

Statistical Network Analysis for Analyzing Policy Networks

Garry Robins, Jenny M. Lewis, and Peng Wang

To analyze social network data using standard statistical approaches is to risk incorrect inference. The dependencies among observations implied in a network conceptualization undermine standard assumptions of the usual general linear models. One of the most quickly expanding areas of social and policy network methodology is the development of statistical modeling approaches that can accommodate such dependent data. In this article, we review three network statistical methods commonly used in the current literature: quadratic assignment procedures, exponential random graph models (ERGMs), and stochastic actor-oriented models. We focus most attention on ERGMs by providing an illustrative example of a model for a strategic information network within a local government. We draw inferences about the structural role played by individuals recognized as key innovators and conclude that such an approach has much to offer in analyzing the policy process.

KEY WORDS: social network models, statistical models, quadratic assignment procedure, exponential random graph models, stochastic actor-oriented models, innovation

Social network methodology has a long research tradition, going back at least as far as the 1930s (e.g., Moreno & Jennings, 1938). In the last 20 years, however, there has been a dramatic growth in new and innovative network methods. An exciting area has been the development of modern statistical approaches for the analysis of social networks. For this special issue of the *Policy Studies Journal*, this article reviews three important statistical methods that are currently widely used in social network analysis and that are illustrated in the articles published in the special issue. The aim is for a tutorial style, relatively nontechnical summary that will assist readers unfamiliar with these techniques. We assume that readers are knowledgeable about basic network terms and concepts and standard statistical methods relevant to the social sciences.

The article is focused on the three statistical methods, but in an actual study—as is the case for any area of social science—the analysis cannot be divorced from broader issues of research questions and hypotheses, research design, and data collection. Other articles in this special issue discuss network-based research questions and design relevant to policy studies, and the empirical articles illustrate how

the methods presented here can be incorporated into specific studies, with examples of different types of network data collection.

The article is structured as follows. After introducing some terminology and notation, we explain the distinctiveness of network data, discuss why special methods are necessary, and illustrate some of the important research questions that might be asked. We then summarize three statistical methods: quadratic assignment procedure (QAP) approaches, exponential random graph models (ERGMs), and stochastic actor-oriented models (SAOMs). We focus most attention on ERGMs by providing an illustrative example of a model for a strategic information network within a local government. We draw inferences about the structural role played by individuals recognized as key innovators and conclude that the statistical modeling of networks is better able to test policy theory than the descriptive network methods often used in policy studies to date.

In this article, we typically concentrate on binary network data (i.e., a network tie is present or not), but the three methods have been extended to a broad range of other types of network data, including multiple, bipartite, and other network data structures. There are a few constraints on the type of data that can be used for the different methods. SAOMs require longitudinal panel data. ERGMs are typically used for cross-sectional network data, but there are longitudinal versions. Although ERGMs for valued data have been developed, current practice is to use binary data for the network in question (exogenous predictors, either networks or attributes, may be valued). As is noted below, more work needs to be done on using binary data in network regressions using QAP, although binary data are fine for QAP network correlations.

Of course, these three methods are not the only means of analyzing networks. For readers requiring detailed information about network analysis more generally, the now classic textbook on basic network concepts and earlier methods is Wasserman and Faust (1994). More recent methodological developments are described in the volumes edited by Carrington, Scott, and Wasserman (2005) and by Scott and Carrington (2011). Butts (2008) and Wasserman and Robins (2012) provide introductions to and reviews of current social network analysis. Kolaczyk's (2009) excellent text on network statistical models covers developments in the social network, statistical, and physics literatures in detail. Pattison and Robins (2008) provide a chapter-length review of network statistical approaches, focusing on ERGMs. Snijders (2011) provides a comprehensive overview of the current state of the statistical modeling of social networks. Jackson (2010) and Newman (2010) have produced recent texts that describe a wide range of network methods.

Some Basic Network Terminology and Notation

The simplest form of a social network is a set of social *actors* with a possible social relationship (or *relational ties*) among them. A social relationship can be quite general: there is no supposition that this is "social" in the sense of "out-of-hours" or "socializing"; nor is it necessarily a product of social media (although it may be), so is not specific to "social networking" websites. In other words, a relational tie can be

any form of social connection between two individuals that is relevant to the purposes of the research. Possible relationships include: communicating, working with, hating, bullying, trusting, and so on.

A network can be represented as a (mathematical) graph, with the actors represented by *nodes* (or *vertices*) and the relational ties by *edges*, visualized as dots with lines between them. These different terminologies (actor/node; tie/edge) are often used interchangeably. A binary tie variable X_{ij} equals 1 if there is a tie between actor i and actor j , and is 0 otherwise. Consistent with statistical practice, a notation is often used whereby $X_{ij} = x_{ij}$ indicates that the variable for a possible tie between i and j takes the value x_{ij} (which might be 0 or 1), with x_{ij} called an instantiation or realization. (In other words, it can be convenient to differentiate the variable and the value it takes in a particular instance.) Accordingly, empirical network data can be written as a binary square *sociomatrix* or *adjacency matrix* $x = [x_{ij}]$ with structural zeroes down the diagonal to indicate that a node does not have a tie with itself. The corresponding matrix of network variables is X .

This basic formulation can be extended in a variety of ways. X may be valued instead of binary, although this article concentrates on binary data. The network may be *undirected* where there is no direction to the tie (so that $x_{ij} = x_{ji}$ and the sociomatrix x is symmetric), or *directed* (where x_{ij} need not be equal to x_{ji} and the sociomatrix x need not be symmetric). For directed networks, the term *edge* is replaced by *arc*. There may be two or more relational ties measured on the same set of nodes, resulting in two (or more) networks x and z , often known as a *multiple* or *multivariate network*. If network data are collected across time in a panel data design, then the network variable becomes X_{ijt} to indicate a tie between i and j at time t . In this case, there will be multiple sociomatrices x_t indicating the network as observed at time points t .

These descriptions are all for *unipartite networks*, where there is only one type of node (i.e., actors). However, relational data can take different forms, including two (or more) types of nodes. Such data can represent membership or association (e.g., actors may be members of groups), in which case there may be ties between nodes of different types but not between nodes of the same type. *Bipartite networks* have two types of nodes. In a bipartite network, the variable X_{ij} indicates a tie from actor i to group j , and the association matrix x may be rectangular (with no structural zeroes).

Of course, the nodes in a network are not just actors in the sense of expressing ties to other actors: they may have certain qualities, motivations, or properties. Such individual-level variables are called *actor* (or *node*) *attributes*. In this article, attribute variables will be denoted by Y_i for a variable pertaining to the attribute of actor i . As before, y_i is a realization and may take binary, categorical, or continuous values. So, the vector $y = [y_i]$ contains data across all actors for the attribute variable Y . Of course, there may be many such attribute variables in the one study.

Why Are Network Datasets Different?

Networks have become a popular concept in the social sciences, given increased attention by the interconnected, globalized world in which we live. Of course, social

networks have always been with us (Padgett & Ansell, 1993), but it is only in recent decades that they have become so prominent a research tool. Many social scientists, coming to network analysis for the first time, see it through the lens of familiar standard statistical approaches. To do so is risky. It has been over two decades since David Krackhardt has shown that even modest network effects can blow out alpha levels dramatically (Krackhardt, 1988). In other words, if your data are networked in some way, your standard null hypothesis significance test may be drastically wrong, and you may not even know it.

The reason can be stated simply, even though it is the principal source of the complexity in network research. Standard statistical approaches assume (at some level) that observations are independent of one another; a network model expressly does not. If a social tie between two actors does not imply some form of dependence, then for the purposes of the research, there might as well be no connection, and the network is irrelevant. So if networks are relevant, there is some level of complex dependence in the data.

Social Influence

So, to take a simple starting illustration, observations on actors (i.e., attribute measures) may be dependent based on the network. In other words, one observation may be associated with another through network connections. An important network research question is often one of *social influence* (sometimes called *contagion* or *network diffusion*): that is, given a social network, do network partners influence each other in some way? In other words, if there is a tie between actors i and j , then it might be theorized that the actors may influence one another, so that their attribute variables become similar to each other. For instance, we may adopt the opinions or behaviors of our workmates, or one person may catch a disease through friendship with an infected person.

Disease transmission is a paradigmatic case of network contagion. Let us suppose that the attribute variable $y_i = 1$ indicates that person i has the disease (whereas 0 indicates good health). Suppose we are uncertain whether a new disease is contagious or not, but we see the pattern of variable values $y_i = y_j = x_{ij} = 1$ across many instances of i and j pairs (or $y_i y_j x_{ij} = 1$). That is, we observe many circumstances where, for pairs of actors i and j with a tie between them, both actors have the disease. Then we can treat this as evidence that the disease is contagious and that it is possible to catch it through the network X . Notice that we can represent this pattern of variables in a simple diagram (Figure 1a), where we have two dotted nodes (representing being diseased) connected by a tie. If for a given i and j , we have $y_i y_j x_{ij} = 1$, then we will see the network pattern in Figure 1a. The pattern is what is called a *network configuration*, in this case, one that includes both tie and node information. This is a basic configuration for network influence.

