, 13*, 155–200.

Dow, M. (1986). Model selection procedures for network autocorrelated disturbance models. *Sociological Methods and Research, 14*, 403–422.

Dozier, D. M. 1977. *Communication Networks and the Role of Thresholds in the Adoption of Innovations*. Ph.D. Thesis, Stanford University, Stanford, CA.

Eliason, S. R. (1993). *Maximum Likelihood Estimation: Logic and Practice*. Newbury Park, CA: Sage.

Entwisle, B., Rindfuss, R. D., Guilkey, D. K., Chamratrithirong, A., Curran, S. R., and Sawangdee, Y. (1996). Community and contraceptive choice in rural Thailand: A case study of Nang Rong. *Demography*, *33*, 1–11.

Erbing, L., and Young, A. (1979). Individuals and social structure: Contextual effects as endogenous feedback. *Sociological Methods and Research*, *7*, 396–430.

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, *83*, 1420–1443.

Griffith, D. A., Layne, L. J., Ord, J. K., and Sone, A. (1999). *A Casebook for Spatial Statistical Data Analysis: A Compilation of Analyses of Different Thematic Data Sets*. New York: Oxford University Press.

Hägerstrand, T. (1967). *Innovation Diffusion as a Spatial Process*. (A. Pred, Trans.) Chicago: University of Chicago Press.

Hamblin, R. L., Jacobsen, R. B., and Miller, J. L. L. (1973). *A Mathematical Theory of Social Change*. New York: John Wiley and Sons.

Katz, E., Levine, M. L., and Hamilton, H. (1963). Traditions of research on the diffusion of innovation. *American Sociological Review*, *28*, 237–253.

Kincaid, D. L. (2000). Social networks, ideation, and contraceptive behavior in Bangladesh: A longitudinal analysis. *Social Science and Medicine*, *50*, 215–231.

Kohler, H. P. (1997). Learning in social networks and contraceptive choice. *Demography*, *34*, 369–383.

Latkin, C. (1998). Outreach in natural setting: The use of peer leaders for HIV prevention among injecting drug users' networks. *Public Health Reports*, *113*(S1), 151–159.

Lomas, J., Enkin, M., Anderson, G. M., Hanna, W. J., Vayda, E., and Singer, J. (1991). Opinion leaders vs. audit feedback to implement practice guidelines: Delivery after previous cesarean section. *Journal of American Medical Association*, *265*, 2202–2207.

Mahajan, V., and Peterson, R. A. (1985). *Models of Innovation Diffusion*. Newbury Park, CA: Sage.

Marsden, P. V. (1987). Core discussion networks of Americans. *American Sociological Review*, *52*, 122–131.

Marsden, P. V. (1990). Network data and measurement. *Annual Review of Sociology*, *16*, 435–463.

Marsden, P. V., and Podolny, J. (1990). Dynamic analysis of network diffusion processes. In J. Weesie and H. Flap (Eds.), *Social Networks Through Time*. Utrecht, The Netherlands: ISOR.

Montgomery, M. R., Agyeman, D., Aglotise, P., Hewett, P., and Kiros, G. (2001). *Social Networks and Contraceptive Dynamics in Southern Ghana*. Paper presented at the annual meeting of the Population Association of America, Washington, DC.

Montgomery, M. R., and Chung, W. (1999). Social networks and the diffusion of fertility control in the Republic of Korea. In R. Leete (Ed.), *Dynamics of Values in Fertility Change*. Oxford, UK: Oxford University Press.

Morris, M. (1993). Epidemiology and social networks: Modeling structured diffusion. *Sociological Methods and Research*, *22*, 99–126.

Myers, D. J. (2000). The diffusion of collective violence: Infectiousness, susceptibility, and mass media networks. *American Journal of Sociology*, *106*, 173–208.

Neiagus, A., Friedman, S. R., Kottiri, B. J., and Des Jarlais, D. C. (2001). HIV risk networks and HIV transmission among injecting drug users. *Evaluation and Program Planning*, *24*, 221–226.

Nischan, P., Ebeling, K., Thomas, D. B., and Hirsch, U. (1993). Comparison of recalled and validated oral contraceptive histories. *American Journal of Epidemiology*, *138*(9), 697–703.

Nyblom, J., Borgatti, S., Roslakka, J. and Salo, M. A. (2003). Statistical analysis of network data – an application to diffusion of innovation. *Social Networks*, *25*, 175–195.

