Social Networks 14 (1992) 5–61
North-Holland

Blockmodels: Interpretation and evaluation *

Katherine Faust

Department of Sociology, University of South Carolina, Columbia, SC 29208, USA

Stanley Wasserman

Department of Psychology, University of Illinois, Champaign, IL 61820, USA

Many methods for the description of social network structural properties are concerned with the dual notions of *social position* and *social role*. Common goals of these methods are to represent patterns in complex social network data in simplified form, to reveal sets of actors who are similarly embedded in networks of relations, and to describe the associations among relations in multirelational social networks. Often these representations take the form of a *blockmodel*. In a blockmodel actors are assigned to positions and network relations are presented among positions, rather than among actors.

The literature on blockmodels is extensive and is overflowing with computation and applications of blockmodels. However, there is a surprising lack of attention to two very important aspects of blockmodel analyses: the interpretation and evaluation of the results. The purpose of this paper is to focus on these topics, primarily reviewing and synthesizing the approaches to interpretation and evaluation currently in use.

Positional analysis of social network data rests on the assumption that the role structure of the group and positions of individuals in the group are apparent in the measured relations present in a set of network data. To date, most research, most methodological advances, and most applications of role-positional analyses (beginning with the pioneering work of Lorrain and White 1971) have emphasized the descriptive aspects of these methods, both to formalize social concepts and to describe the structure of particular social networks. However,

* This research was supported, in part, by grants from the National Science Foundation and the Research Board of the University of Illinois. This paper is based on Chapters 9, 10 and 18 of Wasserman and Faust (1992). We would like to thank Stephen Borgatti and Alaina Michaelson for helpful comments on an earlier draft of this paper and David Krackhardt for sharing his data with us.

integrating these notions into theories and process models requires going beyond description.

The main purpose of this paper is to discuss the interpretation and evaluation of blockmodels. We begin with a general discussion of the steps one must undertake to do a complete positional analysis, particulary the latter steps involving representations, interpretations, and assessments.

1. Overview of positional and role analysis

There are two key aspects to the positional-role analysis of social networks: identifying social positions as collections of actors who are similar in their relations with others, and modeling social roles as systems of relations among actors or among positions. These two aspects are apparent in the foundational works by White *et al.* (1976) who focused on methods for partitioning actors, and Boorman and White (1976), who focused on models for collections of relations. In practice, however, many applications of these methods to substantive problems emphasize one or the other of these tasks. Most analyses emphasize the similarity of actors (that is, the identification of positions) with considerably less attention to the relations among the positions.

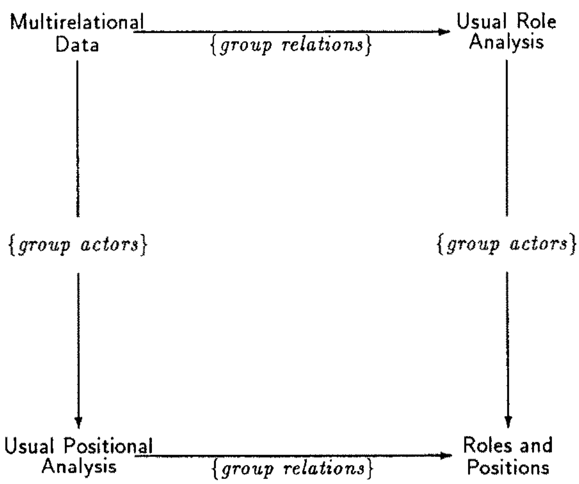

Fig. 1. An overview of positional and role analysis.

Schematically one can present the task of a full positional and role analysis of social network data as in Figure 1 (see Sailer 1978; Pattison 1982). Beginning with a set of network data consisting of a collection of relations (a multirelational data set), the ultimate goals are to "group" actors into positions based on relational similarity, and simultaneously to describe the association among relations based on how they combine to link actors or positions (White *et al.* 1976; Boorman and White 1976; Sailer 1978; Breiger and Pattison 1986; Pattison 1982, 1988). As shown in this figure, the alternative paths involve (from top to bottom) grouping actors, the standard positional analysis, and (from left to right) studying the associations among relations, the usual role analysis. A complete positional and role analysis would result in both an assignment of actors to positions and a model of the system of relations that link these positions.

Let us think about the tasks of analyzing network positions and analyzing network roles separately for the moment. We will start with one set of actors and a collection of binary relations.

First consider the positional analysis problem (the left path from top to bottom in Figure 1). The major task here is to locate sets of actors who are similar across the collection of relations. Similarity will be defined in terms of the equivalence of actors with respect to some formal mathematical property. The formal mathematical property specifies which actors will be "grouped" together in a network position. A positional analysis, the vertical path on the left side of the diagram, maps actors into equivalence classes, where an equivalence class consists of all actors who are identical (or nearly identical) on the specified mathematical property. Structural equivalence, which we define shortly, is one such formal mathematical property for defining equivalence classes.

Once equivalence classes (or positions) of actors have been identified, the relations among these positions must be described. Representing relations among positions is the task represented by the arrow from left to right on the bottom of the figure. Image matrices and density tables, as well as blockmodels (common representations of a positional analysis) are the primary tools for this task.

Now, let us consider the usual role analysis. A role analysis is concerned with the associations among relations. Schematically, a role analysis will traverse the horizontal paths in Figure 1, either along the top or along the bottom of this diagram. The distinction between the top path and the bottom path is related in part to the distinction between "global" roles, which describe associations among relations for an entire group, and "individual" or "local" roles, which describe associations among relations from the perspectives of individual actors or subsets of actors.

The path from left to right along the top of the diagram outlines one approach to role analysis. Starting with a collection of binary relations the task is to describe the association among the relations. For example, in an analysis of kinship relations, one might note that the combination of relations "mother of" and "sister of" gives rise to a meaningful compound relation – "mother's sister" which (in standard American English kinship terms) is labelled "aunt". Modeling the association among relations is the basis for the network role system (Boorman and White 1976).

The final step moving from top to bottom along the right side of the diagram requires grouping actors into equivalence sets based on the description of the role system resulting from the previous step. Here, as on the left side of the diagram, the critical decision is how to measure similarity among actors. The result is both a model of associations among relations (the network roles) and a partition of actors into equivalence classes that relate similarly to one another according to the roles.

In terms of the scheme shown in Figure 1, blockmodels are the result of the vertical path on the left side of the figure. In order to interpret and evaluate a blockmodel we assume that the researcher has already located sets of equivalent actors from multirelational data. Our purpose here is to evaluate and interpret such a grouping. We are not concerned here with the lower horizontal path: describing role systems, based on a prior aggregation of actors using blockmodels.

2. Tasks of a positional analysis

One of the major objectives of a positional analysis is to simplify the information in a network data set. This simplification consists of a representation of the network in terms of the positions identified by an equivalence definition, and a statement of how these positions are related to each other. In this section we focus on the steps that are required for a complete positional analysis.

In practice, a complete positional analysis requires at least four steps. Specifying the equivalence definition by which actors will be assigned to the same equivalence class is only the first step in a positional analysis. At the very least, we also need an assessment or evaluation of how good the representation is, as well as an interpretation of the representation. These steps are:

(1) a formal definition of equivalence;
(2) a measure of the degree to which sets of actors approach that definition in a given set of network data;
(3) a representation of the equivalences; and
(4) an assessment of the adequacy of the representation.

The equivalence definition specifies the formal mathematical conditions under which we will consider actors in a network to be equivalent.

Structural equivalence is one such equivalence definition; others are explored in Breiger and Pattison (1986), Faust (1988), Mandel (1983), Pattison (1988), Wasserman and Anderson (1987), Wasserman and Faust (1992, Ch. 12), White and Reitz (1989), Winship (1988), Winship and Mandel (1983), Wu (1983), among others. In all cases the equivalence definitions can be stated in terms of properties of relational ties among actors in a network.

However, in actual network data, it is unlikely that any actors will be exactly structurally equivalent. Therefore the second step requires measuring the extent to which actors are equivalent. Measuring equivalence allows us to decide, for any given equivalence definition, whether or not (and perhaps to what extent) sets of actors in a network are equivalent according to the given definition.

The third step in a positional analysis is *representation* of the equivalences, including the assignments of actors to equivalence classes, a statement of the relationships between and among the classes, and an interpretation of this representation. The most common kind of representation is a *discrete* model, that provides a partition of the actors in the network into a collection of classes (subsets). Another important aspect of the representation of equivalences is a statement of how the positions relate to each other. The reduced graph, image matrix, and blockmodel are examples of representations.

The fourth component of a positional analysis is the assessment of the adequacy of the representation. Sometimes the assessment of adequacy (usually called goodness-of-fit) requires probability models.

We now turn to a brief discussion of the four steps in a positional analysis, focusing primarily on the third and fourth.

2.1. Structural equivalence

The first step in a positional analysis is to adopt a formal definition of equivalence. Throughout this paper we will use *structural equivalence* (Lorrain and White 1971). Other papers in this volume discuss blockmodels for other equivalences, such as regular equivalence (Everett and Borgatti 1992; Batagelj *et al.* 1992b), or stochastic equivalence (Wasserman and Anderson 1987; Anderson *et al.* 1992). Actors in a social network are structurally equivalent if they have identical relational ties to and from all other actors in a network. We will use the notation, $\mathscr{B}_k$ to denote subsets of equivalent (or approximately equivalent) actors.