The point of this example is that, with network influence, it follows that there must be dependence among observations of the disease, dependence arising from the network ties. Suppose our research question relates to the factors that might lead to an individual being vulnerable to infection with the disease. There are at least two

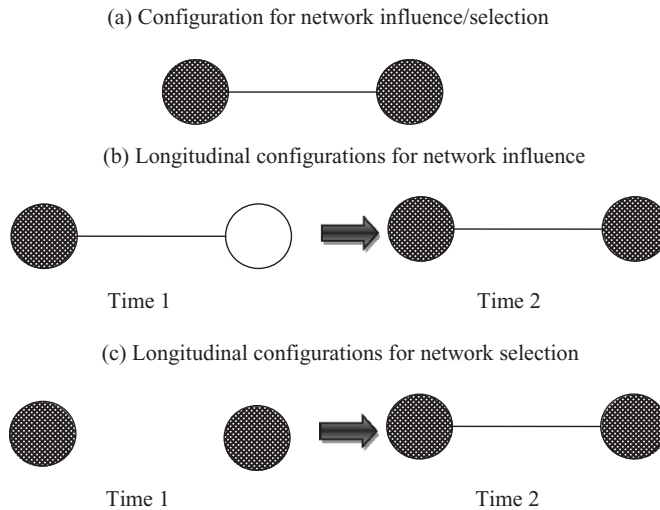


Figure 1. Some Network Configurations Involving Tie and Attribute Variables (Circle Indicate Nodes; Lines Indicate Ties; Dotted Circles Indicate Nodes with Attribute Value 1).

general sources for infection that we might hypothesize: (i) those relating to the individual, i.e., attributes; (ii) the possibility of contagion, i.e., network transmission of the disease. If we ignored the second, one standard approach to predict a binary variable such as disease status would involve a logistic regression to predict y_i from other attribute variables pertaining to i (genetics, age, behavior, and so on). This might be fine for the many diseases that are not contagious.

Dependence among observations, however, undermines this whole analytic strategy. If the disease is transmitted from others in the network, we are obviously missing an important source of infection if we concentrate on attribute variables alone. Yet, it is not simply that we have forgotten an important variable in our logistic regression. Rather, contagion implies that there is a need for a different type of analysis that caters for the dependence. The convenient distinction between predictor and outcome variables, the cornerstone of regressions, is now muddled.

With network contagion, we have to predict y_i from y_j as well as possibly other attribute variables—in fact we have to predict $Y_i = 1$ from $y_j x_{ij}$. And then we have to predict $Y_j = 1$ from $y_i x_{ij}$. Notice how the predictor variable in the first case (y_j) has now become the outcome variable in the second. It is possible to do these predictions in a formal sense with logistic regression procedures (and the authors admit to doing such things in the days before they knew better.) However, the approach is certainly not a standard logistic regression—in some parts of the regression, a variable is a predictor, while in other parts, it is an outcome—so the correctness of the procedure is not immediately obvious. In fact, it is now well known that the resulting estimates may be substantially incorrect, and analysts will not know whether this is the case.¹

Of course, if we have longitudinal network data, we can study this process in a more dynamic way. Let us suppose that we observe the situation in Figure 1b. At time

1, we see a healthy actor who is tied to a diseased actor; whereas at time 2 both actors are diseased. If we see many instances of this pattern across time, we can take this as evidence for a contagion process.

This discussion has focused on contagion through a dyadic network tie. Robins (2009) and Wasserman and Robins (2012) used the term *dyadic influence* to describe this situation. However, it is also possible that individual attributes may be influenced and so change because an individual occupies a particular *network position*, rather than being connected to a particular type of individual. For instance, an important theme in social network theory discusses the advantages that may accrue to *network brokers*, individuals who occupy *structural holes*, “gaps” in the network structure (Burt, 1992). It is frequently argued that individuals who occupy these gaps in the structure may gain personally (e.g., increase in power, wealth, and so on, depending on context). In this case, an individual who occupies a particular structural position achieves a positive change in their attribute values. Robins (2009) and Wasserman and Robins (2012) referred to this situation, where a network position (as opposed to dyadic connections) influences attribute values, as *generalized influence*.

Social Selection

The foregoing presupposes that the attribute variable can change (a person can become diseased) but that the network tie does not. But of course networks are not static, and ties come into and out of existence. One reason why ties might occur is because the two actors have similar attributes, a process known as *homophily*. For instance, you may be more likely to become friends with people of your own age. Suppose that, in our example, diseased individuals come together and become friends through support groups relating to the disease. Then we would expect to see more pairs of friends, both of whom have the disease—in other words, $y_i y_j x_{ij} = 1$. This is also the configuration from Figure 1a. Yet, this is obviously not social influence. In this case, it is the network tie that has been created by the disease status.

We describe *social selection* as the situation where attribute variables lead to the formation or dissolution of network ties; in contrast to social influence, when the presence of a network tie leads to a change in attributes. Homophily is not the only social selection process: For instance, actors with certain attributes might express or receive more ties (or be more *active* or *popular*).

So, we see that Figure 1a is a configuration that can result from either social influence or social selection. Without further information, we cannot determine which process applies. This is akin to a cross-sectional correlation when we infer association, not causality. With Figure 1a, we have a configuration that represents an association between similar actor attributes and network ties, but we cannot usually be sure whether we have influence or selection. (Of course, sometimes we can be sure. For instance, if the attribute is unchanging—, e.g., sex—we know the process cannot be influence.)

Across time, however, homophilous selection can be apparent. The sequence represented in Figure 1c illustrates a tie forming by time 2 based on shared nodal

attributes at time 1. If we observe enough of this type of change, we conclude that the common attributes at time 1 lead to a network tie at time 2, and hence infer a selection process.

Again, discussion of homophily focuses attention on *dyadic selection*, but individuals with particular attributes may seek to obtain a network position. For instance, the motivations of individuals may influence whether or not they seek to occupy structural holes (e.g., Kalish, 2008). Robins (2009) referred to *generalized selection* as the social process whereby individuals with certain attribute values sought particular network positions, rather than dyadic network partners.

Structural Processes: Network Self-Organization

Selection processes suggest new possibilities for research questions that are not available for standard datasets. Selection deals with tie formation arising from patterns of attributes. However, ties may also form because of the patterns of other ties. These *structural processes* lead to network configurations that do not involve attributes at all. The simplest to understand is *reciprocity* in directed networks. The presence of a tie from actor i to j may encourage the formation of a tie from j to i . Reciprocity is a basic social process that we see in many (although not all) social networks. The point is that the force for reciprocation may occur, irrespective of particular actor attributes. This does not mean that it is universal or deterministic, but tendencies toward reciprocity are often very strong.

Network *closure* is another important structural effect that frequently occurs in social network data. There is often a tendency for network ties to triangulate, so that a 2-path from i to j to k may close into a *triangle* with the addition of a tie between i and k . Theoretical discussion of the importance of triads in networks goes back to the beginnings of the twentieth century (Simmel, 1908). Empirical evidence for network closure (sometimes referred to as network clustering) is well established (Davis, 1970). Granovetter (1973) argued that certain types of tie (i.e., strong ties) tend to close, whereas weak ties do not. Burt (1992) pointed out that the presence of nonclosure (*structural holes*) was also important. In a nonclosed 2-path, i to j to k , actor j can reap certain advantages as a *network broker* between nodes i and k .

Network closure can arise because individuals tend to introduce their network partners to each, or because they tend to operate in collaborative team-like structures. Lubell, Robins, and Wang (2011) point to a number of possible outcomes: increased social support and cooperation, closer scrutiny of individual actions, and stronger group norms or localized cultures.

Another important structural feature of networks is the dispersion of the degree distribution, or using different terminology, the spread of the network activity. The degree of a node is the number of ties expressed by that node, and the degree distribution is the distribution of degrees across the entire network. Sometimes a social network may be highly centralized around a few high-degree nodes (known as *hubs*) with most nodes having much lower degree. The result is a highly dispersed or skewed degree distribution, or a wide spread of network activity. A popular version

is the so-called inverse power law or *scale-free* degree distribution, which may emerge from preferential attachment processes whereby the popularity of some nodes attracts further popularity (Albert & Barabási, 2002; Barabási & Albert, 1999). There have been claims that scale-free networks are ubiquitous, but there is plenty of evidence that this is not so (Li, Alderson, Tanaka, Doyle, & Willinger, 2005). Poor estimation procedures when fitting degree distributions has sometimes led to incorrect inference that a scale-free model best represents the data (Jones & Handcock, 2003). Even so, it is not uncommon for social networks to be centralized around a few high-degree nodes, with a positively skewed degree distribution. There are occasions, however, where networks are not so centralized and have a more “equal” distribution of degrees among nodes. Whatever the empirical outcome, the shape of the degree distribution and the level of centralization is an important structural feature of the network. (For directed networks, there are in fact two degree distributions, relating to in- and out-degrees, so it is possible that one may be skewed and the other not.)

