

The Specification of Weight Structures in Network Autocorrelation Models of Social Influence

Roger Th.A.J. Leenders

SOM-theme B Innovation and interaction

Abstract

Many physical and social phenomena are embedded within networks of interdependencies, the so-called 'context' of these phenomena. In network analysis, this type of process is typically modeled as a network autocorrelation model. Parameter estimates and inferences based on autocorrelation models, hinge upon the chosen specification of weight matrix W , the elements of which represent the influence pattern present in the network. In this paper I discuss how social influence processes can be incorporated in the specification of W . Theories of social influence center around 'communication' and 'comparison'; it is discussed how these can be operationalized in a network analysis context. Starting from that, a series of operationalizations of W is discussed. Finally, statistical tests are presented that allow an analyst to test various specifications against one another or pick the best fitting model from a set of models.

(also downloadable) in electronic version: <http://som.rug.nl>

The Specification of Weight Structures in Network Autocorrelation Models of Social Influence

1. Introduction

Many physical and social phenomena are embedded within networks of interdependencies, the so-called ‘context’ of these phenomena. In determining their opinions and behavior, in accordance with the constraints and possibilities imposed by the network, actors are assumed to be responsive to the contextual cues provided by the opinions and behavior of significant others. By appropriately taking into account the opinions and behaviors displayed by their significant others, actors thus establish their own behavior. In the literature, this influence process has been labeled ‘contagion’ (cf. Leenders, 1995, 1997).

Of course, opinions and behavior are not solely determined by those of others (interaction), but also by reaction to various other constraints and opportunities granted by the social system (local effects). In sociology, this type of process is typically modeled as an autocorrelation model¹ of the form

$$y = \rho W y + X \beta + \varepsilon$$

or

$$y = X \beta + \varepsilon, \quad \varepsilon = \rho W \varepsilon + \nu.$$

Parameter estimates and inferences based on such autocorrelation models hinge upon the chosen specification of weight matrix W . This matrix represents the influence process assumed to be present in the network and can be operationalized in many different ways. W is supposed to represent the theory a researcher has about the structure of the influence processes in the network. Since any conclusion drawn on the

1. Autocorrelation of either a variable (or error term) is the situation where the observations of variables (or error terms) for different actors are not independent over time, through space, or across a network.

basis of autocorrelation models is conditional upon the specification of W , the scarcity of attention and justification researchers pay to the chosen operationalization of W is striking and alarming. This is especially so, since different specifications of W typically lead to different empirical results.

In this paper, I attempt to provide researchers with an understanding of various specifications of influence structure W . When specifying W , four steps need to be taken. First, the researcher has to decide whether social influence occurs through the autocorrelation of the dependent variable, through the autocorrelation of disturbances, or through a combination of the two. In Section 2 I describe the substantive difference between these two approaches and present the models that accompany them. Second, the researcher decides on which mechanism governs social influence: communication or comparison (Section 4). Third, a choice is made about which alters exert influence on ego and which alters don't. In other words, this is a choice about which elements of W are zero and which are non-zero. In the fourth step, for the non-zero elements of W , the researcher determines how much influence is exerted. These last two steps are discussed in Sections 5 and 6. Finally, in Section 7 I present some statistical tests that can help a researcher make a choice between rival models.

2. Network autocorrelation models

Let y be a $(g \times 1)$ -vector of values of an endogenous variable for g actors making up a network and let X denote a $(g \times k)$ -matrix of values for the g actors on k covariates (including an optional row of 1's for the constant term). I will refer to y as the *dependent* variable. In determining his opinion, ego takes into account the opinions of his significant others. These significant others make up i 's frame of reference. The opinion of some alters carry more weight to ego than those of others. This is denoted by 'nearness,' referring to the extent to which alter's opinions and beliefs are emulated by ego. In mathematical terms, y_i is related to a weighted combination of the y_j of other actors. If these weights are given in the $(g \times g)$ -matrix W , then y is related to Wy . An entry w_{ij} of W denotes the influence actor² j has on actor i ; the nearer j is to i , the larger w_{ij} . So y_i is written as³

$$y_i = \rho w_{i1}y_1 + \rho w_{i2}y_2 + \dots + \rho w_{ig}y_g + \varepsilon_i.$$

The strength of the influence if i 's alters on i is thus determined by the weights in the i th row of W . In matrix notation:

$$y = \rho Wy + \varepsilon \quad (1)$$

where $E(\varepsilon) = 0, E(\varepsilon\varepsilon') = \sigma^2I$ and ρ is a scalar (Anselin, 1982; Doreian, 1981; Griffith, 1976; Haining, 1978; Mead, 1967; Whittle, 1954). This model represents contagion in a straightforward fashion: ego's opinion is a weighted version of the opinions of his alters.

In most cases both interaction and local effects play a role in the influence process. For example, two organizations may compete for market share, thus revealing between-organization effects (interaction); but they will also react to the general availability of demand, revealing local effects to exogenous factors. Similarly, ego's voting behavior may be influenced by discussing matters with friends (interaction), but will often also depend on ego's status, income, education, and so forth (local effects). Here, (1) is of limited utility and is extended by including covariates:

$$y = \rho Wy + X\beta + \varepsilon \quad (2)$$

where it is assumed that the error terms are normally distributed with zero means and equal variances; so $\varepsilon \sim N(0, \sigma^2I)$. The difference between interaction and local effects is reflected by the difference between the autocorrelation part in the equations (containing W) and the exogenous part (containing $X\beta$). Ord (1975) terms model (2) the regressive-autoregressive model; Doreian (1989b) calls it the network effects model. The model has been studied by, among others, Doreian (1981; Doreian et al., 1984). When $\rho = 0$ the model reduces to the standard regression equation $y = X\beta + \varepsilon$, while for $\beta = 0$ the purely spatial model (1) is obtained.

-
2. In some parts of the literature, w_{ij} is used to denote the influence of i on j . In order to preclude confusion, I will systematically transform W , so that all w_{ij} 's in this paper can unambiguously be interpreted as j 's influence on i . Also, see Section 6.
 3. In social network analysis, standard usage is to exclude the relation from an actor to himself (loops), thus $w_{ii}=0$. However, the inclusion of loops can be theoretically meaningful. The effect of actor i on y_i should then be included.

An alternative way to incorporate both local and interaction effects is the network disturbances model:

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \rho W\varepsilon + \nu \quad \nu \sim N(0, \sigma^2 I) \end{aligned} \quad (3)$$

This model has been studied by, among others, Doreian (1980), Dow et al. (1982), Loftin and Ward (1983), Ord (1975), and White et al. (1981). Due to their similarity to time series models, (2) has also been labeled a SAR (spatial autoregressive) model, and (3) an SMA (spatial moving average) model (e.g., Mur 1999).

Sometimes one may expect two regimes W to be present. Such a model is (Doreian, 1989a,b)

$$y = \rho_1 W_1 y + \rho_2 W_2 y + X\beta + \varepsilon, \varepsilon \sim N(0, \sigma^2 I) .$$

Analogously this can be done in the case of disturbance autocorrelation:

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \rho_1 W_1 \varepsilon + \rho_2 W_2 \varepsilon + \nu, \nu \sim N(0, \sigma^2 I). \end{aligned}$$

See Brandsma and Ketellapper (1979) and Dow (1984) for a discussion of this model. A natural generalization combining network effects and network disturbances is (Doreian, 1982; Rietveld and Wintershaven, 1998):

$$\begin{aligned} y &= \rho_1 W_1 y + X\beta + \varepsilon \\ \varepsilon &= \rho_2 W_2 \varepsilon + \nu \quad \nu \sim N(0, \sigma^2 I) . \end{aligned} \quad (4)$$

Of course, $W_1 = W_2$ is allowed. Anselin (1988: 34-35) studies (4) as a general family of autocorrelation models.⁴ An overview of some network autocorrelation models is given by Doreian (1989b) and Leenders (1995, 1997). An overview and discussion of related network models is given in Marsden and Friedkin (1993).

The substantive choice between modeling contagion through either autocorrelating the dependent term or the disturbance term reflects a theoretical difference of how

4. In fact, Anselin's approach is more general, as he allows $\nu \sim N(0, \Omega)$, where Ω may be heteroskedastic and is allowed to depend on covariates.

contagion is supposed to take place.⁵ The opinion an actor would display in the absence of social influence is an actor's *intrinsic opinion* (Leenders, 1995). In both (2) and (3) an actor's intrinsic opinion is represented by $X\beta$, when ρ is equal to zero. In specification (2), ego builds his own opinion both on his intrinsic opinion and on the opinion of his alters. For instance, ego's political preference would be a function of both ego's socio-economic status, education, and income (intrinsic opinion) and the political views expressed by his family, neighbors, and colleagues (contagion). Together, these effects then simultaneously determine ego's political stance. Alternatively, there may be situations in which ego initially only bases his stance on his intrinsic opinion; his status, education, and income initially determine ego's political views. As ego then observes his significant others to deviate from their intrinsic opinions, ego decides to adapt to their deviation. For instance, ego may find his alters to be more on the left-hand or right-hand side of the political spectrum than their status, education, and income would prescribe. With his alters leaning more towards the left (right) than expected, ego also is inclined to alter his view towards the left (right). In this case, ego does not take the absolute value of the opinion of his alters as a benchmark, but the deviation from their supposed (intrinsic) opinion. The mechanism of adaptation to deviation from the intrinsic opinion, rather than to opinions themselves is, in a statistical context, reflected by autocorrelation of residuals. The residuals ε_i capture latent forces that push an actor's opinion away from his intrinsic standpoint; it is this type of process that is captured by specification (3).