More precisely, assume we have a collection $\mathscr{R} = \{\chi_1, \chi_2, \ldots, \chi_R\}$, containing R binary relations (indexed by $r = 1, 2, \ldots, R$). We will denote the presence of a tie between actors i and j on relation χ_r as $i \xrightarrow{\chi_r} j$.

Formally, actors i and j are *structurally equivalent* if for all actors, $k = 1, 2, \ldots, g$, $k \neq i, j$, actor i has a relational tie to k, if and only if j also has a relational tie to k, and i has a relational tie from k if and only if j also has a relational tie from k. In other words, i and j are structurally equivalent if $i \xrightarrow{\chi_r} k$ if and only if $j \xrightarrow{\chi_r} k$ and $k \xrightarrow{\chi_r} i$ if and only if $k \xrightarrow{\chi_r} j$, for actors, $k = 1, 2, \ldots, g$, $k \neq i, j$, and relations $r = 1, 2, \ldots, R$.

Alternatively, the definition may be expressed using the more usual sociometric notation. Letting x_{ijr} be the binary variable indicating the presence or absence of a relational tie from actor i to actor j on relation χ_r, then actors i and j are structurally equivalent if $x_{ikr} = x_{jkr}$ and $x_{kir} = x_{kjr}$ for $k = 1, 2, \ldots, g$, $k \neq i, j$, and $r = 1, 2, \ldots, R$. If actors i and j are structurally equivalent, then directed ties from i terminate at exactly the same actors as directed ties from j, and directed ties to i originate from the same actors as the directed ties to j.

2.2. Measuring equivalence

Actors in a social network are almost never structurally equivalent. The second task of a positional analysis is a measure of the degree to which pairs or sets of actors approach structural equivalence.

Doreian (1988) makes the useful distinction between the equivalence definition, and the procedure for detecting the property of equivalence (a "detector"). Pattison (1988) makes a similar distinction, between the "model" and the algorithm for fitting the model to data. In the following sections we will refer to the detector as a measure of equivalence.

Measuring degree of structural equivalence is a problem of measuring the similarity (or dissimilarity) of the relational ties to and from pairs of actors on a given set of network data. It is, therefore, a specific instance of a more general issue of measurement of the similarity (or dissimilarity) of two data "profiles". In analyzing network data, the "profiles" are the rows and columns in the sociomatrices corresponding to two actors' relational ties. Numerous authors, both inside and outside the network community, have examined the relationships among alternative measures of similarity and dissimilarity. The relationship between correlation and Euclidean distance is well known (Cronbach and Gleser 1953; Coxon 1982; Fox 1982; Rohlf and Sokal 1965; Sneath and Sokal 1973; Sokal and Sneath 1963).

Measures such as correlation or Euclidean distance, that are commonly used to measure structural equivalence, do not always give the same results. The correlation between two actors may be equal to $+1$, indicating perfect structural equivalence by that measure, while the Euclidean distance between the same two actors on the same relation(s), may be non-zero, indicating that the actors are not perfectly structurally equivalent. Therefore, as detectors of structural equivalence, correlation and Euclidean distance differ in the ways that two actors in a social network may fail to have identical ties, and therefore not be measured as structurally equivalent. It is important to understand the formal properties of these measures in order to make an appropriate choice for a given application (see Faust and Romney 1985a, Burt 1986; and Wasserman and Faust 1992 for further discussion of this issue).

We simply assume that some suitable measure of equivalence has been used and actors will be asigned to positions based on this measure. So, the result of step two might be a $g \times g$ matrix of measures of pairwise actor equivalence (such as the correlations between actors rows and columns in the sociomatrices being studied).

2.3. Representing equivalences

The third step in a positional analysis is representation of the positions, and a statement of how the positions are related to each other. The major goals of the representation are to present the information in a network data set in simplified form, and provide an interpretation for the results. The key step is to partition actors into subsets, so that actors within each subset are closer to being equivalent than are actors in different subsets. Subsets of (nearly) equivalent actors will be referred to as equivalence classes, or *positions*. This partitioning then leads to the problem of deciding how the subsets relate to each other. A density table (or density matrix), an image matrix, and a reduced graph are three ways to present the relational ties among positions.

2.3.1. Partitioning actors

If we look at the matrix of distances or correlations that measure structural equivalence, it is usually impossible to see any pattern in the values. In general, we seek a partition of the actors into subsets so that actors within each subset are more nearly equivalent, and actors in different subsets are less equivalent.

Hierarchical clustering is one data analysis technique that is ideally suited for the task of partitioning actors. There are many hierarchical clustering algorithms, and computer routines are widely available. The input to a clustering program is (usually) a one-mode symmetric matrix in which the entries measure the similarity (or dissimilarity) of pairs of entities. A positional analysis using hierarchical clustering would use a matrix with measures of structural equivalence (either the correlation matrix or the matrix of Euclidean distances) as input.

A second method, CONCOR (an acronym for CONvergence of iterated CORrelations) was specifically developed for analyzing social network data in order to identify subsets of equivalent actors. This method was first used for analyzing social network data by Harrison White and others (Breiger, Boorman, Arabie, Schwartz etc.) in their research on the application of social networks to the algebraic study of roles (Breiger *et al.* 1975; White *et al.* 1976). Both hierarchical clustering and CONCOR give a partition of the actors in a network.

Consider a hypothetical network represented by a sociomatrix. In this form, it is difficult, if not impossible, to see any regularities or patterns which might exist in the data. However, if we were to rearrange (permute) both the rows and the columns of the sociomatrix (in the same way), then we might be able to see considerable regularity in the relational ties among sets of actors. In such a permutation, one looks for subsets of actors that are similar with respect to sending and receiving ties.

2.3.2. Relational ties between and among positions

Once positions have been identified, the second task of a representation is describing how the positions relate to each other. There are three common ways to represent the relational ties between and among positions: a density table, an image matrix, and a reduced graph. A starting point for representing relational ties between and among positions is to use the positions to permute the rows and columns of the original sociomatrix so that actors who are assigned to the same position are adjacent in the permuted sociomatrix. Rows (and columns) corresponding to actors in the same position are arranged so they are adjacent in the permuted matrix. If all actors within each position are structurally equivalent, then when the rows and columns of the original sociomatrix are permuted so that actors who are assigned to the same equivalence class are adjacent, the submatrices corresponding to the relational ties between and among positions are filled with either all zeros or all ones.

If submatrices of a sociomatrix that contain relational ties between positions are filled completely with ones or completely with zeros the decision concerning whether a relational link exists between positions is straightforward.

Another useful way to summarize the relational ties among positions is in a *density table* (or *density matrix*). A density table is a matrix that has positions rather than individual actors as its rows and columns, and the values in the matrix are the proportion of "choices" that are present from the actors in the row position to the actors in column position.

Often we would like to be able to summarize the relational ties among the positions in a more parsimonious way. An *image matrix* is a summary of the relational ties between and among positions, so that each relational tie is coded as either present or absent between each pair of positions. The image matrix is constructed by allowing rows and columns to refer to positions, rather than individual actors. If we let B be the number of positions in the network, then the image matrix is of size $B \times B$. A "1" in row k, column l of this matrix indicates that position $\mathscr{B}_k$ has a directed relational tie to position $\mathscr{B}_l$. When the model is perfect so that all submatrices are either filled with ones or filled with zeros, then there is no ambiguity about whether a relational tie exists between positions.

When submatrices contain both zeros and ones, then not all actors in the position "choose" all actors in the other positions. In that case, the actors within the positions are not perfectly structurally equivalent, and the description of how positions relate to each other is less clearcut.

It is important to note that image matrices are the starting point for a blockmodel of a social network data set. An image matrix (for a single relation), or a set of image matrices (one for each relation in a multiple relational analysis), along with a description of which actors are assigned to which positions is called a *blockmodel*. Blockmodels are the usual representations for the grouping of actors, based on an equivalence definition.

A final useful way to present the relational ties between and among positions is in a *reduced graph*. A reduced graph has positions as nodes, and uses the relational ties in an image matrix to define the arcs between nodes. It therefore has fewer nodes and fewer lines than the original graph. We use the following rule to construct the reduced graph. It is easy to construct the reduced graph from the image matrix. A "1" in the image matrix indicates that there is an arc from the row position to the column position in the reduced graphs. So, in the reduced graph, there is an arc between the nodes representing positions $\mathscr{B}_k$ and $\mathscr{B}_l$ if there is a relational tie between $\mathscr{B}_k$ and $\mathscr{B}_l$ in the image matrix.

2.4. Assessing the adequacy of the representation

The last task of a positional analysis asks the researcher to determine how well a mathematical representation of the positions and the ties among the positions "fits" a given network data set. Such tasks are usually called *goodness-of-fit* problems in statistics. We will present several goodness-of-fit indices below, all of which are designed to measure the fit of a blockmodel to a given network data set.