Additionally, there are novel structural processes possible with multiple network data. With two types of relational tie, X and Z , on the same set of nodes, one simple research question is whether there is an association between the two networks, so that the presence of a tie x_{ij} between nodes i and j is associated with the presence of tie z_{ij} . We call this network *entrainment*, where the ties in one network tend to be aligned with those in another. This is only one of many types of network patterns that may apply to multiple networks. For instance, Padgett and Ansell (1993) argued that the structural position (betweenness centrality) in marriage, commerce, and political networks jointly contributed to the overall influence of the Medicis. In a more recent context, Scholz, Whiteman, and Berardo (2011) analyzed the impact of structural position among government agencies and stakeholders in four separate resource networks on the likelihood that projects will gain funding in a cooperative funding program.

We can present these different types of network structures again through sets of representative network configurations, this time not involving actor attributes. Figure 2 presents configurations representing reciprocity (a pair of reciprocated arcs), closure (a triangle), degrees (*stars* of various order), and entrainment (a pair of edges of different types). The reciprocity configuration only applies to directed networks, whereas the others are presented in Figure 2 for undirected networks. There are of course various directed counterparts of these (Robins, Pattison, & Wang, 2009).

Configurations such as those in Figure 2 can be thought of as the building blocks of network structure. Certain policy theorizations are directly translatable into such basic forms. For instance, Berardo and Scholz (2010) proposed the risk hypothesis whereby actors in governance systems prefer different network positions depending on whether the underlying motive is for coordination or for more risky cooperation. Under coordination conditions, actors prefer to connect to popular nodes for efficiency in information seeking. This type of structure can be represented with the star-like configurations in Figure 2. On the other hand, in more risky circumstances, actors prefer closed network structures, represented by

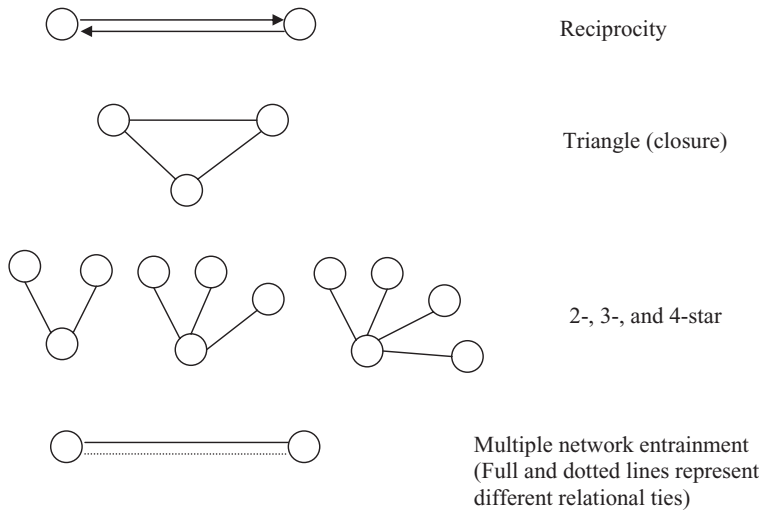


Figure 2. Some Network Configurations Involving Only Tie Variables.

the triangle in Figure 2, where both cooperation and adherence to norms can be more rigorously scrutinized.

These (and other) structural patterns can occur without any effects of actor attributes. In other words, the network ties themselves form into patterns, sometimes termed *network self-organization* (Wasserman & Robins, 2012). Not only may there be dependencies among attribute variables due to network ties, as in social influence, but there may be dependencies among the network tie variables themselves. Without dependence among ties, there can be no regular patterning of ties. So, the presence of network structure accentuates the issue of dependence in social network data, emphasizing the problems in using standard statistical approaches that rely on independence.

Investigating these local network patterns enables us to examine the structure of the network. The structure of a system is not a research issue that often occurs in other areas of social science. We can ask questions about whether, for instance, a more centralized structure is more effective in spreading information, whether a denser network encourages more collaborative outcomes, and whether a more disconnected network is associated with greater levels of competition among sub-groups. These questions of network performance and effectiveness are among the most challenging but exciting research issues for empirical social network analysis.

Throughout this exposition, we have emphasized that we are not talking about universal, deterministic effects. Rather, social networks are changeable, stochastic entities, but they can evolve according to stable self-organizing principles. Accordingly, we require statistical techniques that can cope with the dependence implied by the very use of social network data. In the remainder of this article, we summarize three important approaches that are illustrated further in the empirical papers in this special issue.

QAP (and Related Approaches)

QAP is used as a statistical test to determine whether two matrices are correlated. As such, it is a technique that is not specific to the analysis of social networks. It was originally proposed by the statistician Mantel (1967) to examine whether disease was geographically clustered. QAP was then championed by the mathematical psychologist Hubert (1987), who first coined the phrase “quadratic assignment,” as a quite general data analytic approach (Hubert & Schultz, 1976).

The social network application is obvious, however. We often want to know whether two networks measured on the same set of nodes are associated in some way (perhaps entrained as in Figure 2). For instance, a researcher might ask whether the informal communication network within an organization aligns with the formal organizational hierarchy. Krackhardt (1987) proposed that QAP could be used to address such network research questions.

If we take two adjacency matrices, x and z , we can produce a statistic of association (e.g., a Pearson correlation coefficient) just by treating the corresponding cells of the two matrices (excluding diagonals) as a data point involving two variables (i.e., x_{ij} and z_{ij}).² A correlation or some other association statistic can then be calculated in the normal way.³ If all the network tie variables are independent of each other, then this statistic can be treated in precisely the same way we handle a regular correlation and subjected to a standard null hypothesis test to infer whether the value was significantly different from 0, thereby to demonstrate (or not) an association between the two networks.

However, we know that social networks are self organizing and that there will be dependence among network variables. The standard tests do not apply because the distributional assumptions underpinning them are now not applicable given the dependent observations. QAP provides a nondistributional bootstrapping method of testing whether there is an association between x and z .

The QAP approach involves permutation of the node labels. Nodes are labeled 1, 2, and so on up to n . In a permutation π , node 1 is relabeled by chance to be node $\pi(1)$ —where $\pi(1)$ is a number between 1 and n chosen randomly—node 2 is relabeled to be node $\pi(2)$, and so on. Then the permutation π is applied to one of the matrices, say x , to produce $\pi(x)$. In other words, what was the first row and column in matrix x is now row and column number $\pi(1)$, and so on. What has happened is that the rows and columns have been swapped around in x but of course all the existing structure in the data has been retained because we have not changed any of the entries in the matrix itself, just their ordering. Meantime, matrix z is left unchanged.

The argument is that if there is an association between x and z it will depend on the labeling of the nodes. For instance, suppose x is a network of informal communication and z the formal hierarchy in an organization. We might hypothesize an entrainment association such as workers communicating with their bosses. In that case, for instance, if node 2 is the boss of node 1, then we expect that node 1 will communicate with node 2: that is, there is a dependency between z_{21} and x_{12} such that if $z_{21} = 1$ there is a greater chance that $x_{12} = 1$. However, this obviously depends on the node labeling so that if we relabel node 1 to be (say) node 77 and node 2 to be (say)

node 93, the expected dependence between z_{21} and x_{12} now disappears because nodes 1 and 2 in x are no longer the same nodes in z .

So, if there is an association, we expect that the correlation (or other association statistic) for the real data, x and z , will be larger than the statistic for the permuted $\pi(x)$ and z . Of course, this depends on the extent to which the permutation drastically changes the reordering of the nodes. So we use a simple bootstrapping test by performing many permutations (say, 1,000), thereby creating a sample of 1,000 correlation coefficients from a null distribution where there is no association. If the observed correlation in the real data is larger than 95% of the correlations from permuted x , we infer that there is a significant association between x and z .

The QAP procedure we have described is for a bivariate association between two networks and is the classic form proposed by Mantel (1967) and Hubert (1987). QAP has been extended to a version of multiple network regression, multiple regression QAP (MRQAP), whereby multiple matrices are used to predict a “dependent” network. The predictor matrices may include matrices derived from combinations of actor attributes (e.g., similarity across pairs of attributes to check homophily effects). The regression coefficients are estimated in the same way as for a standard regression. Each of the estimates, however, is tested individually in a QAP-type procedure focusing on the given matrix, while the other matrices remain unpermuted. MRQAP has been more controversial than QAP because previous approaches to permutation in a multiple regression framework have been debatable. Dekker, Krackhardt, and Snijders (2007) proposed a new permutation approach for QAP (“double semi-partialling”) as among the most robust approaches. This has now been implemented in commonly used social network analytic software such as UCINET (Borgatti, Everett, & Freeman, 2002). Dekker et al. (2007), however, pointed to some limitations of MRQAP tests. They noted that their study was limited to models where the dependent variable was continuous or countable integers. They concluded that further work was needed because “we do not know the behaviour of the different MRQAP approaches for models wherein the data are binary or ordinal, as is often found in social network data” (Dekker et al., p. 580). So while the application of QAP to binary data is fine, MRQAP with binary network data should be used with appropriate caution.