A model of contagion of deviations is appropriate in at least three situations. First, the deviation may represent insecurity, uncertainty, or risk. When uncertainty is high, making it difficult for actors to assess the 'right' opinion or 'right' behavior, actors are expected to watch others and observe how others deal with this uncertainty and mimic their adaptive action. Similarly, if actors cannot observe all variables thought to be relevant to the formation of their opinion, ego will adapt to this lack of information by considering how his significant others adapt to this lack.⁶ Finally, when actors do know all the relevant factors, but are not sure as to the relative and behavioral importance of

5. Anselin and Bera (1998: 247) discuss the substantive basis for the choice between the two approaches in a geographical setting.

them, they can take the educated guesses of significant others into account.

Autocorrelation model (1) is highly stylized, representing an extreme form of contagion. It supposes behavior to be fully dependent on interaction, void of any effects intrinsic to the actors in the network. It can caricaturingly be termed a 'lemming' model. It is hard to think of situations in which this model would be substantively appropriate. One case would be when modeling the behavior of members of a sect, where people are stripped of their individuality and are forced into group behavior. Alternatively, this model could be appropriate for modeling situations so unfamiliar to actors, that they can not (or will or dare not) rely on their own experiences, background, or routines for behavioral direction, but fully rely on the behavior of others in determining their own best behavior.

3. Why the specification of W is important

When working with any autocorrelation model, the chosen specification of W is of vital importance. The first reason network autocorrelation models are used is for estimating ρ or β . Many researches aim at estimating β in the situation of (possibly) interdependent variables. In this fashion, the autocorrelation model is used to remove bias due to the interdependence of units from OLS estimates of β . Various procedures exist, an overview of which can be found in Anselin (1988) and Leenders (1995). The estimation procedure with the best overall performance undoubtedly is the Maximum Likelihood estimation procedure. This procedure has been the preferred method of estimation for over two decades and is the default procedure in the available commercial software packages for network (or spatial) autocorrelation models. The Maximum Likelihood (ML) estimators (as do almost all alternative estimators) explicitly incorporate weight matrix W . Change W , and get different estimates. The difference can be substantial.

A second reason autocorrelation models are employed is for testing purposes. One may want to test the statistical significance of particular parameters or simply test for

6. Likewise, from a statistical point of view, when important variables are not included in X , disturbance term autocorrelation may arise. Also, autocorrelation in disturbance terms may also arise when linear relationships between the variables are assumed, but when in fact they are nonlinear.

the presence of autocorrelation. In the latter case, there are two classes of tests. The first class of tests is used when the underlying distribution of the autocorrelated phenomenon is unknown, or when observations are measured on a nominal or ordinal scale. This class consists of statistics such as Moran's I , Geary's c , and the join-count statistic. These do not involve weight matrix W at all, and are therefore unaffected by any specific form chosen for W . The downside to these tests is that, although their null hypothesis is that of absence of autocorrelation, their alternative hypothesis is left vague. It is therefore unclear what the rejection of the null hypothesis really means. The second class consists of tests that involve distributional information and includes, among others, the Lagrange Multiplier (LM) test, the Likelihood Ratio test (LR), the Wald test (Wald), and Moran's I as reformulated by Cliff and Ord (1973; Tiefelsdorf, 1998). In modern literature on autocorrelation models, the LM, LR, and Wald tests are the most commonly used. Since these tests explicitly incorporate W , any inference is conditional upon the chosen W . A parameter may lose or attain statistical significance when inference is based on an alternative specification of W , while still using the same dataset. Change W and draw a different conclusion.

The last reason to use an autocorrelation model is to test theories of social influence in a particular dataset. In this situation, the specification of W should follow naturally from the theory at hand. Change one's theory, change W .

In conclusion, regardless of the purpose served by applying a network autocorrelation model, virtually any conclusion depends on the specification of W . It is therefore of vital importance to have justification for the W applied in the research.

4. Theories of social influence

Social influence occurs when an actor adapts his behavior, attitude, or belief, to the behaviors, attitudes, or beliefs of other actors in the social system. It does not matter whether alter's influence on ego's behavior is intentional or unintentional and is not restricted to direct communication. A precondition for social influence to occur is the availability to ego of information about the attitudes or behavior of other actors. In this paper, social influence is viewed as a dyadic process: ego adapts his behavior to that of alter, leading them to behave similarly.⁷ In the literature, different terms are used to

describe the same thing. In particular, the term *contagion* is often used to describe the social influence processes dealt with in this paper (e.g., Leenders, 1995, 1997). Therefore, in the remainder of this paper, I will use the terms contagion, social influence, and influence interchangeably, each is taken to mean the same thing.

Sociological literature contains many different theories of social influence. Most of these are couched in terms of the idea that the attitudes and opinions of significant others influence the way in which a person comes to view a situation. The opinions of alters are seen as an appropriate standard against which ego evaluates his own opinion. In other words, when forming his own opinion, ego uses other actors as his *frame of reference* and takes their opinions into account. In the remainder of the paper, I will use both the terms *alter* or *significant other* to refer to a member of an actor's frame of reference.

Within the realm of social influence theory, the notion of a frame of reference has crystalized around two processes (Figure 1).

- Communication: actors use actors with whom they are directly tied as their frame of reference.
- Comparison: actors use actors they feel similar to as their frame of reference.

** Figure 1**

4.1. *Communication*

Communication refers to social influence through direct contact between ego and alter. The more frequent and vivid the communication between ego and alter, the more likely it is that ego will adopt alter's ideas and beliefs. Through discussing matters with

-
7. Consequently, if ego and alter express similar opinions after both watching the same television program, this is not a result of social influence between them. However, if alter's opinion changes after watching a TV program, this could cause alter to adhere to and express an opinion that deviates from his intrinsic opinion. The network disturbances model then allows *i* to adapt his opinion as well. Although the direct influence of the TV show on *j* does not fall within my definition of a network social influence process, the adaptation of *i*'s opinion based on *j*'s adaptation does.

alter, ego comes to an understanding of an issue and adds new information to his own.

Homans' work (Homans, 1950, 1961) provides a theoretical foundation for contagion through communication. Classical early empirical work was performed by Festinger (Festinger et al., 1950; Festinger and Kelly, 1951) and Lazarsfeld (Lazarsfeld et al., 1948; Berelson et al., 1954). Lazarsfeld et al. (1948) for instance argue that people rely on personal contacts to help them select relevant arguments in political affairs. Ego trusts the judgement and evaluation of those who are respected around him. Personal obligations and trust are more powerful influences than radio or newspapers. Berelson et al. (1954) show that political preferences of friends and co-workers strongly determine ego's preference and that these alters also affect the strength of conviction with which actor's vote preference is held. Accordingly, they show that young voters have a very strong tendency to vote like their fathers. Baerveldt and Snijders (1994), studying network effects on cultural behavior, find petty crime offences amongst pupils to be correlated with the number of offences committed by their friends. Most studies of social influence assume communication to be the underlying process.

4.2. *Comparison*

The other process of contagion is social *comparison*. In searching for a social identity, ego ascribes to himself those characteristics or feelings that alters would ascribe to him if they would have the same information at their disposal (cf. Bem, 1972; Tajfel, 1972). Putting it differently, ego compares himself to those alters whom he considers similar to him in relevant respects, asking himself 'what would another person do if he were in my shoes?'. Ego perceives (or assesses) alter's behavior and assumes that behavior to be the 'correct' behavior for 'a-person-like-me' or for 'a-person-in-a-position-like-mine.' Burt (1987) argues that comparison is triggered if actors are in competition with one another. By comparison they evaluate their relative adequacy.⁸

Role playing and imitation are similar to comparison; comparison establishes a role-playing frame as alters are imitated or roles are emulated.

4.3. *Attitudes vs. behavior*

In studies on organizations, and in political science in particular, much interest exists in contagion processes. The argument is that if contagion is strong within a group, the likelihood of similar behavior increases, leading to an increase in the group's power. However, there is an important difference between similarity of beliefs and interests and similarity of behavior (Mizruchi, 1989, 1990; Wickes, 1969). Behavior is not solely determined by a set of attitudes and beliefs, but also by restrictions with which the actor is confronted. A change in some of the attitudes and beliefs does not automatically lead to changes in behavior. Similarity in beliefs, therefore, does not necessarily lead to similarity in behavior. Moreover, actors with different beliefs might well behave similarly (Merton, 1936).

In the communication approach to contagion, information is exchanged about an issue at hand, uncertainty is expressed, past experiences are shared, and actors learn from each others' mistakes, leading to a unison of opinions, attitudes, and beliefs, but not necessarily to conformity in behavior. Comparison, on the other hand, which may take place among non-adjacent actors, is explicitly based on role playing and the copying of behavior. Moreover, actors who are not directly tied to one another can rely only on observed behavior, as they cannot discuss with alter which attitudes underly his behavior.

In short, communication yields similarity of beliefs, but not necessarily of behavior, whereas comparison leads to similarity in behavior, but not necessarily in underlying beliefs.

The consequence of this difference is important because similarity of behavior are easily observed by a researcher, but similarity of beliefs and attitudes are not. The methodological problems related to this will not further be discussed in this paper.