There have been two approaches to this goodness-of-fit task in the literature. The first uses a standard data analytic technique of comparing the observed data set (in this case, the R sociomatrices $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_R$) to the predicted data set, which is based on the blockmodel to be evaluated. A number of measures for this comparison have been presented in the literature. We discuss and illustrate several of these measures in this paper. Unfortunately, there is little consensus or agreement on such statistics.

The second approach is more statistical and model-based. One can first assume that a dyadic interaction model, described, for example, in Wasserman and Anderson (1987) and Chapter 15 of Wasserman and Faust (1992), is operating, and then postulate a *stochastic blockmodel*. This strategy then allows one to conduct likelihood-based, statistical tests for goodness-of-fit. Anderson *et al.* (1992) describe stochastic equivalence and stochastic blockmodels in their article in this issue.

3. Blockmodels

Consider now how to model the relationships among the positions found by adopting a specific equivalence definition and measuring how well subsets of actors in a given data set adhere to this definition. The literature contains numerous discussions of blockmodel construction (Arabie and Boorman 1982; Arabie *et al.* 1978, 1990; Breiger *et al.* 1957; Ennis 1982; Light and Mullins 1979; Panning 1982a, 1982b; White *et al.* 1976). We want to consider how to interpret the results of a positional analysis, when the results are presented in a blockmodel. As we will see, the most interesting and useful features of blockmodels are their theoretical interpretations, the potential for validating structural theories, and their use for comparing structural patterns across groups.

3.1. A definition

We begin with a set of R binary relations defined on one set of g actors. A blockmodel consists of two things:

(1) A partition of actors in the network into discrete subgroups – called positions, and
(2) For each pair of positions a statement of the presence or absence of a relational tie within or between the positions on each of the relations.

A blockmodel is thus a *model*, or a *hypothesis* (White *et al.* 1976) about a multirelational network. It presents general features of the network, such as the relations among positions, rather than information pertaining to individual actors.

We can define a blockmodel more precisely in terms of a mapping of the actors in a network onto the positions in the blockmodel. We begin with a collection of R binary relations and their corresponding R binary sociomatrices defined on g actors in $\mathcal{N}$. A blockmodel is a partition of $\mathcal{N}$ into B positions, denoted $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_B$, where $B < g$, and an onto mapping, ϕ, from B onto the collection of positions, ($\phi(i) = \mathcal{B}_k$ if actor i is in position $\mathcal{B}_k$). A blockmodel also specifies the relational ties among the B positions. Relations among positions are specified by a matrix $\boldsymbol{B}$, whose entries (b_{kl_r}) are equal to unity if there is a relational tie from position $\mathcal{B}_k$ to position $\mathcal{B}_l$ on relation r, and equal to zero otherwise.

Whereas the original relational data are presented in the usual $g \times g \times R$ multirelational sociomatrix, the relational ties among positions in the blockmodel can be presented in a smaller array that we denote by $\boldsymbol{B}$. For a partition of the actors in $\mathcal{N}$ into B positions, a blockmodel is a $B \times B \times R$ binary array, with entries b_{kl_r} indicating the presence or absence of a relational tie between positions $\mathcal{B}_k$ and $\mathcal{B}_l$ on relation r.

The matrix $\boldsymbol{B}$ has also been referred to as a blockmodel, since it presents the presence or absence of relational ties among positions.

A blockmodel thus has two components: the mapping, ϕ, that describes the partition of actors in a network into positions, and the matrix, $\boldsymbol{B}$, which specifies the presence or absence of relational ties between and within positions on each relation. Each actor is assigned to one and only one of the positions, and the assignment is the same across relations. The relational ties are presented in a set of R, $B \times B$ binary arrays. There is one array for each of the R relations in the multirelational network, and each array is an image matrix describing the hypothesized relational ties between positions on the specific relation.

Each of the entries in the $B \times B \times R$ matrix $\boldsymbol{B}$ is called a *block*. Each block, b_{klr}, in the blockmodel corresponds to a submatrix of the original sociomatrix that contains the relevant interposition or intraposition relational ties. A block containing a one is called a *oneblock*, and indicates the presence of a relational tie from the row position to the column position. A oneblock may also be referred to as a *bond* (White *et al.* 1976). A block containing a zero is called a *zeroblock*, and indicates the absence of a relational tie from the row position to the column position. More formally, if there is a hypothesized relational tie from position $\mathscr{B}_k$ to position $\mathscr{B}_l$ on relation r then $b_{klr} = 1$ in the blockmodel; b_{klr} is a oneblock. If there is no hypothesized relational tie from position $\mathscr{B}_k$ to position $\mathscr{B}_l$ then $b_{klr} = 0$ in the blockmodel; b_{klr} is a zeroblock.

4. Building blocks

Assigning actors to positions is only one part of constructing a blockmodel. Further, before interpreting a blockmodel, one must also determine whether each block is a oneblock or a zeroblock. Suppose that we have permuted the rows and the columns of the sociomatrix for each relation so that actors who are assigned to the same position occupy adjacent rows, and columns, in the permuted sociomatrix. In the permuted sociomatrix, all entries (x_{ij}) pertaining to relational ties between or within positions, will be contained in submatrices of the sociomatrix. Now, if all actors within each position are *perfectly* structurally equivalent, then all submatrices corresponding to relational ties within and between positions, for all relations, will be filled either all with zeros or all with ones. In such a case, it is easy to determine whether a block should be a oneblock or a zeroblock. However, in real network data, pairs (or collections) of actors are seldom structurally equivalent. In the permuted sociomatrix the submatrices corresponding to inter- and intraposition relational ties will usually contain both ones and zeros. Therefore, determining whether a block in a blockmodel is oneblock or a zeroblock is not straightforward. Constructing a blockmodel requires a rule which governs the assignment of a zero or one to the relational tie between positions in the model.

There are several criteria which have proved useful for deciding whether a block should be coded as a zeroblock or a oneblock. These include:

- perfect fit (fat fit)
- zeroblock (lean fit)
- oneblock
- α density criterion.

In any blockmodel, each of the $B \times B \times R$ elements of $\boldsymbol{B}$ contains the hypothesized value of the relational tie from the row position to the column position on the layer relation. As described above, b_{klr} denotes the value of the hypothesized relational tie from position $\mathscr{B}_k$ to position $\mathscr{B}_l$ on relation r. If the block is a oneblock then $b_{klr} = 1$, and if the block is a zeroblock then $b_{klr} = 0$. The decision about whether a relational tie exists or not in each block of $\boldsymbol{B}$ depends on the observed values of the relational ties between actors in the positions. That is, b_{klr} depends on the values of x_{ijr} for $i \in \mathscr{B}_k$ and $j \in \mathscr{B}_l$. We will let g_k be the number of actors in position $\mathscr{B}_k$ and g_l be the number of actors in position $\mathscr{B}_l$. For distinct $\mathscr{B}_k$ and $\mathscr{B}_l$, there will be $g_k \times g_l$ relational ties from members of position $\mathscr{B}_k$ to members of position $\mathscr{B}_l$. For ties among members of the same position, there will be $g_k \times (g_k - 1)$ relational ties among actors in position $\mathscr{B}_k$. Note that in a blockmodel, relational ties from a position to itself are meaningful, and often quite important theoretically – in contrast to reflexive relational ties for actors and diagonal entries in a sociomatrix, which are often undefined.

The most common criteria for defining oneblocks and zeroblocks are based on the density of ties within a block. The density of ties in block b_{klr} will be denoted as Δ_{klr} and is defined as the proportion of relational ties that are present. For $k \neq l$ this proportion is:

$$\Delta_{klr} = \frac{\sum_{i \in \mathscr{B}_k} \sum_{j \in \mathscr{B}_l} x_{ijr}}{g_k g_l} \tag{1}$$

for $k = 1, 2, \ldots, B$, $l = 1, 2, \ldots, B$, and $r = 1, 2, \ldots, R$. The density of ties within a position, for example, in block b_{kkr}, is equal to:

$$\Delta_{kkr} = \frac{\sum_{i \in \mathscr{B}_k} \sum_{j \in \mathscr{B}_k} x_{ijr}}{g_k(g_k - 1)} \tag{2}$$

for $i \neq j$.

We now specify more formally some useful criteria for defining zeroblocks and oneblocks in a blockmodel. More details on these criteria can be found in Chapter 10 of Wasserman and Faust (1992).