Comparison of Data with Null Distributions of Graphs

Although strictly speaking, the approach we describe now is not a permutation test, it does simulate a distribution of graphs against which the data are compared, and hence bears some similarities to QAP. It also is a procedure that in many ways is a precursor of ERGMs, discussed next.

The approach goes back to Katz and Powell (1957), who proposed a number of conditional uniform graph distributions. The simplest is the uniform distribution of graphs with a fixed number of edges (and nodes). In this distribution, each graph with the right number of edges and nodes has equal nonzero probability with all other graphs having zero probability. In effect, this is a distribution of random graphs conditional on a fixed density. It is an easy distribution from which to

simulate a sample of (say) 1,000 or more graphs. A researcher can then look at any feature of the observed data and check against the same feature from the sample. For instance, the level of reciprocity may be of interest in directed graphs. The distribution of the number of reciprocated ties from the sample can be compared with the observed data. If the number of reciprocated ties is extreme in the distribution from the sample, there is evidence for a specific effect of reciprocation. In other words, the null hypothesis that the presence of reciprocated ties in the data can be explained by a random process is rejected.

This bootstrapping method can be useful in examining the presence of local network patterns in the data (the patterns are often referred to as *configurations* or *motifs*). The approach has a well-established tradition in sociological network methodology but recently has also been introduced more widely across network science (Milo et al., 2002).

ERGMs

ERGMs are basically models for network structure and social selection effects.⁴ They are typically used with cross-sectional network data, although longitudinal versions have been developed. Using these models, one can infer whether a configuration occurs in the network more than expected by chance, given the other effects in the model. For example, if a parameter for triangles is included in the model along with other configurations such as edges and 2-stars (see Figure 2), one can infer whether there are more triangles in the network than expected simply from the number of edges and 2-stars. For a given number of edges, a certain number of triangles will occur just by chance, similarly for a given number of 2-stars. It is not enough to infer the presence of a closure process only by observing a number of triangles. Rather, inference needs to be based on a large proportion of 2-stars closed into triangles and more than expected given the density.

The technical and mathematical details of ERGMs have been covered in a number of recent reviews (e.g., Robins, 2011), so here we concentrate on a broad, intuitive description. Parameters in the models relate to network configurations akin to those we presented in Figures 1 and 2. Statistics for ERGMs are basically the counts of these configurations in the observed data. It is assumed that the observed data come from a *graph distribution*, and parameters are estimated based on maximum likelihood criteria. Because of the assumed dependence in the data, however, maximum likelihood estimates have to be obtained through computer simulation procedures.⁵ A positive (negative) parameter estimate indicates that there are more (fewer) of that configuration in the network than expected by chance, conditional on other effects in the model.

With an ERGM, it is important to step back from the dependent/independent variable distinction so familiar in regressions and other general linear model techniques. The unit of analysis in an ERGM is the network tie variable, so in one sense each network tie is the dependent variable. However, the presence of a tie may be predicted from the presence of other ties, so at one point in the model the network tie is an outcome and at another point the same variable is a predictor. This arises

because of the dependencies among the tie variables. The result is that an ERGM in effect looks for small network patterns—the configurations—in the data. A parameter for a configuration expresses how prevalent this pattern is in the network and that inference provides the interpretation. This is where any analogy with regression breaks down. In a standard regression, a parameter expresses the increase in the predictor that is required to increase the outcome variable by a set amount, holding other variables constant. Because the predictor/outcome variable distinction breaks down, this interpretation is not readily available for ERGMs.⁶

Recall earlier we described the Berardo and Scholz (2010) risk hypothesis, where to minimize risk, actors might prefer closed network structures as in the triangle in Figure 2. An ERGM can provide evidence that triangle configurations are more prevalent in the data than expected by chance (conditional of course on the other effects in the model). So a strong positive triangle parameter would indicate triangulation in the data, hence if we follow Berardo and Scholz, one interpretation might be that this network invokes collaboration that is seen as risky by the actors.

However, sometimes additional data are used as exogenous to the network to assist in predicting the presence of network ties. These may include attributes as well as dyadic covariates (other types of networks, or pairwise measures such as spatial distance between actor locations). When some data are exogenous, then the predictor/outcome distinction is maintained, and we can use analogs of some traditional techniques to aid interpretation for these exogenous effects (e.g., below we illustrate an odds ratio analysis using attribute variables.)

An appealing feature of ERGMs is that they explicitly model dependence among network tie variables. There are various dependence assumptions that can be applied. The current state of the art is to use *social circuit dependence* proposed by Snijders, Pattison, Robins, and Handcock (2006), which is most likely to produce converged parameter estimates in nondegenerate models.⁷ Social circuit parameters include effects for star configurations (degrees) as in Figure 2; but instead of a parameter for a single triangle, incorporate effects for *multiple* or *k-triangles* as depicted in Figure 3. The use of these closed, triangular configurations, with triangles “stacked” on one another, helps technically with modeling larger, denser regions of the network. The *k-triangle* configuration is accompanied by a lower order *k-2path*, which represents multiple local connectivity in the network (Figure 3).

Counts for the three families of configurations—*k-stars*, *k-triangles*, and *k-2paths*—are combined into the one statistic for each family in the way described by Snijders et al. (2006). These are the so-called *alternating* statistics, one for each family. The combining of statistics in this way simplifies the model and controls against model degeneracy. The result is that a basic social circuit model for an undirected network might contain four parameters.

- An *edge* parameter, representing the underlying propensity to form ties, and analogous to an intercept in regression.
- An *activity spread* or *centralization* parameter, based on the *k-star* (degree) effects (alternating stars). A positive parameter estimate suggests skew in the degree

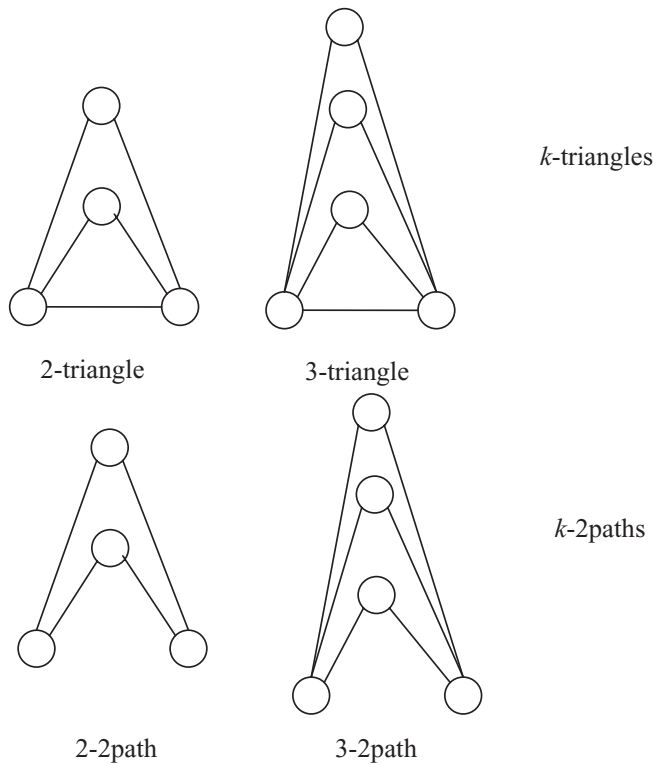


Figure 3. Some Configurations for Exponential Random Graph Model Social Circuit Models.

distribution, whereas a negative estimate indicates a more homogeneous sharing of degrees among nodes.

- A *multiple triangulation* or *closure* parameter based on the *k*-triangle configurations (alternating triangles), with a positive estimate suggesting closure effects in the network.
- A *multiple connectivity* parameter based on the *k*-2path configurations (alternating 2paths), with a positive estimate suggesting nonclosure and the presence of structural holes in the network (network brokerage).⁸

Social selection effects can also be included in an ERGM, for instance the homophily configuration in Figure 1. Additionally, it is often desirable to include basic *attribute activity* effects to represent whether actors with a particular attribute tend to express more ties than other actors. If one class of actors is more active, there will be more ties among them anyway, and there may be no need to postulate a specific homophily effect. Accordingly, one can use ERGMs to assess whether there is a homophily process over and above a simple activity effect.

Attribute effects are sometimes included in models as controls to see whether there are residual structural effects present. Sometimes they are the center of research attention, when the structural effects can then be viewed as controls for the

dependencies in observations, permitting principled inference about the association between attributes and ties (e.g., as in homophily). Once the structural effects are controlled in this way, it is possible to undertake a post hoc analysis, akin to an odds ratio analysis from a logistic regression, to aid interpretation of the attribute effects. We illustrate how to do this in our example.

The effects we have described in this section are basically for undirected graphs. With directed graphs, there is a natural increase in the number of possible configurations. Most fundamentally, there is now a reciprocity parameter in the model to capture tendencies for reciprocated ties within dyads. There are separate alternating star parameters for the in- and out-degree distributions, representing centralization in network popularity and activity, respectively. The number of closure parameters also increases, given that triadic configurations now have direction on the ties; most importantly, transitive triadic structures can be distinguished from cyclic triads. These and other effects are described and interpreted by Robins et al. (2009).