Robertson, T. S. (1971). *Innovative Behavior and Communication*. New York: Holt, Rinehart and Winston.

Rogers, E. M. (2003). *Diffusion of Innovations* (5th ed.). New York: Free Press.

Rogers, E. M., and Kincaid, D. L. (1981). *Communication Networks: A New Paradigm for Research.* New York: Free Press.

Ryan, R., and Gross, N. (1943). The diffusion of hybrid seed corn in two Iowa communities. *Rural Sociology, 8*(1), 15–24.

Scott, J. (2000). *Network Analysis: A Handbook.* Newbury Park, CA: Sage.

Sikkema, K. J., Kelly, J. A., Winett, R. A., Solomon, L. J., Cargill, V. A., Roffman, R. A., et al. (2000). Outcomes of a randomized community-level HIV prevention intervention for women living in 19 low-income housing developments. *American Journal of Public Health, 90*, 57–63.

Soumerai, S. B., McLaughlin, T. J., Gurwitz, J. H., et al. (1998). Effect of local medical opinion leaders on quality of care for acute myocardial infarction: A randomized controlled trial. *Journal of the American Medical Association, 279*, 1358–1363.

Strang, D., and Tuma, N. B. (1993). Spatial and temporal heterogeneity in diffusion. *American Journal of Sociology, 99*, 614–639.

Teachman, J. D., and Hayward, M. D. (1993). Interpreting hazard rate models. *Sociological Methods and Research, 21*(3), 340–371.

Tuma, N. B., and Hannan, M. T. (1984). *Social Dynamics: Models and Methods.* New York: Academic Press.

Urberg, K. A., Degirmencioglu, S. M., and Pilgrim, C. (1997). Close friend and group influence on adolescent cigarette smoking and alcohol use. *Developmental Psychology, 33*, 834–844.

Valente, T. W. (1993). Diffusion of innovations and policy decision-making. *Journal of Communication, 43*(1), 30–41.

Valente, T. W. (1995). *Network Models of the Diffusion of Innovations.* Cresskill, NJ: Hampton Press.

Valente, T. W. (1996). Social network thresholds in the diffusion of innovations. *Social Networks, 18*, 69–89.

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

Valente, T. W., Hoffman, B. R., Ritt-Olson, A., Lichtman, K., and Johnson, C. A. (2003). The effects of a social network method for group assignment strategies on peer led tobacco prevention programs in schools. *American Journal of Public Health, 93*, 1837–1843.

Valente, T. W., and Rogers, E. M. (1995). The origins and development of the diffusion of innovations paradigm as an example of scientific growth. *Science Communication: An Interdisciplinary Social Science Journal, 16*(3), 238–269.

Valente, T. W., and Saba, W. (1998). Mass media and interpersonal influence in a reproductive health communication campaign in Bolivia. *Communication Research, 25*, 96–124.

Valente, T. W., and Saba, W. (2001). Campaign recognition and interpersonal communication as factors in contraceptive use in Bolivia. *Journal of Health Communication, 6*, 303–322.

Valente, T. W., and Vlahov, D. (2001). Selective risk taking among needle exchange participants in Baltimore: Implications for supplemental interventions. *American Journal of Public Health, 91*, 406–411.

Valente, T. W., Watkins, S., Jato, M. N., Van der Straten, A., and Tsitsol, L. M. (1997). Social network associations with contraceptive use among Cameroonian women in voluntary associations. *Social Science and Medicine, 45*, 677–687.

Van den Bulte, C., and Lillien, G. L. (1997). Bias and systematic change in the parameter estimates of macro-level diffusion models. *Marketing Science, 16*, 338–353.

Van den Bulte, C., and Lillien, G. L. (2001). Medical innovation revisited: Social contagion versus marketing effort. *American Journal of Sociology, 106*, 1409–1435.

Wasserman, S., and Faust, K. (1994). *Social Networks Analysis: Methods and Applications.* Cambridge, UK: Cambridge University Press.

White, K., and Watkins, S. C. (2000). Accuracy, stability and reciprocity in informal conversational networks in rural Kenya. *Social Networks, 22*, 337–356.

Models and Methods in Social Network Analysis

Edited by

PETER J. CARRINGTON
University of Waterloo

JOHN SCOTT
University of Essex

STANLEY WASSERMAN
Indiana University

2005

CAMBRIDGE
UNIVERSITY PRESS