Perfect fit (fat fit). The perfect fit blockmodel occurs if all actors in each position are structurally equivalent. This ideal situation would result in submatrices in the permuted sociomatrix that were filled all with ones or all with zeros. The criterion for a perfect fit blockmodel requires that the relational tie between two positions on a given relation is equal to one only if all actors in the row position have relational ties to all actors in the column position, and a relational tie between positions is equal to zero only if there are no relational ties from actors in the row position to actors in the column position (Breiger *et al.* 1975; Carrington *et al.* 1979/80). Specifically,

$$b_{klr} = \begin{cases} 0 & \text{if } x_{ijr} = 0, \text{ for all } i \in \mathscr{B}_k,\ j \in \mathscr{B}_l \text{ and,} \\ 1 & \text{if } x_{ijr} = 1, \text{ for all } i \in \mathscr{B}_k,\ j \in \mathscr{B}_l. \end{cases}$$

Zeroblock (lean fit) criterion. The zeroblock criterion states that the relational tie between two positions on a given relation is zero only if there are *no* relational ties from actors in the row position to actors in the column position on the specified relation, otherwise the block is a oneblock. This criterion was first proposed by White *et al.* (1976) (see also Arabie *et al.* 1978; Arabie and Boorman 1982). Specifically,

$$b_{klr} = \begin{cases} 0 & \text{if } x_{ijr} = 0, \text{ for all } i \in \mathscr{B}_k,\ j \in \mathscr{B}_l \\ 1 & \text{otherwise.} \end{cases}$$

Oneblock criterion. The oneblock criterion focuses on oneblocks rather than on zeroblocks. It requires that the submatrix of the sociomatrix corresponding to the intra- or interposition relational ties be completely filled with ones. All possible relational ties from actors in row position $\mathscr{B}_k$ to actors in column position $\mathscr{B}_l$ need to be present in order to define a oneblock, otherwise it is a zeroblock. We define

$$b_{klr} = \begin{cases} 1 & \text{if } x_{ijr} = 1, \text{ for all } i \in \mathscr{B}_k,\ j \in \mathscr{B}_l \\ 0 & \text{otherwise.} \end{cases}$$

α criterion. As we have noted, real social network data rarely contain (perfectly) structurally equivalent actors, thus blockmodels which are based on the property of structural equivalence are unlikely to contain blocks all of which are either perfect oneblocks or perfect zeroblocks. For various reasons we expect that oneblocks might contain some zeros and zeroblocks might contain some ones. Therefore it is reasonable to define some threshold density, α, such that if the observed block density, Δ_{klr}, is greater than or equal to α then the block will be coded as oneblock, and if the observed block density is less than α then the block is coded as a zeroblock (Arabie *et al.* 1978). We define the α criterion as:

$$b_{klr} = \begin{cases} 0 & \text{if } \Delta_{klr} < \alpha \\ 1 & \text{if } \Delta_{klr} \geq \alpha . \end{cases}$$

One guideline for choosing a value of α is that it should depend on the density of the relations in the analysis. Two commonly used values are the overall (grand) density computed across all relations, or, since all relations are unlikely to have the same density, there could be R separate α's, one for each relation ($\alpha_r = \Delta_r$).

5. Examples

We will now analyze two examples in detail to illustrate these criteria. The examples are taken from Krackhardt (1987a) and Wasserman and Faust (1992). The two data sets consist of a network of relations linking some nations, and advice and friendship among a group of managers in a high-tech company. We will first describe each data set, and then illustrate blockmodels on each of these data sets.

5.1. World systems data

The actors in this network are nations, selected from a list of 63 countries given in Smith and White (1988). We chose countries representing different categories from across several developmental classifications: Snyder and Kick's (1979) core/periphery status, Nemeth and Smith's (1985) alternative world system classification and level of industrialization, and a historical economic base from Lenski (as reported in Breedlove and Nolan 1988). We also chose countries to both span the globe and to represent politically and economically interesting characteristics. Only countries for which data were reported in 1984 commodity trade statistics were eligible for inclusion. We attempted to reduce the number of shared borders between countries, though some politically interesting countries are included even though they share borders (Israel and Syria, for example). Because of data availability, less-developed nations (African nations in particular) are probably under-represented in this set. The 24 countries represented as actors in this network are a geographically, economically, and politically diverse set, chosen to represent a range of interesting features and to span the categories of existing world system/development typologies. We will refer to these data as the "World system network".

In this paper we analyze three relations, two of them are economic and one is political. The relations are:

- imports of crude materials, excluding fuel
- imports of basic manufactured goods
- diplomatic exchange.

The first two relations are taken from the United Nations Commodity Trade Statistics (United Nations 1984). The third relation comes from The Europa Year Book (Europa Publications 1984), which lists for each country those countries that have embassies or high commissions in the host country.

All three relations are binary and directional. The two economic relations were reported on a continuous US$ scale. The reported values indicate the amount of goods (of the specified type) in 100,000 US$ imported by one country from the other. We note that the UN does not list trade amounts under US$100,000. In order to standardize the imports to control for the vastly different economy sizes across countries, we first standardized each value by dividing by the country's total imports on that commodity. If the realized proportion was less that 0.01 percent, we coded the relational tie as absent. Otherwise, the tie was coded as present. This standardization actually had very little impact. Most of the relational ties that were changed from "trade present" to "trade absent" were large countries (US, Japan, UK) importing small amounts from very small countries (Madagascar, Liberia, Ethiopa). The diplomatic relation records a tie as present if one country has an embassy or a high commission in another country. These data are taken from the 1984 Europa Year Book (Europa Publications 1984).

The data set also includes four attribute variables reflecting the economic and social characteristics of the countries. The first two attribute variables measure annual rates of change between 1970 and 1981. They are: Annual population growth rate between 1970 and 1981, and Annual growth rate in GNP per capita between 1970 and 1981. The second two attribute variables measure rates of education and energy consumption. These variables are: Secondary school enrollment ratio in 1980, and Energy consumption per capita in 1980 (measured in kilo coal equivalent). Measurements on these four variables were taken from The World Bank (1983).

Researchers using network data have constructed blockmodels of positions in the world system using data on trade, diplomatic ties, and military interventions for networks of most countries (Snyder and Kick 1979; Nemeth and Smith 1985; Kick n.d.) or for samples of developed (core) nations (Breiger 1981a). In addition, numerous researchers have attempted to validate these blockmodels using characteristics of countries and to compare the positions of countries resulting from blockmodels of relational data with alternative schemes for classifying countries (Snyder and Kick 1979; Kick n.d.; Nolan 1983, 1987, 1988; Lenski and Nolan 1984). Among the variables that have been used to study positions in the world system are four variables that we will use:

- population, annual growth rate from 1979 to 1981;
- GNP per capita, annual growth rate from 1970 to 1981;
- secondary school enrollment ratio in 1980; and
- energy consumption per capita (in kilo coal equivalents) in 1980.

The blockmodel of the world system example is based on three relations: Manufactured goods, Crude materials excluding fuel, and Diplomatic ties. For this analysis we measured structural equivalence using the Pearson product-moment correlation coefficient, calculated on the rows and columns of the three sociomatrices. We used UCINET (MacEvoy and Freeman n.d.) to calculate the correlations. Positions were identified using complete link hierarchical clustering, in the program SYSTAT (Wilkinson 1987). In order to study the positional system in detail we will use a six position model. These six positions and their members are:

- $\mathscr{B}_1$: Japan, United Kingdom, United States
- $\mathscr{B}_2$: China, Czechoslovakia, Indonesia, Spain, Yugoslavia
- $\mathscr{B}_3$: Argentina, Brazil Finland, New Zealand, Pakistan, Switzerland, Thailand
- $\mathscr{B}_4$: Algeria, Egypt, Syria
- $\mathscr{B}_5$: Ecuador, Honduras, Israel
- $\mathscr{B}_6$: Ethiopia, Liberia, Madagascar.

The density tables for this partition are presented in Table 1. These tables show the proportion of relational ties that are present from countries in the row position to countries in the column position. Density tables, and image matrices were constructed using UCINET (MacEvoy and Freeman n.d.).

Notice that these density tables have some values that are equal to 1.00 and 0.00, indicating that some submatrices corresponding to intraposition or interposition relational ties are either completely filled with ones, or completely filled with zeros. Therefore, it is possible to consider using either the zeroblock or the oneblock criteria to construct a blockmodel for these data. However, using the oneblock criterion gives a very sparse blockmodel, since only 21 of the $36 \times 3 = 108$ submatrices have densities equal to one, and would thus be coded as oneblocks in the blockmodel. The zeroblock criterion gives a very dense blockmodel, since only 14 of the submatrices have densities equal to zero, and would be coded as zeroblocks. Therefore, it seems reasonable to use the α density criterion to construct the blockmodel image matrices. The densities of the three relations are:

- imports raw materials from, density = 0.556
- imports manufactured goods from, density = 0.562
- diplomat resides in, density = 0.668.

Since there is some variation in the densities of these relations, it is reasonable to choose α density cutoff values that are specific to the relations. Using the α density rule with relation specific α's gives the set of five image matrices, in Figure 2.