8. Burt's argument is that comparison is used by actors in order to gain strategic advantage. By comparing himself to alter, ego can invent and adopt innovations that would make ego more attractive than alter as the object or source of relations (Burt, 1987: 1291). Naturally, this only makes sense for actors who occupy similar positions in the social system.

5. Operationalizing social influence

5.1. Operationalizing communication

The common approach to operationalizing a communication process in social network analysis is through *cohesion*. Cohesion incorporates the number, length, and strength of the paths connecting actors, leading to concepts such as cliques (Mokken, 1979), *k*-plexes (Seidman and Foster, 1978), and *k*-cores (Doreian and Woodard, 1994, Seidman, 1983). It is then tested whether actors who belong to the same subgroup, clique, *k*-plex, or *k*-core are more alike than actors in different groups. If they are, this is attributed to communication. However, when $k \geq 2$ there is no guarantee that actors, belonging to the same cohesive subgroup, are in communication with each other. In this paper I therefore restrict cohesion to actors who are directly tied. Cohesion then assumes actors who are directly linked to be(come) more similar than actors who are not linked directly.

5.2. Operationalizing comparison

Comparison is most often operationalized by the concept of equivalence. Equivalent actors are similarly embedded in the network. The most widespread conceptualization of equivalence in autocorrelation models is *structural equivalence*⁹ (Lorrain and White, 1971). Actors are structurally equivalent if they have exactly the same ties to and from all actors. Actors need not be directly tied in order to be structurally equivalent. They may never communicate and may not even know of each other's existence.

In practice, actors are seldom exactly structurally equivalent and the equivalence criterion is relaxed to measure the *extent* to which actors are structurally equivalent. A common measure is as follows. Construct vector \vec{i} by stacking the *i*th row and column

9. There are many alternative types of equivalence, some of them substantively being better suited to capture processes of comparison. However, as I discuss elsewhere (Leenders, 1995), structural equivalence is the most practical measure for the purpose. Moreover, with respect to capturing influence processes, structural equivalence measures correlate highly with alternative conceptualizations of equivalence. A related discussion is found in a highly recommended paper by Borgatti and Everett (1992).

of adjacency matrix A , and \vec{j} by stacking the j th row and column of A . Euclidean distance $d_{ij} = \|\vec{i} - \vec{j}\| = [(\vec{i} - \vec{j})'(\vec{i} - \vec{j})]^{1/2}$ is a measure of structural equivalence, equaling 0 for exactly equivalent actors and $\sqrt{2g}$ for completely nonequivalent actors (g being the number of actors in the network). This number is then standardized by dividing by $\sqrt{2g}$. For use in network autocorrelation models, I propose a slightly different approach. Define \vec{i} and \vec{j} as before and define $\tilde{j} = 1 - \vec{j}$, where 1 is a $(2g \times 1)$ vector of ones. The Euclidean distance $\|\vec{i} - \tilde{j}\|$ now is a measure of structural similarity or proximity rather than distance. Again this measure is standardized by dividing by $\sqrt{2g}$. A value of 0 represents exact non-equivalence, a value of 1 exact structural equivalence. I will come back to this in 5.5.

Since the comparison argument states that actors involve in copying each other's behavior, actors should be able to observe each other. One can not derive from a network structure alone whether actors, who are not directly tied, know or can observe each other. It is likely that the shorter the path between two actors, the higher the probability that they know of each other and can observe each other's behavior. It therefore makes sense to restrict equivalence to actors who are proximate, say at a sociometric distance less than or equal to three.¹⁰

In analyses of empirical data, cohesion and structural equivalence often yield similar results since, in order to establish whether actors are structurally equivalent, one only needs to consider to whom they are directly tied. A change in network structure at a sociometric distance of at least two from both actors does not affect their structural equivalence. Thus, two actors cannot be structurally equivalent if they have a sociometric distance of more than two between them.¹¹ Thus, structural equivalence and cohesion are often strongly correlated.

10. Sociometric distance 3 (friend-of-a-friend-of-a-friend) intuitively seems a useful threshold for approximating awareness. As a practical test, the reader might consider of how many friends-of-a-friend-of-a-friend-of-a-friend (sociometric distance 4) (s)he can really assess opinions, attitudes or behavior. A researcher may also consider setting the threshold at distance 2 (friend-of-a-friend).

11. There is one exception to this rule: isolates (nodes that have no ties to and from other actors) are also mutually structurally equivalent.

5.3. *Communication or comparison?*

Although communication and comparison reflect theoretically fundamentally different mechanisms of social influence, they are not easily separable empirically. They are different, but not disjoint. The communication argument relates similarity to the direct ties between actors: if adjacent actors are observed to be(come) alike, this is attributed to communication. The comparison argument, on the other hand, relates similarity to the fact that actors are correspondingly embedded in the social structure: if equivalent actors are observed to be(come) alike, this is attributed to comparison. However, contrasting the similarity of adjacent actors with the similarity of structurally equivalent pairs does not establish whether similarity is a result of either communication or comparison. Actors who are directly tied can become alike by both communication and comparison. Besides this, they can become alike by communicating with a mutual friend (indirect communication) or by comparing themselves to the same alter (indirect comparison).¹² It is usually not possible to empirically distinguish among these four processes (communication, comparison, indirect communication, indirect comparison) when it is only observed that adjacent actors are or become alike. Even though in some cases it may be possible to theoretically rule out one or two of them, it will hardly ever be possible to discard all but one.

Actors who are equivalently located in the network may become alike through comparison, indirect communication or indirect comparison. Equivalent actors can be adjacent and become alike through communication. Again, these effects can usually not be distinguished.

The problem of distinguishing between communication and comparison has led several researchers to arbitrarily discard either communication or comparison as a contagion mechanism and strictly adhere to the other.¹³ It is more interesting, however, to ask whether it is possible to determine in advance whether communication or

12. Indirect communication and indirect comparison are not contagion, but a consequence of contagion. However, it appears to be impossible to formulate practical measures of contagion that exclude all indirect flows of contagion.

13. Indirect communication and indirect comparison are seldom addressed.

comparison is driving contagion. Unfortunately, in the literature not much is known about the relative prominence of either of the two processes in various situations. Burt and Doreian (1982) find support for both the communication and comparison arguments in separate analyses of the same dataset. Galaskiewicz and Wasserman (1989) also find proof of comparison, as they find that organizational decision makers mimic the behavior of other decision makers. However, they also report that managers are especially likely to mimic behavior of the organizations they are directly linked to. Comparison or communication?

An interesting empirical debate centers around the diffusion of innovations study of Coleman et al. (1957, 1966) who show that communication was driving the adoption of a new drug among physicians in the Midwest. Burt (1987) reanalyzes the data and concludes comparison rather than communication to be the source of contagion. In turn, Marsden and Podolny (1990) find hardly any evidence of contagion at all, neither communication nor comparison. Finally, Strang and Tuma (1993) show effects of *both* communication and comparison! This sequence of contradictory results on the same dataset certainly highlights that the empirical distinction between the two processes is not straightforward and that they may be strongly interrelated. Although, to my knowledge, this is the only dataset that has evoked such major controversy, it is not likely an exceptional one.

Sometimes an attempt is made to test communication against comparison by testing for equivalence after first controlling for cohesion, the idea being that cohesion has already picked up communication effects, so equivalence will only estimate comparison. This, however, is not correct. Cohesion will still summarize direct and indirect communication and comparison effects between adjacent actors, and equivalence will then summarize comparison and indirect communication among non-adjacent actors. In some cases, it may be possible to argue on theoretical grounds whether communication or comparison is driving contagion, but even then it will most often still be impossible to empirically test one against the other.

5.4. *An alternative distinction: adjacency versus non-adjacency*

Only in rare cases is it possible to strictly (theoretically or empirically) distinguish

between the two types of processes, attempting to do so in other cases only yields spurious results. Therefore, I will not make the empirical distinction between communication and comparison, but make a distinction between contagion among directly tied actors and contagion among actors who are not directly tied (Figure 2).

**** Figure 2****

Contagion among directly tied actors picks up both direct and indirect communication and comparison effects. Contagion among actors who are not directly tied summarizes direct and indirect comparison and indirect communication effects. The greater the sociometric distance between the actors, the smaller the effect of indirect communication and comparison, thus the stronger the direct comparison component.

The distinction between contagion through direct and indirect ties does not resolve the issue of communication versus comparison entirely. It does, however, isolate the effects of contagion through direct communication,¹⁴ which is one of the two major theoretical explanations of contagion. Table 1 summarizes the argument.

**** Table 1 ****

Contagion among adjacent actors is simply operationalized by investigating whether adjacent actors tend to share similar opinions or behavior. I propose two different operationalizations for contagion among non-adjacent actors. In the first, equivalence between untied actors is considered; the second considers the number of paths of various lengths between actors.

5.5. *Effects between non-adjacent actors*

Structural equivalence between non-adjacent actors can be defined following Section 5.2. Define equivalence proximity ep_{ij} by

14. Note, however, that it will never be possible, using only relational data, to distinguish between influence through direct communication and influence by comparison between adjacent actors.

$$ep_{ij} = [(\tilde{i} - \tilde{j})'(\tilde{i} - \tilde{j})]^{1/2}(1 - a_{ij})/\sqrt{2g} . \quad (5)$$

Non-adjacent actors who are exactly equivalent have proximity 1, other non-adjacent pairs of actors have equivalence proximity between zero and one. Adjacent actors always have proximity zero. The reasoning behind the inclusion of $(1 - a_{ij})$ in (5), is that this term prevents equivalence being blended with adjacency. When equivalence proximate actors are found to be similar in opinions, this is not an effect of direct communication. Note that (5) takes into account the directionality of the tie, which is lost in the standard formulae for structural equivalence. This means that if $a_{ij} = 1$, but $a_{ji} = 0$, then the influence of i on j by direct communication is filtered out, but j can still influence i by equivalence. Definition (5) shifts the focus to the zero relations in the network.