Once an ERGM has been fitted to network data, it is possible to simulate from the parameter estimates to obtain a sample of graphs consistent with the model. Then features of these simulated graphs can be compared with features of the data to see how well the model can reproduce them (Hunter, Goodreau, & Handcock, 2008). This is a diagnostic goodness of fit procedure that tells us which features of the observed data are not well explained by the model and may indicate the type of additional effect to include in a subsequent model to improve fit.

Because ERGMs model quite complex dependencies, they do not always result in converged parameter estimates, or models may in some other way be degenerate or near-degenerate (see Snijders et al. [2006] for a discussion). In these cases, the model is not adequate to the data in some way and needs to be adjusted (see footnote 7). So, it is important to ensure that there has been successful convergence. If it is not possible to obtain convergence, then sometimes it may be suitable to use a simpler convergent ERGM (with fewer effects) as a null distribution of graphs against which to compare the data and draw inferences about particular structural features. This is analogous to the conditional uniform graph distribution approach discussed above.

ERGMs have been applied to a wide range of relational data, including bipartite networks (Wang, Sharpe, Pattison, & Robins, 2009) and multiple networks (Pattison & Wasserman, 1999). Early work on ERGMs involving missing network data and on snowball sampling of networks has shown considerable potential (Handcock & Gile, 2010; Koskinen, Robins, & Pattison, 2010).

Influence Models

The models we have described are for social selection. As we noted above, it may be difficult to distinguish selection from influence in cross-sectional data. There are influence extensions of the ERGM approach, however. Autologistic actor attribute models model interdependent actor attributes Y_i as a function of exogenous tie variables X_{ij} (and possibly other exogenous attribute variables and dyadic covariates), whereas an ERGM models the interdependent tie variables X_{ij} (Daraganova & Robins, in press).

Autologistic approaches to influence processes are not the only possibility. There is a long tradition in network methodology of using autoregressive and other models for similar purposes (e.g., Friedkin, 1998; Friedkin & Johnsen, 1990, 1997; Valente, 1995; see Mason, Conrey, & Smith, 2007, for a review).

An ERGM Example: A Strategic Information Network in a Local Government

The topic of innovation is most often approached as one that relates to the private sector, and especially to corporations with high levels of flexibility and the freedom to pursue big ideas without the layers of oversight that accompany government and publicly funded services. Where government is discussed in relation to innovation, it is generally about its role in stifling innovation, and in policy studies, there is often examination of what governments should do to encourage others to innovate. The topic of governmental innovation (or how governments innovate) has had less attention paid to it and is of interest in its own right, with a number of unresolved issues, such as the kinds of actors that are involved and the different contributions that politicians and bureaucrats (insiders) and outsiders make as innovators. A study of 11 local governments in Australia examined innovation inside government, with networks expected to be a key variable. This study is briefly summarized here, and the network data collected from one of these governments, referred to as Parkside (a pseudonym) is used to demonstrate the insights that can be generated from ERGMs. The full study is reported in Considine, Lewis, and Alexander (2009).

There are many lines of inquiry that could be taken to this rather broad and little researched topic. In this study, four aspects were central. The first was one of normative frames, related to how those inside government understand the notion of innovation. The second was about the perceived impacts that the procedures of government have on innovation. The third was the familiar notion of hierarchy in organizations, and the fourth was interpersonal networks, in contrast to the formal structural relationships between individuals. Normative frames, the procedures of government, and the effects of hierarchy are all seen as important in shaping an individual's ability to achieve things in organizations through his/her position. Networks provide a way forward in examining power relations shaped by informal patterns of interaction rather than formal positional power, which is the often the starting point in policy studies.

Networks provide access to embedded resources (Lin, 2001). As others have argued, the ways in which different groups and different types of network ties matter are important to innovation adoption (e.g., Coleman, Katz, & Menzel, 1966; Rogers & Kincaid, 1981; Valente, 1995). It was expected that the collection of network data and its analysis would contribute substantially to understanding innovation inside government. In the context of this article, the data collected for this study are used to analyze networks inside government, and the other three dimensions—innovation norms, innovation procedures, and formal hierarchical positions—are not directly addressed, although as will be shown, hierarchy is impossible to ignore in this case.

If we examine formal positions and roles, a reasonably common approach to locating important actors in studies of the policy process, we can discover who has power by sitting on committees and having particular portfolio responsibilities and who is more senior in the hierarchy. But who actually interacts with whom, and how does this affect policy? Actors are connected into relationships through networks, which significantly shape political power (Knoke, 1990) and policy choices (Laumann & Knoke, 1987). Networks of influential actors have access to a set of resources that can be used to wield informal power within a policy sector (Lewis, 2006) and to significantly shape the policy agenda (Lewis, 2005). In one attempt to integrate studies of policy diffusion processes with considerations of policy networks, Mintrom and Vergari (1998) demonstrated the importance of different types of networks for different phases of innovation. "Entrepreneurs" (akin to the innovators here) used external and internal networks to get new items onto the policy agenda (external for generating new ideas from elsewhere, internal for shaping proposals so that they gain attention) and internal networks to get the required approval for the innovation.

The data for this study were collected from 11 municipalities in the state of Victoria, Australia. A call for expressions of interest to participate in this research went out through the regular newsletter of an umbrella industry association for local governments, community groups, and individual citizens. In response, 11 municipalities volunteered to participate, and they were all included.⁹ The top four levels of bureaucrats (the Chief Executive Officer, directors, managers, and team leaders/coordinators) at each of the 11 participating local governments were surveyed using a self-completed questionnaire. The senior bureaucrats referred to in the analysis below are the CEOs and directors, and the junior bureaucrats are the managers and team leaders/coordinators. All the politicians in each municipality were invited to participate. The questionnaire was distributed either at meetings of staff, through the internal mail system, or in person by the research team.¹⁰ Overall, 765 responses were received (an 81 percent response rate). For Parkside, 89 of the 102 eligible bureaucrats completed the survey, as did 5 of the 7 politicians (86 percent). We have ignored the missing nodes for this illustrative analysis, a not uncommon approach in network analysis. There are nevertheless problems in doing so, and ERGM methods for dealing with missing network data are being developed (Handcock & Gile, 2010; Koskinen et al., 2010), but we leave these complexities to one side for the purposes of this article.

Network information was collected using a name generator (open nominations), with respondents asked to nominate up to five people they went to most, first when they wanted to get advice on a work-related issue and second when they wanted to get strategic information about something in their government. The strategic information question is the one that is used in the following analysis. For this question, the nominations were limited to people from within their own organization. Network ties were recorded simply as present or absent, with no attempt to measure weighting of ties. The network tie data are directed based on who nominated whom but has been symmetrized to create an undirected network (for purposes of a simpler exposition here).

Following the first survey-based stage of this study, four of these 11 governments were then selected to participate in a more detailed second stage of the research. We conducted follow-up interviews with 104 respondents from across these four governments to gain a more detailed understanding of the nature of innovation at each location. Parkside was one of these four, and 25 people (politicians, and senior and junior bureaucrats) were interviewed in that government. Respondents were asked to identify important innovations in their municipality to detail who was involved in each case, and to nominate the key innovators in their municipality.¹¹ No limitations were placed on the number of “key innovator” nominations that could be made, and these could include politicians, bureaucrats, or people outside the government.¹²

To enable us to identify the innovators at each municipality, a “key innovator” score was calculated. Using these nominations, politicians and bureaucrats at each municipality who had completed the initial survey were assigned a score that is the percentage of total “key innovator” nominations they received from the respondents in their municipality. Those who were not nominated at all were assigned a score of zero. In Parkside, scores on this scale ranged from 0 to 12.4, with a mean score of 1.1 and a standard deviation of 2.2. Key innovators in the models presented below are those who received two or more nominations, so we use a simple binary indicator of innovation, again for the purposes of simplicity. It would also be possible to develop models that used the innovator score as a continuous actor attribute variable.

We fitted three models with an increasing number of actor attribute effects. All models contain the following social circuit structural parameters: edge, activity spread, closure, and multiple connectivity. For each actor attribute variable, we include homophily and activity effects. Parameter estimates for all three models are in Table 1.

Our first model included the structural effects plus the attribute parameters for the innovator variable (KI = “key innovator”) because of the importance of innovation to the study. In terms of structural effects, we see the following:

Table 1. Parameter Estimates (Standard Errors in Brackets) for Three Exponential Random Graph Models for the Strategic Information Network

Parameter	Model 1	Model 2	Model 3
Structural			
Edge	-3.44 (0.59)*	-3.42 (0.61)*	-3.16 (0.64)*
Activity spread (centralization)	-0.31 (0.19)	-0.28 (0.19)	-0.15 (0.21)
Multiple triangulation (closure)	1.06 (0.12)*	1.04 (0.11)*	0.88 (0.11)*
Multiple connectivity	0.02 (0.02)	0.01 (0.02)	-0.08 (0.03)*
Attributes			
Innovator activity	0.13 (0.13)	0.22 (0.14)	0.16 (0.18)
Innovator homophily	0.84 (0.32)*	0.76 (0.34)*	0.63 (0.34)
Politician activity		-0.31 (0.16)	-0.15 (0.19)
Politician homophily		1.18 (0.78)	1.31 (0.78)
Senior bureaucrat activity			1.39 (0.28)*
Senior bureaucrat homophily			-1.93 (0.77)*

*Indicates significant estimate.