The collection of three image matrices, along with the assignment of countries to positions constitutes the blockmodel for these data. For the moment, notice that no two relations have the identical image matrices, though there are features common to all three image matri-

Table 1
Density tables for manufactured goods, crude materials and diplomatic ties

	Manufactured goods					
	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{B}_4$	$\mathscr{B}_5$	$\mathscr{B}_6$
$\mathscr{B}_1$	1.000	1.000	0.952	1.000	1.000	1.000
$\mathscr{B}_2$	1.000	1.000	0.914	0.933	0.467	0.533
$\mathscr{B}_3$	0.952	0.857	0.810	0.667	0.571	0.286
$\mathscr{B}_4$	0.444	0.400	0.095	0.000	0.000	0.111
$\mathscr{B}_5$	0.556	0.133	0.286	0.000	0.000	0.111
$\mathscr{B}_6$	0.222	0.067	0.000	0.000	0.000	0.000
	Crude materials					
	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{B}_4$	$\mathscr{B}_5$	$\mathscr{B}_6$
$\mathscr{B}_1$	1.000	1.000	0.952	1.000	0.778	0.667
$\mathscr{B}_2$	0.867	0.800	0.657	0.600	0.267	0.133
$\mathscr{B}_3$	0.952	0.917	0.762	0.571	0.476	0.048
$\mathscr{B}_4$	0.556	0.867	0.238	0.333	0.111	0.000
$\mathscr{B}_5$	0.778	0.333	0.238	0.000	0.167	0.000
$\mathscr{B}_6$	1.000	0.333	0.143	0.556	0.222	0.000
	Diplomatic ties					
	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{B}_4$	$\mathscr{B}_5$	$\mathscr{B}_6$
$\mathscr{B}_1$	1.000	1.000	0.952	1.000	1.000	1.000
$\mathscr{B}_2$	1.000	0.900	0.943	1.000	0.400	0.600
$\mathscr{B}_3$	0.952	0.857	0.714	0.714	0.429	0.238
$\mathscr{B}_4$	1.000	1.000	0.667	0.333	0.111	0.667
$\mathscr{B}_5$	1.000	0.333	0.476	0.222	0.833	0.111
$\mathscr{B}_6$	0.889	0.267	0.000	0.333	0.000	0.333

ces. In all image matrices all positions relate *to* block $\mathscr{B}_l$, and overall, positions $\mathscr{B}_l$ and $\mathscr{B}_2$ are involved in more relational ties than are the remaining positions.

5.2. *Krackhardt's high-tech managers*

These data were gathered by David Krackhardt (1987a) in a small manufacturing organization on the west coast of the US. This organization had been in existence for ten years and produced high-tech machinery for other companies. The firm employed approximately 100 people, and had 21 managers. These 21 individuals are the set of actors for this data set. We will refer to this example as "Krackhardt's high-tech managers". Krackhardt's interest in these data focused on

Manufactured goods

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Crude Materials

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Diplomatic Ties

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Fig. 2. Blockmodel image matrices for three relations in the world system example.

the perceptions of all 21 managers of the entire network of informal advice and friendship relations. Specifically, he was interested in the perceptions held by the actors of the structure of the entire group. Here, we are interested only in the reports made by each manager of their own advice seeking and friendships.

Each manager was given a questionnaire and asked two questions: "Who would [you] go to for advice at work?" and "Who are your friends?" Each actor was given a roster of the names of the other managers, and asked (in a free-choice setting) to check the other managers with respect to advice seeking, and friendship. We will also study two actor attributes: Age, and Length of time employed by the organization. Both variables are measured in years.

The blockmodel of Krackhardt's high-tech managers is based on the two relations, advice and friendship. We split the actors into a four position blockmodel, using CONCOR in the program UCINET (MacEvoy and Freeman n.d.). These four positions are:

- $\mathscr{B}_1$: 1, 8, 11, 12, 16, 17
- $\mathscr{B}_2$: 3, 4, 5, 9, 15, 20
- $\mathscr{B}_3$: 2, 6, 7, 14, 21
- $\mathscr{B}_4$: 10, 13, 18, 19.

Table 2
Density tables for advice and friendship

	Advice			
	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{B}_4$
$\mathscr{B}_1$	0.200	0.083	0.467	0.208
$\mathscr{B}_2$	0.944	0.367	0.900	0.625
$\mathscr{B}_3$	0.200	0.100	0.750	0.150
$\mathscr{B}_4$	0.500	0.708	0.500	0.750
	Friendship			
	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{B}_4$
$\mathscr{B}_1$	0.500	0.389	0.300	0.208
$\mathscr{B}_2$	0.278	0.133	0.233	0.167
$\mathscr{B}_3$	0.167	0.067	0.300	0.100
$\mathscr{B}_4$	0.292	0.375	0.150	0.000

These subsets show the mapping, $\phi(i) = \mathscr{B}_k$, for each of the 21 managers.

The second step in the blockmodel analysis is to describe the relational ties between and within positions. The density tables for the Advice and Friendship relations are presented in Table 2. Since the four criteria for assigning oneblocks and zeroblocks for the blockmodel all depend on the density of relational ties within and between positions, the density tables contain all of the information that is necessary for constructing the blockmodel.

Let us consider the four criteria in turn. Notice that since the densities in the submatrices are not *all* equal to either zero or one, the perfect fit criterion will not yield a blockmodel for this partition of actors. Similarly, since there are no submatrices with density equal to one, the oneblock criterion would give an uninteresting blockmodel, one all filled with zeros. The zeroblock criterion also gives an uninteresting blockmodel. Only the single block containing the relational ties within position $\mathscr{B}_4$ on the advice relation has a density of zero. Thus, we will use the α criterion, with α_r for each relation equal to the density of the relation, Δ_r. The density of the Advice relation is equal to 0.452, so any submatrix with a density greater than or equal to $\alpha_1 = \Delta_1 = 0.452$ will be coded as a oneblock in the Advice image matrix. The density of the Friendship relation is equal to 0.243, so any submatrix with a density greater than or equal to 0.243 will be coded as a oneblock in the Friendship image matrix.

Advice

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Friendship

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

Fig. 3. Blockmodel image matrices of the advice and friendship relations for Krackhardt's high-tech managers.

The image matrices for this blockmodel are presented in Figure 3. Each of these image matrices may also be presented in the form of a reduced graph, in which nodes represent positions, and the arcs are the relational ties between positions. Figure 4 gives these graphs. For the moment, simply notice that the image matrices and graphs for the two relations are quite different. We will examine these differences in more detail in the remainder of this paper.

6. Blockmodel interpretation

Blockmodels are hypotheses about the structure of relations in a social network. Although blockmodels may appear deceptively simple, in that they usually consist of rather small arrays of zeros and ones, the patterns of relations among positions can present important structural properties. Rules for interpreting blockmodels are quite

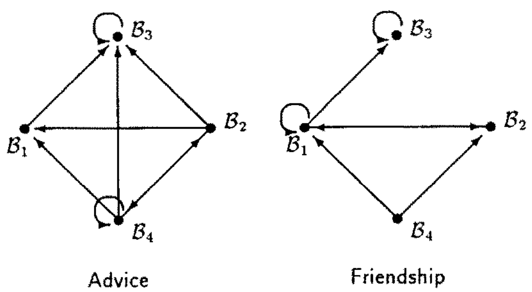

Fig. 4. Reduced graphs of advice and friendship relations for Krackhardt's high-tech managers.

important. In the following sections we discuss three different ways to interpret a blockmodel:

(1) validation of a blockmodel using actor attributes;
(2) descriptions of individual positions; and
(3) descriptions of the overall blockmodel.

The first way to interpret a blockmodel uses exogenous actor attribute variables to describe the positions in the blockmodel. The later two ways provide statements about the form of the blockmodel, $\boldsymbol{B}$, without reference to the attributes of the actors.

6.1. Actor attributes

One of the most straightforward ways to interpret a blockmodel is to use attributes of the actors to describe the positions. If there are systematic differences between positions in the characteristics of their members, then we have some external validation for the blockmodel.

There are many examples of network analyses that have used actor attributes to help interpret blockmodels. Investigations of positions in the world economic and political system (via network models) have used growth in GNP per capita measured on countries to help understand the positional structure (Snyder and Kick 1979). Researchers studying scientific communities have used the date of a scientist's professional degree, the number of articles each has published, the number of citations made to their published work, and the dollar amount of grant money they have received, to help understand the structure of scientific networks (Mullins *et al.* 1977; Breiger 1976). In his investigation of the social structure among prison inmates, Arabie (1984) used ethnicity, level of education, and drinking habits to validate a blockmodel, employing a discriminant analysis to study whether positional assignments could be predicted from the actor attributes.

Depending on one's theoretical orientation, one might argue that the characteristics of the actors are an important determinant of their network relations which then led to the observed positional structure. Or (on the other hand) that the structural positions of the actors (and network processes) were influential in determining the characteristics of the actors in the model. For example, world system theory argues that the position of a country in the world system influences the rate of development of the country. On the other hand, social psychologists might argue that similarity between actors in their characteristics leads to mutual attraction and the formation of relational ties among actors, and thus the structure of the group. In either case, the actor attributed are related to the network structure.

6.1.1. Examples

We will first examine the four position model of the corporation studied by Krackhardt, and then look at the six position model of the world system data. In each case we will present the average value of each attribute variable, calculated within each of the positions in the blockmodel.

Krackhardt's high-tech managers. In addition to relational data, Krackhardt recorded information about the characteristics of the managers in the corporation. This includes information about the age of each manager, and the number of years each manager had been with the company (tenure). One might reasonably expect that patterns of advice seeking would be related to the experience of the managers, and that experience would be reflected in the age of the managers and/or in their length of service (tenure) with the company. For the company as a whole the mean age is 39.71 years, and the mean length of service is 11.75 years. To examine whether age and tenure vary across positions we computed the mean and standard deviation of age and tenure for the managers within each of the positions. These statistics are reported in Table 3.

Notice that members of position $\mathscr{B}_3$ are oldest on average (47.00 years) and have the longest tenure in the company (18.236 years).