An alternative approach starts from the paths between i and j . When i and j are not directly tied but many short paths exist between them, there are many possibilities for indirect communication. When there are relatively many long shortest paths between actors, comparison is predominant. Again, i and j should be able to observe one another, I therefore suggest considering only non-adjacent actors i and j who are at a sociometric distance less than or equal to two or three. With $a_{ii} = 0$, the number of (shortest) paths¹⁵ of length two from i to j is given by entry (i, j) of A^2 . Similarly, the number of (shortest) paths of length three from i to j is given by A^3 . The number of (shortest) paths of length two or three is found by simply adding A^2 and A^3 . After having constructed the matrix with shortest path lengths, the matrix is multiplied element-wise by $(1 - a_{ij})$, in order to set the path length between adjacent actors equal to zero.

It is also possible to calculate all paths between actors i and j , by determining entry (i, j) of A^2 and A^3 after setting $a_{ij} = 0$ before multiplication. This calculation needs to be performed separately for each pair of adjacent actors. In this fashion, adjacent

15. Mathematically, the k th power of A gives k -sequences and not k -paths. However, the paths between non-adjacent actors of length smaller than or equal to three are exactly equal to the corresponding 2-sequences and 3-sequences. Also, for these lengths the paths are equal to the shortest paths.

actors are allowed to influence each other through indirect paths, but the influence flowing through paths involving a direct link between them is filtered out.

A situation in which these approaches are especially useful is when both a factor based on adjacency and a factor based on equivalence proximity/shortest paths is incorporated in a model of attitudinal actor similarity. Equivalence proximity will not absorb any of the social influence based on direct adjacency.

6. Operationalization of weight matrix W

Social influence enters network autocorrelation models through the weight matrix W , also called the *structure matrix*. Entry w_{ij} represents the extent to which y_i is dependent on y_j , thus to what extent actor j influences i .

All operationalizations that follow below are ultimately based upon adjacency matrix A , with $a_{ij} = 1$ meaning that j influences i . Different specifications of W can represent different theoretical mechanisms of social influence. In addition, also from a technical point of view it is important how W is specified since estimates of and tests for the various parameters of the models all depend on the specification of W and their properties are conditional on W .

With regard to an operationalization of W , two components play a role: the choice for an operationalization of nearness and the choice for a particular normalization which, given a definition of nearness, allocates influence over the network. In other words, nearness defines *which alters* constitute ego's frame of reference (zero and non-zero cells in W), whereas the chosen normalization determines how social influence is *allocated among* these alters. For instance, cohesion suggests that actors are influenced by adjacent actors, normalization then decreases the individual strength of influence with the number of influencers. I will first discuss the difference between row and column normalization. Then the focus will advance to discussing ways to operationalize nearness, since many of these were developed with a particular normalization already in mind.

6.1. Row versus column normalization

The influence structure in a network is represented by a weight matrix where each row displays the influence exerted on an actor and the column displays the influence

exerted by an actor. The normalizations discussed below do not pose restrictions on the value of the diagonal of A .

Row normalization of adjacency matrix A gives

$$w_{ij}^{[1]} = \frac{a_{ij}}{a_i}$$

with $a_i = \sum a_{ij}$, the i th row sum of A . Thus a_i denotes the number of actors to whom i has a tie. With row normalization the same weight is attached to every outgoing tie of i , proportional to the outdegree of i . If actor i has three outgoing ties, each of his alters will have weight $1/3$. An actor with only one outgoing tie will be fully influenced by this one alter.

A straightforward, but less common variation to $w_{ij}^{[1]}$ is the column normalization of A :

$$w_{ij}^{[2]} = \frac{a_{ij}}{a_j}$$

with a_j representing the j th column sum of A . The strength of influence actor j has over actor i now depends on the number of actors influenced by j , instead of on the amount of actors influencing i . In row normalization every actor undergoes the same total amount of influence from all actors: *accepted* influence of j by i decreases with the number of actors influencing i . In column normalization every actor exerts the same total amount of influence on all actors: *exerted* influence of j on i decreases with the number of actors j influences. The important differences between row- and column-normalization are presented in Table 2.

** Table 2 **

Note that the weight matrix resulting from either row or column normalization is likely to become asymmetric since $\sum_j w_{ij} \neq \sum_i w_{ji}$, even though the original matrix may have been symmetric.

6.2. Influence between adjacent actors: Cohesion

A non-normalized weight matrix alternative to $w_{ij}^{[1]}$ was introduced by French (1956) as

$$w_{ij}^{[3]} = \frac{a_{ij}}{a_{i.} + 1}.$$

By adding a one to the numerator the effect an actor has on his own position, termed *resistance*, is taken into account. When four actors try to influence i , each of them only has weight 1/5.

So far I have not made a distinction in the power or abilities of actors in influencing others. These abilities, indicated by the resources available to actors, can be incorporated into a weight matrix by allowing actors to have different resources (Hoede 1979):

$$w_{ij}^{[4]} = \frac{r_j a_{ij}}{r_i + \sum_{\substack{k=1 \\ k \neq i}}^g r_k a_{ik}}$$

If every actor has the same resources, $r_1 = r_2 = \dots = r_g = r$, the Hoede matrix equals French's weight matrix. One difficulty with $w_{ij}^{[4]}$ is the assumption that the resources of the actors are known a priori, requiring the researcher to find a way of measuring/postulating the values of r_i . Applying $w_{ij}^{[4]}$ with $a_{ii} = 1$ yields the control matrix as used by Stokman and Van den Bos (1992). An alternative to $w_{ij}^{[4]}$ is

$$w_{ij}^{[5]} = \frac{r_j a_{ij}}{r_i + \sum_{\substack{k=1 \\ k \neq i}}^g r_k a_{kj}}.$$

6.3. *Influence between non-adjacent actors: Equivalence*

Two actors are structurally equivalent to the extent that their structural distance is small. These distances, denoted here by d_{ij} , can be used to construct W matrices. An example of this is given in Burt and Doreian (1982: 117, 125) by

$$w_{ij}^{[6]} = \frac{l_{ji}}{1 - l_{ii}}$$

where

$$l_{ji} = \frac{(d_j - d_{ij})^v}{\sum_{q=1}^g (d_j - d_{qj})^v}$$

with v a constant,¹⁶ w_{ii} is set to zero and $d_j = \max_i(d_{ij})$. Burt and Doreian employ $w_{ij}^{[6]}$ to model how interests expressed by scientists in a scientific journal are converted into a journal norm. The journal norm is defined as the level of interest in the journal expected of scientists because of their structural position in ‘invisible colleges.’ Weight $w_{ij}^{[6]}$ represents the extent to which actor j is the only other member of the college that i perceives to be his structural peer (Burt and Doreian, 1982: 125).

In 5.5, I proposed an alternative approach to equivalence, termed equivalence proximity. The operationalization of equivalence proximity excludes effects of direct communication, by assigning zero proximity to adjacent actors. This measure can be transformed into a weight matrix by normalizing the matrix with equivalence proximities. Since the matrix is asymmetric for an asymmetric A , the device of asymmetrizing it by $w_{ij}^{[6]}$ is unnecessary.

6.4. *Influence between non-adjacent actors: indirect paths*

16. The magnitude of \mathbf{V} represents the extent to which ego is conservative in adjusting his interests to those held by his significant others. Values of \mathbf{V} much larger than 1 indicate that ego takes into account only the interests of his closest alters. Values near zero indicate that ego’s interests are affected by nearly everyone in the network. See Stevens (1957, 1962) for the theory behind \mathbf{V} .

A second route into modeling influence between non-adjacent actors starts from the indirect paths between them. Contagion is then considered to flow from alter to ego through ties with third actors. In 5.5 two ways of measuring these paths were discussed. In the first, the adjacency matrix is squared or cubed and entries for adjacent actors are set to zero afterwards. In the second, entries (i, j) of A^2 and A^3 for adjacent actors are determined separately after setting $a_{ij} = 0$ (for only this particular pair (i, j)) in the original adjacency matrix. After calculating A^2 , A^3 , and $A^2 + A^3$, these can be transformed into weight matrices employing any of the transformations $w_{ij}^{[1]}$ - $w_{ij}^{[5]}$.

6.5. *Other specifications*

Numerous other specifications of nearness exist, many of them emanating from the geography literature. A straightforward measure is the actual geographic distance between two actors, e.g., measured in miles. For a taste of geographic approaches, try Cliff and Ord (1973, 1981), Loftin and Ward (1983), and Bavaud (1998). Unfortunately, many of the specifications in the geography literature are hard to translate into a social network context. Other specifications, within the realm of social network analysis, can be built upon the numerous (social) distance measures, status scores, prominence measures, and so forth, that have been developed in the social network literature over the years.¹⁷ Many of these can be seen as special cases of cohesion or equivalence.