- a significantly negative edge estimate, which simply indicates that isolated ties are relatively unlikely (this is akin to an intercept term in a regression);
- a nonsignificant activity spread estimate, suggesting that conditional on the other effects, there is no tendency for a highly centralized network (in other words, to the extent there is centralization, it must be due to triangulation or attribute effects); in short, there is no evidence for preferential attachment in this data;
- a positive and significant estimate for closure: this is a network built around collaboration and control (Berardo & Scholz, 2010; Lubell et al., 2011); and
- a miniscule effect for multiple connectivity: there is little evidence for substantial network brokerage.

In terms of the KI attribute, we see significant and positive homophily and a non-significant activity effect. So while this model suggests that key innovators do not have more communication partners than others, they do tend to communicate more with other innovators.

For binary attributes, a simple post hoc analysis of conditional odds ratios can aid interpretation, analogous to the odds ratio calculations available in logistic regression procedures. The difference here is that we are using the attribute estimates to calculate the conditional odds of a tie being present for a pair of actors with given attributes, compared with the conditional odds of a tie being present when both actors score 0 on the attribute variable. In this case, the comparison category is when both actors are not innovators. So, we can calculate the conditional odds ratio of a tie between an innovator and noninnovator, and between two innovators, compared with two noninnovators. The conditionality here is equivalent to the statement that all other effects are equal, including that the comparable pairs of actors have the same network neighborhood of configurations (i.e., the same number of triangles and so on).

The calculation of the odds ratio is easy for binary attributes in undirected networks. If θ_a is the activity estimate and θ_h the homophily estimate, then the conditional odds ratio for a tie between an innovator and noninnovator is simply $\exp(\theta_a)$, and between two innovators is $\exp(2\theta_a + \theta_h)$.¹³ So for the first model, the conditional odds ratio for a tie between an innovator and a noninnovator is about 1.1; in other words, the propensity for ties between innovators and noninnovators is not much different than for ties between noninnovators. However, for ties between two innovators, the conditional odds ratio is nearly 3. So, according to this model, a tie is much more probable between two innovators than a tie involving a noninnovator. (Of course, there are more noninnovators, so there can be more noninnovator ties, but the propensity for ties between innovators is much greater.)

We do not report all the goodness of fit details here. Suffice to say that this model reproduces many graph features well but is very bad at predicting when ties involve politicians and senior bureaucrats. In the second model, we have included activity and homophily effects for politicians. The significant effects from the previous model remain. While the estimate for politician homophily seems quite large, the standard error is also substantial, so we cannot be confident that this estimate is reliably

different from zero. Goodness of fit suggests that this model is unable to predict when ties involve senior bureaucrats.

Our third and final model includes parameters for the senior bureaucrat attribute. Now we see some important changes in the estimates. The multiple connectivity estimate is now significant (and negative). The homophily effect for innovators is no longer significant, and instead we have a negative significant effect for bureaucrat homophily and a positive significant effect for bureaucrat activity. So senior bureaucrats tend to express more network ties, but—taking this into account—they tend not to communicate with each other. The odds ratio for a senior–senior bureaucrat tie is over 2 but for a senior–junior bureaucrat tie is nearly 4. Strategic communication between bureaucrats in this local government does cross divisional boundaries at the same level within the hierarchy, but it is more likely to go up and down the hierarchy.

Moreover, the disappearance of innovator homophily as a significant effect suggests that the cross-level bureaucratic communication subsumes the innovation effect. In this bureaucracy, there are four senior innovators and five junior innovators. The density of ties between these two groups is 0.153, much larger than the overall network density of 0.065. These innovators are communicating among themselves, but this is better explained by the strong effects for cross-level communication than for innovator homophily. This result emphasizes the value of having a statistical model in which a number of effects can be considered simultaneously, to protect against spurious results that may be better explained by other effects.

The appearance of a small but significant negative connectivity parameter suggests that once one controls for cross-level bureaucratic communication, there is a small tendency against network brokerage. In other words, we infer that brokerage in this network, when it occurs at all, tends to occur among the bureaucrats.

These models point to some network characteristics of this government that could be expected—there is no strong centralization around a single actor, but several clusters that reflect the formal structure, and a substantial amount of closure that is not surprising in an environment where the actors have to work together closely. They also confirm the well-established divisions between the political and administrative arms of governments and the hierarchical nature of much communication within the second of these. Most interestingly, the third model suggests that these divisions are more important than innovation status and innovator homophily. In policy theory terms, we might infer that hierarchy is important for innovation inside government, because for this local government, most communication between innovators flows up and down the hierarchy. This is in contrast to what is often claimed about innovation in the private sector—that it is more likely to occur outside formal structures.

It also leads us to a point of circularity: Are you seen to be an innovator because you hold a position that has enough inbuilt scope for you to be innovative, or do you get to hold such a position because you are recognized as innovative? To unpick such a question would require longitudinal data.

SAOMs

SAOMs (also known as stochastic actor-based models) are specifically designed for longitudinal network data.¹⁴ The original versions of SAOMs were intended to examine the dynamics of network self-organization (Snijders, 1996, 2001). More recently, these models have become a major tool for investigating the coevolution of network structures and individual behaviors, so that they are one of the few principled methods to test whether selection or influence, or both, are present in network data (Snijders, Steglich, & Schweinberger, 2007; Steglich, Snijders, & West, 2006). A recent (2010) special issue of the journal *Social Networks* has been dedicated to these models, so we restrict ourselves here to a conceptual summary of the modeling approach and refer interested readers to the special issue, especially to Snijders, van de Bunt, and Steglich (2010), which provides an excellent tutorial overview. Snijders et al. (2010) explains how SAOMs fit in with other longitudinal methods proposed for social network data.

SAOMs are used with network panel data, that is, observations of a network, as well as actor attributes, at two or more specific time points. The focus is on directed networks because actors are assumed able to make choices about where to direct their outgoing ties. Network ties are not assumed to be immediate transactions (such as an e-mail exchange or a telephone call) but a relationship of indefinite duration, albeit one that can change across the time period of the study.

Snijders et al. (2010) describe the four basic assumptions underpinning SAOM¹⁵:

- Time is continuous, so that even though the network is observed at discrete time points, changes to network ties (or to attributes) may have happened at any intervening moment between observations. This assumption permits configurations of network ties to be built up, tie by tie, between observations.
- The change in the network is a Markov process, which means that the appearance or disappearance of a network tie, or a change in attribute, depends only on the state of the network at that particular moment in time, and not on previous moments.
- Actors control their outgoing ties.
- At any moment, one randomly selected actor can change one tie, or alter one attribute. This keeps the modeling relatively simple, so that there is only one basic change at any moment.

SAOMs model the change process, and so condition on the first observation. In other words, a SAOM assumes the network at the first time point and does not attempt to model the structure of that first network. It is the change from that first observation to the second (and perhaps subsequent) networks that is modeled.

In using a SAOM, we are in effect assuming that there are many tie (and attribute) changes that we have not directly observed, in getting from the first to the second (and subsequent) observations. The model estimates parameters through simulation. Many possible sequences of changes are simulated according to certain

rules implicit in model specification. These rules are specified in the *objective function*, which determines whether a change is likely or not. The objective function can be seen as an actor's set of preferences for their structural position. For instance, suppose there is a tendency for closure in the network. Then we want an objective function where there is a relatively high probability for a tie to come into existence when it closes a 2-path into a triangle.

So, a number of effects in the objective function are chosen by the researcher in modeling a given dataset. Parameters are the weights given to these effects. To estimate parameters, an initial guess of parameter values is made, sequences of changes are simulated, and if there is a difference from the second network, the parameter values are adjusted accordingly. If the parameter values eventually converge, these become the estimates for that model. The full details of the estimation algorithm are given by Snijders (2001).

When modeling data for the coevolution of networks and behaviors, there is a separate objective function for changes in ties and for changes in attributes. Attribute variables in the objective function for tie changes are in effect social selection predictors, whereas the same variables in the objective function for attribute changes describe social influence effects. In this way, by estimating parameters for the two objective functions, influence and selection processes can be simultaneously assessed.

Some of the most fundamental parameters relate to effects we have discussed above, such as closure, degree-based processes, and homophily. In that sense, these models bear relationship to ERGMs: in fact, an ERGM can be seen as a SAOM in equilibrium (Snijders et al., 2010). However, there are also many other possible effects that can be included in a SAOM. Snijders et al. (2010) provide a list of some of the more important in their appendix.

The results from the SAOM will produce estimates for structural effects (network self-organization) as well as attribute effects for both selection and influence processes. It is not uncommon to see evidence for both processes simultaneously in empirical datasets. For instance, Snijders et al. (2010) present an example of friendship and delinquency among school children. They concluded that there was clear evidence for delinquent children to tend to become friends as well as evidence for the children to influence the delinquent behavior of their friends.