Table 3
Mean age and tenure of actors in positions from CONCOR analysis of advice and friendship relations, standard deviations in parentheses

	Position			
	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{B}_4$
Age	34.00	40.83	47.00	37.50
	(6.48)	(10.93)	(9.62)	(7.33)
Tenure	12.28	8.18	20.10	5.86
	(7.69)	(3.60)	(8.84)	(4.26)

Position $\mathscr{B}_1$ has, on average, the youngest members (34 years), but, they have the second longest tenure on average (12.28 years). Members of positions $\mathscr{B}_2$, and $\mathscr{B}_4$ are intermediate in age, and have been with the company for the shortest time, on average.

The world system. Let us turn now to the world system example, and examine the characteristics of the countries in each of the positions. Considerable research has focused on whether, and how, the position of a country in the world system affects its social and economic development. One prediction is that dependency status within the world political and economic system affects the rate of economic development of countries.

At the outset, we expect that that GNP per capita growth rate, secondary school enrollment ratio, and energy consumption per capita will be higher in core (industrialized) nations than in peripheral (developing) nations. In contrast, population growth rate is expected to be higher in peripheral developing nations than in core industrialized nations. Table 4 shows the means and standard deviations of these variables within each of the six positions. Notice that there is a tendency for the means to be ordered across positions. Positions $\mathscr{B}_1$, $\mathscr{B}_2$, and $\mathscr{B}_3$ have the lowest annual growth rate in population, the highest secondary school enrollment ratio, and the highest energy consumption. Positions $\mathscr{B}_4$, $\mathscr{B}_5$, and $\mathscr{B}_6$ have the highest annual growth rate in population, the lowest secondary school enrollment ratio, and the lowest energy consumption. Annual growth rate in GNP

Table 4
Means of variables within positions for world system example

	Position					
	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{B}_4$	$\mathscr{B}_5$	$\mathscr{B}_6$
Population	0.73	1.30	1.60	3.17	3.13	2.70
Annual growth rate	(0.55)	(0.63)	(1.06)	(0.61)	(0.46)	(0.76)
GNP per capita	2.33	4.30	2.21	4.70	2.20	−0.57
Annual growth rate	(0.94)	(1.83)	(1.97)	(1.47)	(2.18)	(1.26)
Secondary school	90.00	57.00	51.14	43.67	47.33	14.33
Enrollment ratio	(7.55)	(26.37)	(27.70)	(9.71)	(21.94)	(4.93)
Energy consumption	7217.67	2615.40	2892.29	791.00	1265.67	200.00
Per capita	(3838.70)	(2625.32)	(2527.21)	(185.57)	(1354.87)	(262.73)

per capita varies, but less systematically, across the positions, from a high of 4.3 for position $\mathscr{B}_2$ to a low of -0.57 for position $\mathscr{B}_6$.

These examples have demonstrated that the positions of actors, identified on the basis of their relational ties, differ in the attributes of their actors. Exogenous attribute variables vary across the positions. However, a more complete interpretation of the blockmodel requires examining how the positions are related to each other.

6.2. Describing individual positions

A second way to interpret the result of a blockmodel analysis is to describe how the individual positions relate to each other. This requires examining how the positions are involved in the relations, and how each position sends and receives relational ties in the blockmodel. One useful and informative strategy relies simply on the relational ties to and from the positions in the model (see Marsden 1989; Burt 1976). Descriptive typologies of positions are useful for summarizing tendencies for positions to receive choices or to make choices within or outside the position. However, this approach does not allow the researcher to test whether these tendencies are statistically large.

Graph theorists have used the indegrees and outdegrees of nodes to describe types of nodes in a directed graph. We can use nodal indegrees and outdegrees to distinguish four different types of nodes (see Harary *et al.* 1965; and Marsden 1989):

- isolates: nodes with neither indegree nor outdegree;
- transmitters: nodes with only outdegree;
- receivers: nodes with only indegree; and
- carriers or ordinary points: nodes with both indegree and outdegree.

As Marsden (1989) has noted, this classification is also useful for describing positions in networks. Thus the same labels may be used to describe how positions relate to each other.

These rather neutral labels refer simply to the presence or absence of relational ties to or from positions. If we also take into account the prevalence of "choices" that are made within a position, rather than to or from other positions, then the description can be more informative. Burt (1976) provides a typology of positions that is useful for positively valued, affective, interpersonal ties, such as respect, liking, or esteem. His typology takes into account both whether the relational ties occur primarily within a position, and whether the ties are directed to the position from others.

First, Burt distinguishes between positions that receive "choices" and positions that do not receive "choices". Second, he distinguishes between positions that make less than half of their total "choices" to their own members, and positions that make half or more of their "choices" to their own members. By making these two distinctions one can determine whether each position receives "choices" or not, and whether each position makes more "choices" within the position rather than outside the position. These two distinctions result in a classification into four types of positions. Isolate positions neither give many "choices" nor do they direct many choices to other positions. Sycophants give more "choices" to other positions than to themselves, and do not receive many "choices". Brokers both receive "choices" and send "choices" to other positions. The Primary position receives "choices" both from other positions, and from its own members.

The use of an arbitrary cutoff value of 0.5 for the proportion of a position's "choices" that are within the position is problematic, since the size of the position, g_k, relative to the size of the group, g, affects the our expectation for the proportion of ties that are within the position. If a position is large relative to the size of the entire group, then one would expect many of the "choices" made by position members to be to other members of the position, simply because of their prevalence in the group, even if there were no ingroup bias in making "choices". Similarly, a small position would be expected to have a low proportion of "choices" within the position, simply because there are relatively fewer actors in the position. It is therefore useful to consider the relative size of the position when examining the tendency for position members to make choices within the position. A reasonable value is the proportion of the total "choices" that are made within the position, compared to the proportion that would occur if there were no within or outside position bias in choices.

Consider the "choices" made by members of position $\mathscr{B}_k$. If there are g_k actors in position $\mathscr{B}_k$ then there are $g_k \times (g_k - 1)$ possible "choices" to be made within the position. In the whole group, there g actors, so there are $g_k \times (g - 1)$ possible choices to be made in total by actors in position $\mathscr{B}_k$ (recall that self choices are undefined). So, if

Table 5
Typology of positions, adapted from Burt (1976)

	$\frac{\sum_{j \notin \mathscr{B}_k} \sum_{i \in \mathscr{B}_k} x_{jir}}{\sum_{j=1}^{g} \sum_{i=1}^{g} x_{jir}} \sim 0$	$\frac{\sum_{j \notin \mathscr{B}_k} \sum_{i \in \mathscr{B}_k} x_{jir}}{\sum_{j=1}^{g} \sum_{i=1}^{g} x_{jir}} > 0$
$\frac{\sum_{i \in \mathscr{B}_k} \sum_{j \in \mathscr{B}_k} x_{ijr}}{\sum_{i \in \mathscr{B}_k} \sum_{j=1}^{g} x_{ijr}} \geq \frac{g_k - 1}{g - 1}$	isolate	primary
$\frac{\sum_{i \in \mathscr{B}_k} \sum_{j \in \mathscr{B}_k} x_{ijr}}{\sum_{i \in \mathscr{B}_k} \sum_{j=1}^{g} x_{ijr}} \leq \frac{g_k - 1}{g - 1}$	sycophant	broker

there were no bias toward (or away) from making choices within the position, then we would expect that the proportion of a position's total choices that are made within the position would be:

$$\frac{g_k \times (g_k - 1)}{g_k \times (g - 1)} = \frac{g_k - 1}{g - 1}. \tag{3}$$

One can use this proportion (rather than 0.5) as a baseline for evaluating the tendency for within position choices. Since this proportion depends on the number of actors in the position it will probably differ across positions.

Table 5 summarizes the typology. The columns refer to the first distinction (receiving "choices" or not) and the rows refer to the second distinction (proportion of "choices" within the position).

The labels, isolate, sycophant, broker and primary position, depend on the content of the relation. If the relation is negatively valued (blame, dislike and so on) then the primary position would be more appropriately interpreted as a scapegoat, or pariah. If the relation involved the flow of material goods (such as trade among nations, or buying and selling among corporations) then a position with a high ratio of "choices" made to "choices" received (in other words, relatively a high ratio goods sent) would be interpreted as a supplier or source, and a position with a relatively high ratio of "choices" received (in other words, relatively high ratio of goods bought) would be a consumer (Galaskiewicz and Krohn 1984), or enduser, and the brokers would be middlemen in the transaction.

For communication networks, where the relation is usually the transmission of a message or information, Richards (1989) has presented a typology that is also based on indegree and outdegree conditions, and makes a similar distinction among positions. Since communication is often symmetric, it is likely to be represented as a nondirected graph. Richards distinguishes first between participant and nonparticipant positions. Nonparticipants are either isolates (neither indegree nor outdegree) or "tree nodes" (relate to only one other node); participants are liaisons (both indegree and outdegree, that is, they link two or more others) or group members. This typology is a fundamental part of the network analysis program NEGOPY (Richards 1989).

Marsden (1989) has extended Burt's (1976) typology by distinguishing between the level of "choices" *made* by a position, the level of "choices" *received* by a position, and the position's *ingroup preference*. His typology combines features of the graph theoretic classification with the distinctions made in Burt's typology. Making a binary (high versus low) distinction on each of these three dimensions, gives a typology with eight different kinds of positions. Marsden proposes log linear models for examining these three properties.