A useful measure for the study of policy and power networks is formulated as follows. Define r_i as (political) resources available to actor i . These might include resources i can draw from the political organization he is part of. Part of i 's resources might also be available to j , for instance through j 's membership of the same organization, denote these shared resources by $r_{i(j)}$ (note that $r_{i(j)}$ does not necessarily

17. Social distances are typically calculated from the presence or absence of links among actors or from the difference between a profile of characteristics of an actor and his perception of other actors' profiles on the same characteristics. Distances can be calculated from status scores or prominence scores by viewing the status or prominence of an actor (however derived) as an actor attribute. Differences in these scores can then be used as the basis of the weights in a weight matrix.

equal $r_{j(i)}$). Next, introduce a measure of nearness between i and j , d_{ij} , representing how difficult it is for j to claim i 's resources. This measure may be based on adjacency, social distance, role equivalence, and so forth. An appropriate weight matrix is

$$w_{ij}^{[7]} = \frac{r_{i(j)}^\alpha}{d_{ij}^\beta}$$

where α and β are parameters. Weight $w_{ij}^{[7]}$ may be interpreted as the extent to which i 's resources can be 'claimed'¹⁸ by j . The weight matrix is normalized straightforwardly. The measure is based on an object-oriented approach to network analysis, that has become increasingly popular over the recent years (among others Snijders, 1996; Stokman and Van Oosten, 1994; Zeggelink, 1993, 1994). It also fits well within the theory of core/adjoint networks (Leenders, 1995: 33-36). An adjoint network typically represents an interest group or corporate actor. A member of this network 'inherits' part of his resources from the resources of this corporate actor; the rest of his resources are 'private.' As an example, imagine an actor from an interest group who becomes a member of the political decision taking core, perhaps a leader of a labor union who is also affiliated with a political party. This leader now can exert more influence on both the party and the union by drawing on the resources offered to him by his membership to both.

An entire network consisting of core and adjoint networks should be modeled by using different weighting mechanisms in one weight matrix, by employing multiple weight matrices, or by a weight matrix consisting of partitions that reflect the influence pattern within and between subgroups (Doreian, 1989a: 377-380; Leenders, 1995: 84-91). For instance, the influence process may be driven by $w_{ij}^{[7]}$ in adjoint networks and by $w_{ij}^{[4]}$ in the core network.

A specification known as a 'gravity model' can be used to incorporate the 'size' of actors. For example, some organizations in an organization network, by means of their size, power, visibility, marketing budgets, or status may be able to systematically exert more influence on their alters than can others of smaller size, less power, or lower

18. It is assumed that i cannot use these resources as long as they are in use by j .

status. Their weight of influence on others should thus be larger. Similarly, some organizations, such as governmental or other politically representative organizations may be required to systematically be more open to societal influence than others. This can be modeled by

$$w_{ij}^{[8]} = U_i V_j F_{ij}$$

where V_j is the ‘size’ of influencer j , U_i is the ‘size’ of influencee i , and F_{ij} captures the facility with which j can influence i . This measure traces back to Carey (1858) and is still in popular use (Nijkamp, 1997; De Vries et al., 2000) and varied upon (e.g., Boyle and Flowerdew, 1997).

An alternative influence structure reflects a discontinuous pattern where connectedness ‘skips over’ the first connected actor. In this sense, i and j are connected if i has a relationship to k and k has a relationship to j . This model yields a new adjacency matrix \tilde{A} with elements

$$\tilde{a}_{ij} = \max_k (a_{ik} a_{kj}), \text{ for } k \neq i, j \quad (6)$$

and a weight matrix can be constructed according to any of the methods discussed above (e.g., Brandsma and Ketellapper, 1979). This type of adjacency matrix may be useful for the study of kinship or heredity networks. Mathematically (6) is equal to dichotomizing A^2 , if $a_{ii} = 0, i = 1, \dots, g$. A row normalized weight matrix is

$$w_{ij}^{[9]} = \frac{\tilde{a}_{ij}}{\tilde{a}_i}$$

A different way of specifying influence incorporates the concept of ‘thresholds of influence.’ Let d_{ij} denote some metric representing the nearness between i and j . In modeling inter-municipal influences in labor market-policies, Van Dam and Weesie (1991) use the following specification, employing only two levels of influence separated by a nearness-threshold D

$$w_{ij}^{[10]} = \begin{cases} 0 & i = j \\ \alpha_1 & d_{ij} < D \\ \alpha_2 & d_{ij} \geq D \end{cases}$$

Getis (1984) follows a similar approach. An alternative is to allow w_{ij} to be a continuous function of d . An example is

$$w_{ij}^{[11]} = \begin{cases} 0 & i = j \\ \alpha_0 + \alpha_1 d_{ij} + \alpha_2 d_{ij}^2 & \text{elsewhere} \end{cases}$$

Weights can also be taken as an exponential function of distance (Fotheringham et al., 1998; Talen and Anselin, 1998; White et al., 1981). In fact, any non-increasing function of distance can be applied.

In this section, many different operationalizations for the weight of the relation between two actors have been discussed. Yet other specifications are discussed by Friedkin (1998), starting from a social psychological point of view.

It is clear that in empirical situations the choice for a certain matrix W is not at all obvious. There are an infinity of possible representations. The choice for a certain weight matrix is often debatable. Weight matrix W allows a researcher to choose a set of weights that are appropriate from *prior* considerations. This allows great flexibility in defining the structure of influence in networks. Further, if different hypotheses are proposed about the degree of influence between units, alternative sets of weights might be used to investigate these hypotheses. This will depend on the study at hand. The *generally correct* weight-matrix does not exist. Only substantive knowledge can lend guidance to possible appropriate specifications.

7. Statistical tests on the weight structure

Sometimes the story ends here. In that case, the researcher has been able to construct an appropriate W -matrix, consistent with his or her theoretical beliefs, and can proceed to the host of analyses that have been developed for working with network autocorrelation models. However, in two cases, an additional step is needed. In the first, the researcher was able to narrow down the large set of potential weight matrices

to only a few, but not to one single one, and is not able to distinguish between them on theoretical grounds. In the second case, the researcher has narrowed down the possible influence structures to one and wants to test his theory, as embodied by a specific weight matrix, against alternative influence structures, represented by alternative W 's. In both cases, statistical procedures are required to test various specifications against one another. Yet another approach is to try to estimate directly the elements of W . However, if theoretical considerations have not enabled the researcher to make a justified choice of W -matrix, they will certainly unlikely provide the researcher with enough guidance for specifying an estimation model for the w_{ij} 's sensibly. In addition, the statistical issues connected to such an estimation are still largely unsolved, so the researcher has to make too many assumptions to make the likelihood functions tractable.

In Section 7.1 I will consider the situation in which the researcher considers one structure as his null hypothesis and wants to test that against alternative structures. In Section 7.2 the situation is considered in which in series of influence structures are equally likely from a theoretical point of view, and the final choice between them has to be made on statistical grounds.

7.1. *When a null hypothesis is available*

For ease of presentation, I will focus on the network effects model (2). The competing hypotheses H_0 and H_1 , which can represent different influence structures (in terms of a W matrix) as well as different explanatory variables (the X 's) can be expressed as:

$$\begin{aligned} H_0: y &= \rho_0 W_0 y + X_0 \beta_0 + \varepsilon_0 \\ H_1: y &= \rho_1 W_1 y + X_1 \beta_1 + \varepsilon_1 \end{aligned} \tag{7}$$

where W_0 and W_1 are two $(g \times g)$ -weight matrices, and X_0 and X_1 represent $(g \times k_0)$ and $(g \times k_1)$ -matrices of explanatory variables, some of which may be included in both X_0 and X_1 ; k_0 need not be equal to k_1 .

Classical testing procedures are invalid here. Usually, the competing formulations in (7) cannot be considered as limited forms of a more general expression, or as

restrictions on a general model (conditional upon the latter being true). Therefore, traditional tests do not fully apply in this situation (Anselin, 1984).

The test of one formulation against one or more alternatives, used to falsify the original specification, is performed by using tests for non-nested hypotheses. After a brief discussion of two approaches to testing non-nested hypotheses, I will focus more closely on a third approach.

A first approach is the so-called modified likelihood ratio test. For this test, Maximum Likelihood estimation of both models under their respective null hypotheses (H_0 and H_1) is performed and compared to its expected value under the null hypothesis, H_0 . In other words, it involves the estimation of the alternative model, with the null hypothesis assumed correct, using pseudo maximum likelihood estimation. Anselin (1984) provided the appropriate Cox-statistic for (7) and concluded that its variance is mathematically intractable. As a result, this approach is not recommended for network autocorrelation. A second approach, popular in spatial sciences, involves the use of instrumental variable (IV) estimators. This approach is sensitive to the choice of instruments. The disadvantage of applying IV-estimators for the testing purpose at hand, is that, in order for the competing formulations to be fully comparable, the same set of instruments should be used for all. However, as is obvious in the formulation of (7), one wants to allow for the inclusion of (possibly totally) different covariates with different W-matrices. In that case, using the same set of instruments for all specifications may be far from optimal, both theoretically and statistically. Therefore, I will limit myself here to simply referring to the appropriate literature (Anselin, 1984, 1988; Ericsson, 1983; Godfrey, 1983).