SAOMs are a principled way to model network panel data. Because they are computationally intensive in estimation, their application to very large network datasets with hundreds of nodes may be limited. Nevertheless, for very many empirical longitudinal network datasets in the social sciences, they are the preferred method.

Conclusions

Network-based research in the social sciences is increasingly popular. We need novel statistical methods to analyze empirical network data because a network conceptualization implies that there will be dependency among observations, often in complex ways that mean standard statistical approaches are not applicable. This article has described some common statistical methods that are used in the articles in

this special issue. There are, however, an increasing number of other statistical modeling techniques being developed. Even the methods we have described are experiencing ongoing methodological development, so that the current state of the art is not the final word. Statistical modeling of social networks is one of the most rapidly growing areas of network science. We can expect to see further advances into the future.

It is worthwhile concluding with a few comments about causality, which is sometimes a contentious issue in network research. In this article, we have tried to steer away from making claims about whether and how causality may be attributed. The debate about causality in networks is often related to distinguishing selection from influence. In this article, in describing SAOMs, we have implied that causal inferences are better obtained from longitudinal data, as is the case in other areas of social science, although of course this is not always definitive. Controlled experiments to infer causality, a common approach in medical and related sciences, are not always easy to design for network research. Taking parts of a social system (e.g., individuals or dyads) into a laboratory removes their systemic embeddedness and thereby undermines inference. Complex systems science provides a different perspective: if networks represent a complex adaptive social system, the notion of traditional causality may be inappropriate; rather, the system evolves according to its own internal dynamics (this is a perspective that fits in quite well with ERGM and SAOM approaches). As a result, for the purposes of this article, we have been deliberately agnostic about relating the specific methods to causal inference. These issues go beyond network studies and apply to social science more generally. Readers will have their own points of view. For what it is worth, our own position is that, if it is necessary to infer causality, this is best done in a programmatic way, with a number of studies reinforcing causal interpretations in different ways.

The statistical modeling of social networks has been little used in policy studies, which is rather surprising given the rise of research interest in policy networks and network governance since the 1990s and the increase in network forms of governance in many countries over the last two decades (Lewis, 2011). There have been relatively few attempts to map and understand the network structures that exist inside governments and play an important role in shaping policymaking, although these have increased in the last 5 years. There are even fewer attempts to model these networks and test policy theory. Research on the many networks introduced by governments and others to coordinate activities of many kinds across public, private, and nonprofit sectoral boundaries is burgeoning as a descriptive enterprise. Statistical network models promise to help illuminate the policy process by moving from description to theory testing. Such advances should improve the analysis of policy, and so will surely continue to grow.

Garry Robins, PhD, is a mathematical psychologist and social network methodologist in the School of Psychological Sciences at the University of Melbourne, Australia. His research has concentrated on developing exponential random graph models for social networks but he is also involved in a wide range of empirical social network projects.

Jenny M. Lewis, PhD, is professor of public administration and public policy in the Department of Society and Globalisation at Roskilde University, Denmark. Her research is concentrated on the policy process, new forms of governance, and the consequences of performance measurement. Her most recent book is *Connecting and Cooperating: Social Capital and Public Policy*.

Peng Wang is a research fellow in the School of Psychological Sciences at the University of Melbourne, Australia. His research focuses on the development of exponential random graph models for bipartite and multilevel networks, and the implementation and optimization of the software package PNet for statistical analysis of social networks.

Notes

1. This approach amounts to pseudo-likelihood estimation for an ERGM-based social influence model (Robins, Pattison, & Elliott, 2001). There has been plenty of work over the last 10 years on pseudo-likelihood estimation for ERGMs. The best that can be said is that it may be a first approximation. Van Duijn, Gile, and Handcock (2009) concluded unequivocally that it is always better to use the maximum likelihood estimate (through the estimation procedures described below) than pseudo-likelihood estimation even when it is corrected for bias.
2. The QAP procedure is most commonly applied to unipartite networks, but there is no reason why it could not be extended to bipartite data structures.
3. Of course, one needs to pick an association statistic that is appropriate to the measurement of the network ties. For instance, a Pearson correlation would be best applied to valued rather than binary data.
4. There are also social influence versions of these models, sometimes called Autologistic actor attribute models (Daraganova & Robins, in press), which we briefly describe below.
5. The so-called Markov chain Monte Carlo maximum likelihood estimation.
6. Nevertheless, it is possible to compute a conditional log-odds or probability of a tie being present, conditional on the neighborhood of other ties that are present in the data. However, it is not possible to (say) consider an increase in the number of triangles in a neighborhood without increasing other tie configurations as well, so this is not always that useful as a guide to interpretation.
7. The previously popular Markov random graph models (Frank & Strauss, 1986) are problematic in a number of ways and in most circumstances should be avoided. Even though a social circuit parameterization is much to be preferred, there is no guarantee that a social circuit model will be nondegenerate, so there needs to be careful assessment of the estimation procedure (e.g., Robins et al., 2009). Degeneracy in ERGMs is a topic in its own right: suffice to say that parameter estimates may not converge, or some sets of parameter estimates might imply simple graphs (e.g., full or empty) that do not resemble the data. In such a case, the parameterization is not correct for the data and an alternative model specification could be tried.
8. In the statnet program, the alternating parameters and statistics have a slightly different form and are referred to as geometrically weighted (GW): the GW degree distribution, GW edgewise shared partner distribution, and GW dyadwise shared partner distribution parameters, respectively. See Snijders et al. (2006) and Hunter (2007).
9. Taking volunteers in this fashion skews the sample toward those more interested in this topic. Because the purpose was to study innovation, this bias represents no threat, although it limits how far the findings can be generalized to governments with little interest in innovation.
10. The sample group was identified by an internal liaison officer at each council based on internal staff lists. Response rates were maximized by a minimum of two follow-up calls to all potential respondents within the sample.
11. Nomination-based methods for locating key actors have attracted criticism for being highly subjective. When dealing with actors who work in close proximity to one another, it is reasonable to expect that they will be well informed about one another's reputations for work-related performance. This

assumption was supported by the close correlation between “key innovator” nominations and the prominence of those identified as important actors in the 16 innovation case studies examined as part of the larger study.

12. In each of the four local governments, no nominations were directed to actors outside.
13. The factor of 2 arises because the formation of a tie between two innovators creates one homophily configuration and two activity configurations.
14. Because these models are implemented in the SIENA software (Ripley, Snijders, & Preciado, 2011), they are sometimes also called SIENA models.
15. For modeling coevolution of networks and behaviors, the assumptions are slightly more complicated but nevertheless are a natural extension of these four basic ideas.

References

- Albert, Reka, and Albert-Laszlo Barabási. 2002. “Statistical Mechanics of Complex Networks.” *Review of Modern Physics* 74: 47–97.
- Barabási, Albert-Laszlo, and Reka Albert. 1999. “Emergence of Scaling in Random Networks.” *Science* 286: 509–12.
- Berardo, Ramiro, and John T. Scholz. 2010. “Self-Organizing Policy Networks: Risk, Partner Selection, and Cooperation in Estuaries.” *American Journal of Political Science* 54: 632–49.
- Borgatti, Stephen P., Martin G. Everett, and Lin C. Freeman. 2002. *Ucinet 6 for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.
- Burt, Ronald S. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.
- Butts, Carter T. 2008. “Social Network Analysis: A Methodological Introduction.” *Asian Journal of Social Psychology* 11: 13–41.
- Carrington, Peter, John Scott, and Stanley Wasserman, eds. 2005. *Models and Methods in Social Network Analysis*. New York: Cambridge University Press.
- Coleman, James S., Elihu Katz, and Herbert Menzel. 1966. *Medical Innovation: A Diffusion Study*. New York: Bobbs Merrill.
- Considine, Mark, Jennifer M. Lewis, and Damon Alexander. 2009. *Networks, Innovation and Public Policy: Politicians, Bureaucrats and the Pathways to Change inside Government*. Houndmills, UK: Palgrave Macmillan.
- Daraganova, Galina, and Garry Robins. in press. “Autologistic Actor Attribute Models.” In *Exponential Random Graph Models for Social Networks: Theories, Methods and Applications*, ed. Dean Lusher, Johan Koskinen, and Garry Robins. New York: Cambridge University Press.
- Davis, James A. 1970. “Clustering and Hierarchy in Interpersonal Relations: Testing Two Theoretical Models on 742 Sociograms.” *American Sociological Review* 35: 843–52.
- Dekker, David, David Krackhardt, and Tom A. B. Snijders. 2007. “Sensitivity of MRQAP Tests to Collinearity and Autocorrelation Conditions.” *Psychometrika* 72: 563–81.
- Frank, Ove, and David Strauss. 1986. “Markov Graphs.” *Journal of the American Statistical Association* 81: 832–42.
- Friedkin, Noah E. 1998. *A structural theory of social influence*. New York: Cambridge University Press.
- Friedkin, Noah, and Eugene Johnsen. 1990. Social influence and opinions. *Journal of Mathematical Sociology* 15: 193–205.
- . 1997. Social positions in influence networks. *Social Networks* 19: 210–22.
- Granovetter, Mark. 1973. “The Strength of Weak Ties.” *American Journal of Sociology* 81: 1287–303.
- Handcock, Mark, and Krista Gile. 2010. “Modeling Networks from Sampled Data.” *Annals of Applied Statistics* 4: 5–25.
- Hubert, Lawrence J. 1987. *Assignment Methods in Combinatorial Data Analysis*. New York: Dekker.