Focusing on the level of indegree, outdegree, and within position relational ties can give a fairly interesting, and useful description of the positions that relies simply on the relational ties that each position is involved in. Since a blockmodel is likely to contain several relations, with quite different substantive meanings or contents arriving at a consistent description of a given position might be difficult. The "label" for a position (in one of the above schemes) might not be the same across the different relations.

6.2.1. An example

Now, let us look at an example to illustrate the typology of positions in Table 5. We will use the four position blockmodel of Krackhardt's high-tech managers. In order to classify the positions, it is necessary to count the number of "choices" made by members of each position to other actors, both within and outside the position. These counts can be made by examining the sociomatrix with rows and columns permuted so that actors in the same position are adjacent in the permuted sociomatrix. Table 6 gives the frequency of "choices" within each block.

First, notice that all positions receive at least some "choices" on both relations. Therefore on neither relation is there an isolate or a

Table 6
Frequency of "choices" within and between positions for advice and friendship

	Advice				total
	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{B}_4$	
$\mathscr{B}_1$	6	3	14	5	28
$\mathscr{B}_2$	34	11	27	15	87
$\mathscr{B}_3$	6	3	15	3	27
$\mathscr{B}_4$	12	17	10	9	48
total	58	34	66	32	190
	Friendship				total
	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{B}_4$	
$\mathscr{B}_1$	15	14	9	5	43
$\mathscr{B}_2$	10	4	7	4	25
$\mathscr{B}_3$	5	2	6	2	15
$\mathscr{B}_4$	7	9	3	0	19
total	37	29	25	11	102

sycophant position (though position $\mathscr{B}_4$ does receive relatively few friendship choices). Now, consider position $\mathscr{B}_3$, and its "choices" on the Advice relation. There are $g_3 = 5$ actors in this position, so we would expect that the proportion of their "choices" that would be within the position would be equal to $(5-1)/(21-1) = 0.20$. In fact, members of position $\mathscr{B}_3$ make 15 out of their total 27 Advice "choices" to their own members, for a proportion of $15/27 = 0.56$. Since this proportion is higher than we would expect, this position is a Primary position on the Advice relation. In contrast, consider the Friendship "choices" made by members of position $\mathscr{B}_2$. We would expect that since there are $g_2 = 6$ actors in this position, that they would make $5/20 = 0.25$ of their "choices" to their own members. However, only 4 of their 25 Friendship "choices" (a proportion of $4/25 = 0.16$) are to their own members. Thus, position $\mathscr{B}_2$ is a Broker on the Friendship relation. Table 7 gives the classification of the four positions on each of the two relations, using the typology in Table 5.

Notice that position $\mathscr{B}_3$ is a Primary position on both relations, and position $\mathscr{B}_2$ is a Broker on both relations. Referring to Table 3, we see that position $\mathscr{B}_3$ has the oldest managers, and managers with the longest tenure, on average, whereas managers in position $\mathscr{B}_2$ are intermediate on both attributes.

Table 7
Typology of positions for Krackhardt's high-tech managers

Position	Advice	Friendship
$\mathscr{B}_1$	Broker	Primary
$\mathscr{B}_2$	Broker	Broker
$\mathscr{B}_3$	Primary	Primary
$\mathscr{B}_4$	Primary	Broker

For the most part descriptions of single positions do not take into account the properties of the positions to which a given position is related. A Broker position gives and receives "choices" rather than making "choices" within the position, but the kinds of positions to which it is tied are unimportant. Similarly, a "transmitter" (a position with both in and outdegree) could be either at the bottom of a very long chain of command or pecking order, or toward the top. Thus, intermediate levels in a hierarchy would be indistinguishable (since they would have both indegree and outdegree). Although labels for kinds of positions, that we discussed in this section, are quite useful as a starting point for interpreting the results of a blockmodel analysis, they capture only a limited amount of information about the structural position of the given position. They do not describe the network as a whole.

6.3. *Image matrices*

The third way to study a blockmodel is in terms of the entire configuration of relational ties among positions that is expressed in the image matrix. Many structural theories posit patterns of relational ties among aggregates of actors. For example, the properties of balance and transitivity can be expressed in blockmodels. A network system with a center and a periphery, such as has been proposed for the world economic and political system (Snyder and Kick 1979), can be expressed in a blockmodel. Similarly, systems characterized by a hierarchy, the domination of one or more positions over others, or cohesive subgroups can be represented by blockmodels. We will describe and illustrate these patterns in this section. Theories that are expressed in terms of these patterns may be evaluated by examining the results of a blockmodel analysis to see whether the observed blockmodel is consistent or inconsistent with the predicted pattern. This is approach is quite useful for describing overall patterns.

6.3.1. Image matrices for two position blockmodels

Some of the simplest possible blockmodels can give quite powerful representations of theoretical statements. For example, even a simple two position model, presented in a 2×2 image matrix, can represent quite interesting theoretical properties. In their introduction of blockmodels, White *et al.* (1976) present the 16 possible arrangements that could arise in a two position blockmodel. Since there are two positions, the image matrix for this blockmodel has $2 \times 2 = 4$ cells, each of which may be either a zero (zeroblock) or a one (a oneblock), so there are 2^4 possible arrangements of zeros and ones. Since the order of the positions is arbitrary, there are in fact only 10 distinct images (the others are isomorphic to one of these images).

We begin by describing the images for 2×2 blockmodels, and then discuss some more complicated theoretical patterns that can occur in blockmodels with more than two positions. Figure 5 shows the 16 possible image matrices for a two position model.

Some of these patterns have clear interpretations in terms of structural theories. When there is a theoretical prediction about the arrangement of relational ties between positions, this gives rise to a posited image matrix. White *et al.* (1976) provide useful descriptions for many of these images. Image B, in Figure 5, is a single cohesive subgroup and an isolate position (assuming positive affective relation). Image C could indicate deference directed from members of one position to members of the other. In terms of individual position labels described in the previous section, these would be a sycophant and a primary position. Image D is "pure" reflexivity, and for a positive relation would indicate two cohesive subgroups. Image D could also represent an endogamous system in which all relational ties exist within subgroups, or homophily where all friendship choices are between actors with similar characteristics. In the context of world trade systems, Breiger (1981a) described this pattern as representing separate trading areas. Image E is "pure" symmetry. For a negative relation it would indicate opposition or hostility. Image E could also represent an exogamous system in which all relational ties are directed to members of another group (for example, "seeks a spouse from" in an exogamous system where marriages are between rather

A. Null

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

B. One reflexive arc

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

C. One arc between positions

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

D. Two arcs, reflexive

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

E. Two arcs, symmetric

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

F. Two arcs, reflexive and "out"

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

G. Two arcs, reflexive and "in"

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

H. Three arcs, 2 between positions

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

I. Three arcs, 2 reflexive

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

J. Complete

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Fig. 5. Ten possible image matrices for a two position blockmodel.

than within clans or villages). The combination of image D (for a positive relation) and E (for a negative relation) would be consistent with balance theory which predicts that a actors in a balanced system can be clustered so that all positive "choices" are within subsets and negative "choices" are between subsets. Image F distinguishes between an "active" position and a "passive" position, in terms of "choices" made. Image G combines aspects of a cohesive subgroup (image B) and a deference structure (C), and resembles a core-periphery system (again with one primary position, and a sycophant position). This pattern can also be interpreted as a hierarchy (Breiger 1981a). Image H is complete except for one reflexive tie. White *et al.* also interpret this as a center-periphery or hanger-on pattern. Image I is complete except for one directed relational tie from one position to the other. White *et al.* describe this form as a hierarchy, with deferential ties within each of the two levels of the hierarchy in addition to deferential ties from one position to the other. Image H is somewhat similar to image C, which has ties only between positions. Finally, image J is complete, and therefore shows no differentiation among positions.

6.3.2. Image matrices with more than two positions

Certainly not all blockmodels have only two positions. More interesting, but also more complex, systems arise when there are more positions. However, for a 3-position model there would be $2^9 = 512$ possible image matrices for a single relation, and 104 distinct image matrices (isomorphism classes). As the number of positions increases, the number of distinct image matrices increases rapidly. Instead of enumerating all of the possible images for larger blockmodels, let us examine a few theoretically important, ideal images that display interesting structural properties. In particular we will illustrate images that display the properties of cohesive subgroups, a center-periphery structure, a hierarchy, a transitive system, and a centralized system. Figure 6 shows some of these ideal patterns.

One of the most straightforward patterns is a system that is composed of *cohesive subgroups*. Such a system will have an image matrix, for a single positively valued relation, that consists primarily of intra-position relational ties. Thus, the image matrix for a system of cohesive subgroups will have oneblocks on the main diagonal. The image matrix for a cohesive subgroup system is reflexive at the position level (even though at the level of individual-actor relational ties, self-self "choices" may be undefined). However, the positions in the blockmodel may not be graph theoretic cliques. Oneblocks may contain some zeros (they may not be complete subgraphs), and actors from one position may be connected to all of the actors in another position (the positions may not be "maximal").