Considering the problems associated with the first two approaches, I suggest using the approach of augmented regressions. Consider again the case of testing H_0 against H_1 , as in (7). Null hypothesis H_1 is not nested within H_0 , and H_0 is not nested within H_1 . Thus, the truth of H_0 implies the falsity of H_1 , and vice versa. Now consider the augmented equation

$$y = (1 - \alpha)(\rho_0 W_0 y + X_0 \beta_0) + \alpha(\hat{\rho}_1 W_1 y + X_1 \hat{\beta}_1) + v \quad (8)$$

where $\hat{\rho}_1$ and $\hat{\beta}_1$ are ML estimates calculated from separate estimation of H_1 . If H_0

is true, then the true value of nesting parameter α is equal to 0. It is easily shown that $\hat{\rho}_1 W_1 y + X_1 \hat{\beta}_1$ is independent of v . Therefore, one can simply test whether $\alpha = 0$ in (8) by means of a conventional t -test or related tests. Test (8) is known as the J -test, since it involves estimating nesting parameter α and model parameters $\{\rho_0, \beta_0\}$ *jointly*. A number of variations to (8) have been proposed, but the J -test is the easiest to use. Although all variations are asymptotically equivalent, they tend to differ in finite samples. From Monte Carlo simulation results, the J -test seems to be preferable with regard to its finite sample properties (e.g., Anselin, 1986).

Three remarks are important here. First, since my focus here is on the statistical testing of one specification against one or more alternatives, individual measures of fit are not appropriate, as they compare the fit of a model to the fit of a null model, rather than to the fit of a specific alternative. Second, the approach of specifying one model as the null hypothesis and an alternative model as an alternative hypothesis is fundamentally different from testing for the presence of network autocorrelation (i.e. the significance of ρ for a given structure W). In the latter case, the null hypothesis is that ρ is equal to zero, and the alternative states that ρ is not. Failure to reject that null hypothesis does not necessarily mean that autocorrelation is absent, it simply means that the presence of autocorrelation is rejected for this W -structure. In the non-nested test, the W -matrices are not considered in isolation, but in direct relation to the specific alternative structures for W . Similarly, rejection or non-rejection of H_0 in (7) does not say anything about the presence of autocorrelation in the data. It may very well be that W_0 and X_0 are ‘accepted,’ but ρ_0 is found to not significantly differ from zero.

Third, formulation (7) only includes two model specifications, one of which is considered as the null hypothesis. In practical cases, one may have several alternative hypotheses that should be tested against H_0 . In that case, one essentially has a situation of multiple comparisons and the usual adjustments to significance levels apply. One can also extend (8) to test H_0 against m alternatives through

$$y = \left(1 - \sum_{i=1}^m \alpha_i \right) (\rho_0 W_0 y + X_0 \beta_0) + \sum_{i=1}^m \alpha_i (\hat{\rho}_i W_i y + X_i \hat{\beta}_i) + v$$

and test the hypothesis that all α_i ’s are zero.

7.2. *When no null hypothesis is available*

I now consider the situation where there is a fixed number of equally plausible models and the analyst has no reason to prefer any of them. A test then is needed to choose between these alternative specifications. None of the available specifications is taken as a null hypothesis.

Initially, one may be tempted to apply J -test (8) to this problem. After all, if H_1 is true, then α in (8) will asymptotically converge to one. Although it might seem that that would enable one to test the truth of H_1 directly from (8), this is not correct. The t -statistics for (8) are conditional on the truth of H_0 , not on the truth of H_1 . The alternative is to then simply reverse the roles of H_0 and H_1 and carry out the test again. It is possible that one of the two specifications is rejected in favor of the other one. It, however, is also possible that both hypotheses may be rejected, or that neither may be rejected. The reason for this is that the tests of 7.1 are tests on model specification, and not model discrimination tests. This may not be a problem when testing two specifications against one another, but when more specifications are available, drawing conclusions may become very complicated and unclear. Still, this approach is advocated by Anselin (Anselin and Can, 1986), who also devised a qualitative approach for choosing between alternatives.

Since no null hypothesis and alternative hypotheses are specified, it makes sense to base the model choice upon goodness-of-fit measures. The properties of traditional measures of fit, such as R^2 , do not directly carry over to the situation of network autocorrelation. It therefore seems reasonable to select that model that minimizes a quantity such as the Kullback-Leibler information criterion. This approach is applicable to nested and non-nested hypotheses and avoids the need to carry out multiple pairwise comparisons. Information theoretic approaches deal with measuring the closeness of the assumed model to the true, but unknown, model, while taking into account the trade-off between fit and parsimony of parameters. In contrast to the tests discussed in Section 7.1, which are based on testing models against given alternative specifications, information theoretic measures are associated with each model by itself. Since the true model is unknown, an estimate is needed. The Akaike information criterion (AIC), has been shown to be a useful estimate (Akaike, 1974, 1981; Judge et

al., 1985; Sawa, 1978):

$$\text{AIC} = -2L(\hat{\rho}, \hat{\beta}) + q(k) \quad (9)$$

where $L(\hat{\rho}, \hat{\beta})$ is the loglikelihood at the maximum and $q(k)$ is a penalty function of the number of unknown parameters in the model. The function $q(k)$ varies among different versions of AIC, the most commonly specifications being $q(k) = 2k$ or $q(k) = \ln(gk)$, with g being the number of actors in the network. Comparing the values of AIC for all models, the minimum AIC represents the model with maximum fit.

Thus, selecting one specification from a set of specifications on a statistical basis is really easy. First estimate all of the models separately. Take the value of the loglikelihood (as reported by the software used for estimating the autocorrelation models), choose a penalty function and calculate the AIC. Then pick that specification that minimizes it.

8. Empirical example

In this section I will provide a brief example of the statistical issues dealt with in this paper. First I will show that the choice for an alternative specification of W , or the choice of a network effect model versus a network disturbances model, can lead to different statistical results. I will also show how the specification tests discussed above can be applied. For these purposes I reanalyze the Louisiana voting data from Doreian (1980), in which it is argued that a spatial analysis of the data is appropriate. The example is for illustrative purposes only. The dependent variable y is the proportion of support in a parish for Democratic presidential candidate Kennedy at the 1960 elections. The covariates are B (percentage black in a parish), C (percentage Catholic), U (percentage urban), and BPE (a measure of black political equality). Each of these can be seen as predictors of electoral turnout and partisan electoral behavior in presidential elections. The spatial model describes the extent to which electoral behavior is contagious between adjacent parishes. Adjacency matrix A is a simple binary matrix describing adjacency among 64 parishes (counties).

Table 3 contains the results of four autocorrelation analyses, all containing the same covariates (B , C , U , and BPE) and differing only in the chosen W . The OLS

column contains the results for the non-spatial model $y = X\beta + \varepsilon$. When applying the network effects model, only weighting scheme $w_{ij}^{[1]}$ yields a statistically significant autocorrelation parameter. When the adjacency matrix is column normalized, is based on structural equivalence, or mimics a heredity structure, the autocorrelation parameter loses statistical significance. Note that the parameter estimates for the covariates and, more importantly, their statistical significance, only vary slightly over the various influence structures.¹⁹

**** Table 3 ****

The results for the network disturbances model are very different. Here, three out of four influence regimes yield statistically significant autocorrelation parameters, including $w_{ij}^{[2]}$ and $w_{ij}^{[9]}$. Again, the results for the covariates do not vary much.

**** Table 4 ****

Next, suppose the researcher has decided on substantive grounds that the network effects is the appropriate model and wants to test the situation in which exerted influence decreases with the number of actors influenced (row normalization) against the situation in which accepted influence decreases with the number of actors influencing (column normalization). Now (7) can be used to test $H_0: w_{ij}^{[1]}$ against $H_1: w_{ij}^{[2]}$. Estimating (8) with $W_1 = W^{[1]}$ and $W_2 = W^{[2]}$ and $X_1 = X_2$ yields $\hat{\alpha} = .14$ (.92). So, H_0 is not rejected and the appropriateness of $w_{ij}^{[1]}$ can not be rejected in favor of $w_{ij}^{[2]}$.

If the researcher had narrowed down on substantive grounds the set of probable weight structures to $w_{ij}^{[1]}$, $w_{ij}^{[2]}$, $w_{ij}^{[6]}$, and $w_{ij}^{[9]}$, but had absolutely no substantive reason to prefer one of the others, Akaike information criterion (9) would be appropriate. For both the network effects the network disturbances models, the resulting order of weight matrices is $w_{ij}^{[1]}$, $w_{ij}^{[2]}$, $w_{ij}^{[9]}$, $w_{ij}^{[6]}$ (see Table 5). The AIC criterion can also be used to choose between competing models. In total, nine models

19. This result can not be generalized. Neither is it warranted to generalize the present result that the statistical significance of the covariates is similar for the network effects models and the network disturbances model. For instance, in the analyses in Doreian (1980) statistical significance does vary over both models.

were estimated (four network effects models, four network disturbances models, and one non-spatial OLS model), the AIC values of which are reported in Table 5. The penalty function used is $q(k) = \ln(gk)$. The four regimes with statistically significant autocorrelation parameters achieve the best scores, the OLS model comes in last (although the differences between the last models is only minor).

**** Table 5****

9. Conclusions

In this paper I have attempted to structure the thought that goes into specifying a weight structure for network autocorrelation models. Several steps need to be taken in the specification of an appropriate weight structure. First, the researcher has to choose what mechanism is driving social influence: communication or comparison and/or adjacency or non-adjacency. The choice between adjacency and non-adjacency is especially important, because it is very difficult to empirically distinguish between communication or comparison as the drivers of social influence. A connected choice is related to the source of influence: do actors mimic alter's behavior or opinion itself, or do they mimic the adaptation of alter's behavior or opinion? The first source is captured by the network effects model, the second source by the network disturbances model.