- Hubert, Lawrence J., and James Schultz. 1976. "Quadratic Assignment as a General Data Analysis Strategy." *British Journal of Mathematical and Statistical Psychology* 29: 190–241.
- Hunter, David. 2007. "Curved Exponential Family Models for Social Networks." *Social Networks* 29: 216–30.
- Hunter, David, Steven Goodreau, and Mark Handcock. 2008. "Goodness of Fit of Social Network Models." *Journal of the American Statistical Association* 103: 248–58.
- Jackson, Matthew. 2010. *Social and Economic Networks*. Princeton, NJ: Princeton University Press.
- Jones, James, and Mark Handcock. 2003. "An Assessment of Preferential Attachment as a Mechanism for Human Sexual Network Formation." *Proceedings of the Royal Society, B* 270: 1123–28.
- Kalish, Yuval. 2008. "Bridging in Social Networks: Who Are the People in Structural Holes and Why Are They There?" *Asian Journal of Social Psychology* 11 (1): 53–66.
- Katz, Leo, and James Powell. 1957. Probability distributions of random variables associated with a structure of the sample space of sociometric investigations. *Annals of Mathematical Statistics* 28: 442–48.
- Knoke, David. 1990. *Political Networks: The Structural Perspective*. New York: Cambridge.
- Kolaczyk, Eric. 2009. *Statistical Analysis of Network Data: Methods and Models*. New York: Springer.
- Koskinen, Johan, Garry Robins, and Philippa Pattison. 2010. "Analysing Exponential Random Graph (P*) Models with Missing Data Using Bayesian Data Augmentation." *Statistical Methodology* 7: 366–84.
- Krackhardt, David. 1987. "QAP Partialling as a Test of Spuriousness." *Social Networks* 9: 171–86.
- . 1988. "Predicting with Networks: Nonparametric Regression Analysis of Dyadic Data." *Social Networks* 10: 359–81.
- Laumann, Edward, and David Knoke. 1987. *The Organizational State: Social Choice in National Policy Domains*. Madison: University of Wisconsin Press.
- Lewis, Jennifer M. 2005. *Health Policy and Politics: Networks, Ideas and Power*. Melbourne: IP Communications.
- . 2006. "Being around and Knowing the Players: Networks of Influence in Health Policy." *Social Science and Medicine* 62: 2125–36.
- . 2011. "The Future of Network Governance Research: Strength in Diversity and Synthesis." *Public Administration* 89: 1221–34.
- Li, Lun, David Alderson, Reiko Tanaka, John Doyle, and Walter Willinger. 2005. *Towards A Theory of Scale-Free Graphs: Definition, Properties, and Implications (Extended Version)*. Technical Report CIT-CDS-04-006. Pasadena: California Institute of Technology, Engineering and Applied Sciences Division.
- Lin, Nan. 2001. *Social Capital: A Theory of Social Structure and Action*. Cambridge: Cambridge University Press.
- Lubell, Mark, Garry Robins, and Peng Wang. 2011. *Policy Coordination in an Ecology of Water Management Games*. Fourth Annual Political Networks Conference, Ann Arbor, MI. http://opensiuc.lib.siu.edu/pnconfs_2011/22/. Accessed August 11, 2011.
- Mantel, Nathan. 1967. "The Detection of Disease Clustering and a Generalized Regression Approach." *Cancer Research* 27: 209–20.
- Mason, Winter, Frederica Conrey, and Eliot Smith. 2007. "Situating Social Influence Processes: Dynamic, Multidirectional Flows of Influence Within Social Network." *Personality and Social Psychology Review* 11: 279–300.
- Milo, Ron, Shai Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. 2002. "Network Motifs: Simple Building Blocks of Complex Networks." *Science* 298: 824–27.
- Mintrom, Michael, and Sandra Vergari. 1998. "Policy Networks and Innovation Diffusion: The Case of State Education Reform." *The Journal of Politics* 60: 126–48.
- Moreno, Jacob, and Helen Jennings. 1938. "Statistics of Social Configurations." *Sociometry* 1: 342–74.
- Newman, Mark. 2010. *Networks: An Introduction*. New York: Oxford University Press.
- Padgett, John F., and Christopher K. Ansell. 1993. "Robust Action and the Rise of the Medici, 1400–1434." *American Journal of Sociology* 98: 1259–319.

- Pattison, Philippa, and Garry L. Robins. 2008. "Probabilistic Network Theory." In *Handbook of Probability: Theory and Applications*, ed. T. Rudas. Thousand Oaks, CA: Sage, 291–312.
- Pattison, Philippa, and Stanley Wasserman. 1999. "Logit Models and Logistic Regressions for Social Networks, II. Multivariate Relations." *British Journal of Mathematical and Statistical Psychology* 52: 169–94.
- Ripley, Ruth M., Tom A. B. Snijders, and Paulina Preciado. 2011. *Manual for SIENA Version 4.0*. Oxford: University of Oxford, Department of Statistics, Nuffield College.
- Robins, Garry. 2009. "Understanding Individual Behaviors within Covert Networks: The Interplay of Individual Qualities, Psychological Predispositions, and Network Effects." *Trends in Organized Crime* 12: 166–87.
- . 2011. "Exponential Random Graph Models for Social Networks." In *Handbook of Social Network Analysis*, ed. J. Scott and P. Carrington. London: Sage, 484–500.
- Robins, Garry, Philippa Pattison, and Peng Wang. 2009. "Closure, Connectivity and Degrees: New Specifications for Exponential Random Graph (P^*) Models for Directed Social Networks." *Social Networks* 31: 105–17.
- Robins, Garry L., Philippa Pattison, and Peter Elliott. 2001. "Network Models for Social Influence Processes." *Psychometrika* 66: 161–90.
- Rogers, Everett M., and D. Lawrence Kincaid. 1981. *Communication Networks: A New Paradigm for Research*. New York: Free Press.
- Scholz, John T., Meredith Whiteman, and Ramiro Berardo. 2011. *Network Capital and the Self-Organizing Provision of Public Goods*. Working paper. Florida State University.
- Scott, John, and Peter Carrington, eds. 2011. *The Sage Handbook of Social Network Analysis*. London: Sage.
- Simmel, Georg. 1908. "Soziologie, Untersuchungen Uber Die Formen Der Vergesellschaftung." In *The Sociology of Georg Simmel*, ed., trans. K. H. Wolff. (1950). New York: Free Press, 87–408.
- Snijders, Tom A. B. 1996. "Stochastic Actor-Oriented Dynamic Network Analysis." *Journal of Mathematical Sociology* 21: 149–72.
- . 2001. "The Statistical Evaluation of Social Network Dynamics." *Sociological Methodology* 31: 361–95.
- . 2011. "Statistical Models for Social Networks." *Annual Review of Sociology* 37: 131–53.
- Snijders, Tom A. B., Philippa Pattison, Garry L. Robins, and Mark Handcock. 2006. "New Specifications for Exponential Random Graph Models." *Sociological Methodology* 36: 99–153.
- Snijders, Tom A. B., Christian E. G. Steglich, and Michael Schweinberger. 2007. "Modeling the Co-Evolution of Networks and Behavior." In *Longitudinal Models in the Behavioral and Related Sciences*, eds. K. van Montfort, H. Oud, and A. Satorra. Mahwah, NJ: Lawrence Erlbaum, 41–71.
- Snijders, Tom A. B., Gerhard G. van de Bunt, and Christian E. G. Steglich. 2010. "Introduction to Stochastic Actor-Based Models for Network Dynamics." *Social Networks* 32: 44–60.
- Steglich, Christian E. G., Tom A. B. Snijders, and Patrick West. 2006. "Applying SIENA: An Illustrative Analysis of the Coevolution of Adolescents' Friendship Networks, Taste in Music, and Alcohol Consumption." *Methodology* 2: 48–56.
- Valente, Thomas. 1995. *Network Models of the Diffusion of Innovations*. Creskill, NJ: Hampton Press.
- Van Duijn, Maritje, Krista Gile, and Mark Handcock. 2009. "A Framework for Comparison of Maximum Pseudo Likelihood and Maximum Likelihood Estimation of Exponential Family Random Graph Models." *Social Networks* 31: 52–62.
- Wang, Peng, Ken Sharpe, Philippa Pattison, and Garry Robins. 2009. "Exponential Random Graph (P^*) Models for Affiliation Networks." *Social Networks* 31: 12–25.
- Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.
- Wasserman, Stanley, and Garry Robins. 2012. "Social Network Research: The Foundation of Network Science." In *APA Handbook of Research Methods in Psychology, Vol. 3*, ed. Harris Cooper et al. Washington, DC: APA Press, 451–69.