Since actors usually have a limited influencing capacity (for instance, because the resources required for exerting influence are limited) or a limited readiness for accepting influence, a model choice is often made to normalize either the rows or columns of W . Normalizing the rows decreases the influence each alter has on ego with each additional alter. Normalizing the columns decreases the influence alter has on ego with the number of actors influenced by alter. An additional effect of normalizing W is that it may asymmetrize the influence structure, even though the original structure was symmetric.

Finally, for each mechanism of influence, each source of influence, and each normalization, a large set of potential weight structures remains. In this paper I have presented several of them, but many more are possible. I do not generically prefer one of them, the choice has to be made on substantive grounds. If substantive

considerations fail to limit the number of weight matrices to only one, or when alternative models can not be discarded on substantive grounds, several statistical techniques can assist the researcher in making a final choice or in testing models or specifications against alternatives.

Of course, it is impossible to present a complete overview. Hopefully the present overview will incite analysts to put more thought into the specification of W , since the usefulness of the entire approach of network autocorrelation models hinges upon it. In this paper, I have tried to provide the researcher with the tools and the thought that go into the specification of an appropriate weight structure. The analysis of social influence through network autocorrelation models is promising and has proved its potential over the years. The sensitivity of the results to the specification of weight matrix W , the key element of these models, dictates that careful thought be used in its construction. Unfortunately, the effort devoted by researchers to the appropriate choice of W pales in comparison to the efforts devoted to the development of statistical and mathematical procedures. I certainly do not question the usefulness of statistical progress in this area (e.g., Leenders 1995, 1997), but with this paper I do want to stress that these useful procedures lose their usefulness when applied to a model with an ill-specified weight matrix W . Many approaches to estimating parameters in network autocorrelation models have been devised over the years. Simulation studies have investigated their behavior in a wide range of situations. Accurate estimates and inferences of complex network autocorrelation models are no longer problematic with current computer technology. But, at the end of the day, any autocorrelation model is useless when W is not specified with explicit attention and care.

References

- Akaike, H., 1974, A new look at statistical model identification, *IEEE Transactions on Automatic Control* AC 19: 716-723
- Akaike, H., 1981, Likelihood of a model and information criteria, *Journal of Econometrics* 16: 3-14
- Anselin, Luc, 1982, A note on the small sample properties of estimators in a first order autoregressive model, *Environment and Planning A*: 1023-1030
- Anselin, Luc, 1984, Specification tests on the structure of interaction in spatial econometric models, *Papers of the Regional Science Association* 54: 165-182
- Anselin, Luc, 1986, Non-nested tests on the weight-structure in spatial autoregressive models: Some Monte Carlo results, *Journal of Regional Science* 26: 267-284
- Anselin, Luc, 1988, *Spatial Econometrics: Methods and models* (Kluwer: Norwell, MA)
- Anselin, Luc, Bera, Anil K., 1998, Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah, Aman, Giles, David E.A. (Eds.), *Handbook of applied economic statistics* (Marcel Dekker: New York)
- Anselin, Luc, Can, A., 1986, Model comparison and model validation issues in empirical work on urban density functions, *Geographical Analysis* 27: 179-197
- Baerveldt, Chris, Snijders, Tom A.B., 1994, Influences on and from the segmentation of networks: hypotheses and tests, *Social Networks* 16: 213-232
- Bavaud, François, 1998, Models for spatial weights: a systematic look, *Geographical Analysis* 30: 153-171
- Bem, Daryl J., 1972, Self perception theory. In: Berkowitz. Leonard (ed.), *Advances in*

- experimental social psychology, Vol. II (Academic Press, New York)
- Berelson, Bernard R., Lazarsfeld, Paul F., McPhee, William N., 1954, Voting (University of Chicago Press, Chicago)
- Borgatti, Stephen P., Everett, Martin G., 1992, Notions of position in social network analysis In: Marsden, Peter V. (ed.), Sociological Methodology 1992 (Jossey-Bass, San Francisco)
- Boyle, Paul J., Flowerdew, Robin, 1997, Improving distance estimations between areal units in migration models, Geographical Analysis 29: 93-107
- Brandsma, Andries S., Ketellapper, Ronald H., 1979, A biparametric approach to spatial autocorrelation, Environment and Planning A: 51-58
- Burt, Ronald S., 1987, Social contagion and innovation: cohesion versus structural equivalence, American Journal of Sociology 92: 1287-1335
- Burt, Ronald S., Doreian, Patrick, 1982, Testing a structural model of perception: conformity and deviance with respect to journal norms in elite sociological methodology, Quality and Quantity 16: 109-150
- Carey, H.C., 1858, Principles of social science (Lippincott, Philadelphia), Reprint 1963 (Kelley, New York)
- Cliff, Andrew D., Ord, J. Keith, 1973, Spatial autocorrelation (Pion, London)
- Cliff, Andrew D., Ord J. Keith, 1981, Spatial processes: models and applications (Pion, London)
- Coleman, James S., Katz, Elihu, Menzel, Herbert, 1957, The diffusion of an innovation among physicians, Sociometry 20: 253-270
- Coleman, James S., Katz, Elihu, Menzel, Herbert, 1966, Medical Innovation (Bobbs-Merrill, New York)

- De Vries, Jacob J., Nijkamp, Peter, Rietveld, Piet, 2000, Alonso's general theory of movement, Tinbergen Institute Discussion Paper 2000-062/3
- Doreian, Patrick, 1980, Linear models with spatially distributed data, *Sociological Methods and Research* 9: 29-60
- Doreian, Patrick, 1981, Estimating linear models with spatially distributed data. In: Leinhardt, Samuel (ed.), *Sociological Methodology* 1981: 359-388 (Jossey-Bass, San Francisco)
- Doreian, Patrick, 1982, Maximum likelihood methods for linear models, *Sociological Methods and Research* 10: 243-269
- Doreian, Patrick, 1989a, Network autocorrelation models: Problems and prospects. In: Griffith, Daniel A. (ed.), *Spatial Statistics: past, present, future* (Michigan Document Services, Ann Arbor)
- Doreian, Patrick, 1989b, Models of network effects on social actors. In: Freeman, Linton C., White, Douglas R., Romney, Kimball (eds.), *Research methods in social analysis: 295-317* (George Mason University Press, Fairfax)
- Doreian, Patrick, Teuter, Klaus, Wang, Chi-Sein, 1984, Network autocorrelation models: some Monte Carlo results, *Sociological Methods and Research* 13: 155-200
- Doreian, Patrick, Woodard, Katherine, 1994, Defining and locating cores and boundaries of social networks, *Social Networks* 16: 267-293
- Dow, Malcolm M., 1984, A biparametric approach to network autocorrelation: Galton's problem, *Sociological Methods and Research* 13: 201-217
- Dow, Malcolm M., Burton, Michael L., White, Douglas R., 1982, Network autocorrelation: a simulation study of a foundational problem in regression and survey study research, *Social Networks* 4: 169-200
- Ericsson, Neil R., 1983, Asymptotic properties of instrumental variables statistics for

- testing non-nested hypotheses, *Review of Economic Studies* 50: 287-304
- Festinger, Leon, Kelly, Harold H, 1951, *Changing attitudes through social contact: an experimental study of a housing project* (Research Center for Group Dynamics, Ann Arbor)
- Festinger, Leon, Schachter, Stanley, Back, Kurt, 1950, *Social pressures of an informal groups; a study of human factors of housing* (Harper, New York)
- Fotheringham, A.S., Charlton, M.E., Brunson, C., 1998, Geographically weighted regression, *Environment and Planning A* 30: 1905-1927
- French, John R.P., Jr., 1956, A formal theory of social power, *Psychological Review* 63: 181-194
- Friedkin, Noah E., 1984, Structural cohesion and equivalence explanations of social homogeneity, *Sociological Methods and Research* 12: 235-261
- Friedkin, Noah E., 1998, *A structural theory of social influence* (Cambridge University Press, Cambridge)
- Galaskiewicz, Joseph, Wasserman, Stanley S., 1989, Mimetic processes within an organizational field: an empirical test, *Administrative Science Quarterly* 34: 454-479
- Getis, Arthur, 1984, Interaction modeling using second-order analysis, *Environment and Planning A* 16: 173-183
- Godfrey, L.G., 1983, Testing non-nested models after estimation by instrumental variables or least squares, *Econometrica* 51: 355-365
- Griffith, Daniel A., 1976, Spatial autocorrelation problems: some preliminary sketches of a structural taxonomy, *The East Lakes Geographer* 11: 59-68
- Haining, Robert, 1978, The moving average model for spatial interaction, *Transactions of the Institute of British Geographers* 3: 202-225

- Hoede, Cees, 1979, A new status score for actors in a social network (Twente University of Technology, Dept. of applied mathematics)
- Homans, George C., 1950, The human group (Harcourt Brace, New York)
- Homans, George C., 1961, Social behavior: its elementary forms (Harcourt Brace, New York)
- Judge, George G., Griffiths, W.E. , Hill, R. Carter, Lütkepohl, Helmut, Lee, Tsoung-Chao, 1985, The theory and practice of econometrics, 2nd. edition (Wiley, New York)
- Lazarsfeld, Paul F., Berelson, Bernard, Gaudet, Hazel, 1948, The people's choice: how the voter makes up his mind in presidential campaigns (Columbia University Press, New York)
- Leenders, Roger Th.A.J., 1995, Structure and influence: statistical models for the dynamics of actor attributes, network structure and their interdependence (Thela Thesis Publishers, Amsterdam)
- Leenders, Roger Th.A.J., 1997, Longitudinal behavior of network structure and actor attributes: modeling interdependence of contagion and selection. In: Doreian, P., Stokman, F.N. (eds.), Evolution of social networks (Gordon and Breach, New York)
- Loftin, Colin, Ward, Sally K., 1983, A spatial correlation model of the effects of population density on fertility, *American Sociological Review* 48: 121-128
- Lorrain, François, White, Harrison C., 1971, Structural equivalence of individuals in social networks, *Journal of Mathematical Sociology* 1: 49-80
- Marsden, Peter V., Friedkin, Noah. E., 1993, Network studies of social influence, *Sociological Methods and Research* 22: 127-151
- Marsden, Peter V., Podolny, Joel, 1990, Dynamic analysis of network diffusion processes. In: Weesie, J., Flap. H. (eds.), *Social networks through time* (Isor, Utrecht)
- Mead, R., 1967, A mathematical model for the estimation of inter-plant competition, *Bi-*

ometrics 23: 189-205

Merton, Robert K., 1936, The unanticipated consequences of purposive action, *American Sociological Review* 1: 894-904

Mizruchi, Mark S., 1989, Similarity of political behavior among large American corporations, *American Journal of Sociology* 95: 401-424

Mizruchi, Mark S., 1990, Cohesion, structural equivalence, and similarity of behavior: an approach to the study of corporate political power, *Sociological Theory* 8: 16-32

Mokken, Robert J., 1979, Cliques, clubs, and clans, *Quality and Quantity* 13: 161-173.

Mur, J., 1999, Testing for spatial autocorrelation: moving average versus autoregressive processes, *Environment and Planning A* 31: 1371-1382.

Nijkamp, P., 1997, Gravity models of spatial interaction behavior by Ashish Sen and Tony E. Smith, Springer 1995 (book review), *Journal of Economic Literature* 35: 1435-1436

Ord, J. Keith, 1975, Estimation methods for models of spatial interaction, *Journal of the American Statistical Association* 70:120-126

Rietveld, Piet, and Wintershoven, Patrick, 1998, Border effect and spatial autocorrelation in the supply of network infrastructure, *Papers in Regional Science* 77: 265-276.

Sawa, T., 1978, Information criteria for discriminating among alternative regression models, *Econometrica* 46: 1273-1291

Seidman, Stephen B., 1983, Network structure and minimum degree, *Social Networks* 5: 269-287

Seidman, Stephen B., and Brian L. Foster, 1978, A graph-theoretic generalization of the clique concept, *Journal of Mathematical Sociology* 6: 139-154

Snijders, Tom A.B., 1996, Stochastic actor-oriented dynamic analysis, *Journal of Math-*

ematical Sociology 21: 149-172

Stevens, S.S., 1957, On the psychophysical law, Psychological Review 64: 153-181

Stevens, S.S., 1962, The surprising simplicity of sensory metrics, American psychologist 17: 29-39

Stokman, Frans N., Van den Bos, Jan M.M., 1992, A two stage model of policy making with an empirical test in the U.S. energy policy domain. In Moore, G., Whitt, J.A. (eds.): The political consequences of social networks, Research in politics and society 4 (JAI Press, Greenwich)

Stokman, Frans N., Van Oosten, Reinier, 1994, The exchange of voting positions: an object oriented model of policy networks. In: Bueno de Mesquita, B., Stokman, F.N. Stokman (eds.), European Community decision making: models, applications, and comparisons (Yale University Press, New Haven)

Strang, David, Tuma, Nancy Brandon, 1993, Spatial and temporal heterogeneity in diffusion, American Journal of Sociology 3: 614-639

Tajfel, Henri, 1972, La catégorisation sociale. In: Serge Moscovici (ed.), Introduction à la psychologie sociale (Larousse, Paris)

Talen, E., Anselin, L., 1998, Assessing spatial equity: an evaluation of measures of accessibility to public playgrounds, Environment and Planning A 30: 595-613

Tiefelsdorf, Michael, 1998, Some practical applications of Moran's I 's exact conditional distribution, Papers in Regional Science 77: 101-129

Van Dam, Marcel, Weesie, Jeroen, 1991, Intergemeentelijke beïnvloeding bij lokaal arbeidsmarktbeleid, ICS Prepublicatie 8 (ICS, Utrecht)

White, Douglas R., Burton, Michael L., Dow, Malcolm M., 1981, Sexual division of labor in African agriculture: a network autocorrelation analysis, American Anthropologist 83: 824-849

Whittle, P., 1954, On stationary processes in the plane, *Biometrika* 41: 434-449

Wickes, A.W., 1969, Attitude versus actions: the relationship of verbal and overt behavioral responses to attitude objects, *Journal of Social Issues* 25: 41-78

Zeggelink, Evelien P.H., 1993, *Strangers into friends* (Thela Thesis Publishers, Amsterdam)

Zeggelink, Evelien P.H., 1994, Dynamics of structure: an individual oriented approach, *Social Networks* 16: 295-333

Figures

FIGURE 1 Social influence.

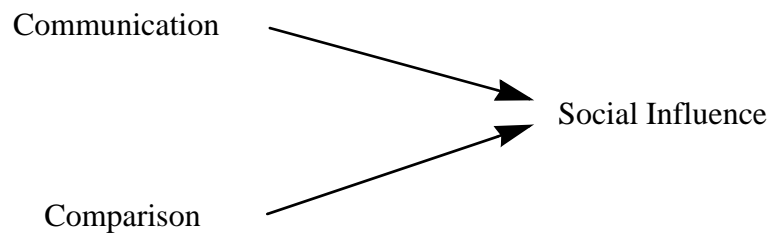
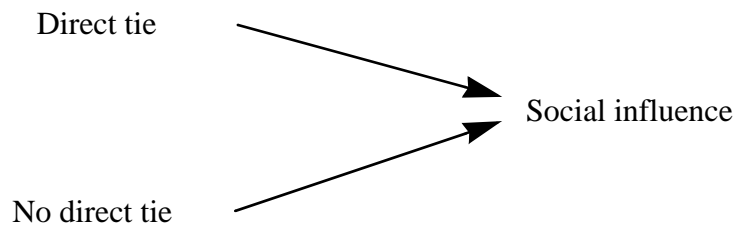


FIGURE 2 Alternative approach to social influence.



Tables

Table 1
Social influence among adjacent and non-adjacent actors.

ADJACENT ACTORS	NON-ADJACENT ACTORS
Communication	
Indirect communication	Indirect communication
Comparison	Comparison
Indirect comparison	Indirect comparison

Table 2
Row normalization versus column normalization.

ROW NORMALIZATION	COLUMN NORMALIZATION
<ul style="list-style-type: none"> • each outgoing contact has equal influence for each actor • weight proportional to outdegree • total amount of accepted influence equal for all actors • deals with accepted/received influence 	<ul style="list-style-type: none"> • each incoming contact has equal influence for each actor • weight proportional to indegree • total amount of exerted influence equal for all actors • deals with exerted/executed influence

Table 3
 Network effects model for the Louisiana voting data^a

	OLS		$w_{ij}^{[1]}$		$w_{ij}^{[2]}$		$w_{ij}^{[6]}$		$w_{ij}^{[9]}$	
ρ	--	--	.31*	(.10)	.07	(.06)	.12	(.25)	.04	(.12)
const.	21.03*	(4.40)	13.87*	(4.67)	19.83*	(4.34)	16.78	(10.06)	19.80*	(5.62)
B	.01	(.08)	-.00	(.07)	.00	(.08)	.01	(.08)	.01	(.08)
C	.30*	(.04)	.22*	(.05)	.28*	(.04)	.29*	(.05)	.29*	(.05)
U	-.11*	(.04)	-.10*	(.04)	-.11*	(.04)	-.11*	(.04)	-.11*	(.04)
BPE	.39*	(.06)	.30*	(.06)	.37*	(.06)	.38*	(.06)	.38*	(.06)

^a. An asterisk denotes statistical significance at $p < .05$.

Table 4
 Network disturbances model for the Louisiana voting data^a

	$w_{ij}^{[1]}$	$w_{ij}^{[2]}$	$w_{ij}^{[6]}$	$w_{ij}^{[9]}$
ρ	.69* (.10)	.53* (.13)	.22 (.42)	.74* (.15)
const.	26.99* (4.50)	24.98* (4.22)	21.52* (4.30)	24.51* (5.06)
B	-.11 (.07)	-.07 (.07)	-.00 (.08)	-.09 (.08)
C	.37* (.05)	.35* (.04)	.31* (.04)	.38* (.04)
U	-.07* (.03)	.08* (.03)	-.11* (.04)	-.10* (.04)
BPE	.24* (.06)	.30* (.06)	.38* (.06)	.29* (.06)

^a. An asterisk denotes statistical significance at $p < .05$.

Table 5
 Order of W-matrices and autocorrelation models according to AIC

	Weight matrix	AIC	Order within model	Overall order
Network effects model	$w_{ij}^{[1]}$	439.12	1	3
	$w_{ij}^{[2]}$	445.52	2	5
	$w_{ij}^{[6]}$	446.78	4	8
	$w_{ij}^{[9]}$	446.44	3	6
Network disturbances model	$w_{ij}^{[1]}$	431.92	1	1
	$w_{ij}^{[2]}$	436.33	2	2
	$w_{ij}^{[6]}$	446.69	4	7
	$w_{ij}^{[9]}$	440.95	3	4
OLS	--	446.82	--	9