A. Cohesive subgroups

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

B. Center-periphery

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

C. Centralized

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

D. Hierarchy

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

E. Transitivity

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

Fig. 6. Some ideal images for blockmodels with more than two positions.

Another common structural pattern to emerge is a *center-periphery* structure. This consists of a core position which is internally cohesive, and one or more other positions connected to the core position, but not to each other (Mullins *et al.* 1977). The peripheral positions may or may not be internally cohesive. Examples of core-periphery systems include an elite position and hangers-on in a social group, or the proposed three "levels" in the world system consisting of the core, periphery and semiperiphery. In general, a center-periphery pattern is apparent in a blockmodel if the positions in the image matrix can be permuted so that the oneblocks are primarily in the upper left triangle of the image matrix, and the zeroblocks are primarily in the lower right triangle. The center-periphery pattern has been found, for example, in the trade relations in the world economic and political system (Snyder and Kick 1979; Breiger, 1981a).

A related pattern is a *centralized* system. A centralized system has all relational ties going toward (or away) from a single position. In an image matrix, all oneblocks would be in the same column (if all relational ties are to the same position), or all oneblocks would be in the same row (if all relational ties are from the same position). Reflexive ties may also be present. A centralized pattern was found by Doreian and Fararo (1985) in their study of citations among major journals in sociology. The most prestigious journals were in the central position and were cited by journals in all other positions. This pattern has also been found by Knoke and Rogers (1979) in their study of an interorganizational network.

Another common pattern is a *hierarchy*. A hierarchy would appear as asymmetric, positive, relational ties directed from each position to one position immediately "above" it. A hierarchy could represent a chain of command in an organization.

A system that is *transitive* at the level of the positions is similar to a hierarchy, but all interposition ties that are implied by the property of transitivity are also present. If there is a directed relational tie from position $\mathscr{B}_k$ to position $\mathscr{B}_l$ and there is a directed relational tie from $\mathscr{B}_l$ to $\mathscr{B}_m$, then there is a directed relational tie from $\mathscr{B}_k$ to $\mathscr{B}_m$. In a fully transitive model, the rows and columns of the image matrix can be permuted so that all oneblocks are in the lower left triangle (or in the upper right triangle) of the matrix. Depending on the substance of the relation, a transitive image could indicate dominance or deference between positions.

These patterns display aspects of theoretically "pure" or ideal structures. It is likely that any actual social network data blockmodels will show some variation around these patterns, or might combine features of two or more patterns.

Examples. Let us first examine the image matrices for the three relations in the world system example. These images are displayed in Figure 2, and appear to show two different patterns. The image matrix for the Manufactured goods relation shows that positions $\mathscr{B}_1$, $\mathscr{B}_2$, and $\mathscr{B}_3$ are the source of manufactured goods imported by all positions, whereas positions $\mathscr{B}_4$, $\mathscr{B}_5$, and $\mathscr{B}_6$ only import, but do not export manufactured goods. This is similar to a centralized system, with three positions in the center. The image matrices for Crude materials, and Diplomatic ties look similar to each other, and different from the image matrix for Manufactured goods. Although neither of these two images perfectly matches one of the ideal types, both matrices are arranged so that the oneblocks are concentrated primarily in the upper left triangle of the matrix, and the zeroblocks are primarily in the lower right triangle. This pattern indicates a center-periphery system. In the world system example, position $\mathscr{B}_1$ is in the center, positions $\mathscr{B}_5$ and $\mathscr{B}_6$ are on the periphery, and the other positions are intermediate.

In the blockmodel for Krackhardt's high-tech managers, the two relations, Advice and Friendship have different patterns. The advice relation is nearly transitive, at the level of the positions (the single "violation" is the reflexive tie at position $\mathscr{B}_4$). If we think of positions seeking advice from other positions that are more prominent in the organization, then position $\mathscr{B}_3$ is at the top, followed by positions $\mathscr{B}_1$, $\mathscr{B}_2$ and $\mathscr{B}_4$. The pattern for Friendship is not as clear. Although positions $\mathscr{B}_3$ and $\mathscr{B}_1$ have intraposition Friendship ties, the system as a whole does not appear to be characterized by cohesive subgroups (at least for this blockmodel).

6.3.3. Image matrices for multiple relations

Interpreting blockmodels with multiple relations can be tedious. The researcher could propose separate interpretations for each image, but in the absence of a theoretical foundation, this seems *ad hoc*. One possible way to interpret multirelational blockmodels is to study pairs of image matrices to see whether they exhibit common kinds of multirelational patterns, such as multiplexity or exchange. Multiplexity of relations is the tendency for two or more relations to occur together. For example, "is a friend of" and "spends time with" are two relations that tend to occur together. Multiplexity in a blockmodel would be apparent if two or more image matrices were identical (or nearly identical). Exchange occurs when one relation "flows" one way, and the second relation "flows" back. For example, "pays money to" and "delivers goods to" are two relations that form an exchange in an economic transaction. The property of exchange would be apparent in a blockmodel if one image matrix were the transpose of the other. This would indicate that whenever one kind of relational tie is present from the row position to the column position, the second kind of relational tie is present from the column position to the row position.

7. Blockmodel evaluation

We now turn to the last question raised at the beginning of the paper. We want to measure the adequacy of a representation of a positional analysis. A blockmodel, one such representation, consists of a partition of the actors in $\mathscr{N}$ into positions and a statement of how the positions relate to each other. The adequacy of this construct can be studied with the methods presented here.

7.1. Background

There are two approaches for assessing how well a blockmodel, or another mathematical representation of the positions among a set of actors, fits. As mentioned earlier, one of these approaches is model-based, and the other is more descriptive. The model-based approach is more statistical, and certainly, more parametric.

The primary difference between these two approaches lies at the center of a positional analyses. The definition of structural equivalence and the algorithms (such as CONCOR) commonly used to find a blockmodel, make no use of statistical ideas. Positional analysis is not a statistical method. This limitation prevents standard, parametric (that is, based on a specific parameterized family of probability distributions) statistical tests and measures from being used to determine directly how well a blockmodel fits a data set. Some researchers use blockmodel representations of a network data to summarize other aspects of a network data set; however, unless the blockmodel representation is independent of these other aspects, any statistical tests will not have accurate error rates. If statistical tests are desired in a network analysis, we recommend the use of statistical methods from the beginning of the analysis. Such methods can be used to find partitions of actors, and lead to proper statistical tests and measures of goodness-of-fit.

There is a compromise between parametric statistical models and the positional analyses described earlier. Nonparametric tests can be used to test specific hypotheses. Some methodologists, such as Hubert and Schultz (1976), Hubert and Baker (1978), Arabie *et al.* (1978), Baker and Hubert (1981), Panning (1982a), and Noma and Smith (1985), propose the use of the common nonparametric randomization test in which all possible ways of placing g objects (or actors) into B

cells (or blockmodel positions) are considered. For each permutation of the data, an index can be computed, comparing the particular permutation to the blockmodel "prediction". An index is then computed, measuring how close each permutation of the data is to the prediction (or in general, a "hypothesis" matrix), thus generating an entire distribution of indices (called the permutation distribution). One of these permutations is the observed blockmodel or partition, actually derived from the relational data.

A permutation test of how well the predicted blockmodel or hypothesis fits the data is conducted simply by determining the fraction of the permutations that fit worse than the one actually observed (that is, the fraction of permutations that have indices indicating fits that are worse). The p-value for the test is this fraction, which is read from the permutation distribution as the tail probability beyond the index calculated for the observed blockmodel. This approach to data analysis, sometimes referred to as *combinatorial data analysis* (see Hubert and Schultz 1976; Hubert 1983, 1985, 1987; and Hubert and Arabie 1989), is quite similar to the approach to "testing" or evaluating blockmodels advocated by White *et al.* (1976) and White (1977), which is implemented by the BLOCKER algorithm of Heil and White (1976). Further, since permutation tests are nonparametric (that is, they make no assumptions about underlying distributions for the data), they can be used in a very wide range of network data analysis situations. Network researchers such as Laumann *et al.* (1974, 1977) and Krackhardt (1987b, 1988) have questioned the use of standard significance tests for the comparison of networks. Permutation tests are a nice response to these concerns.

The second approach, as mentioned, is based on statistical theory for social network data. This idea uses a statistical or *stochastic blockmodel* to mathematically represent the equivalence classes defined on the actors. A stochastic blockmodel is a direct generalization of the p_1 class of probability models for social networks. This approach was introduced by Holland *et al.* (1983) and Wasserman and Anderson (1987), and generalized by Breiger (1981b), Frank *et al.* (1985a, 1985b) and Wang and Wong (1987).

Stochastic blockmodels use a different definition of equivalence. Specifically, they are based on *stochastic equivalence*. In brief, one first assumes a random directed graph distribution, such as p_1, and then focuses on the actor parameters. Two actors are stochastically equivalent if we can interchange their parameters, without changing any of the probabilities of the distribution. Clearly, this approach is useful if a researcher is willing to assume that his or her data have been generated by a stochastic process; further, the task is simplified if he or she is willing to adopt p_1 for this stochastic mechanism. As we have mentioned, it is relatively easy to assess how well a stochastic blockmodel "fits" a data set, since goodness-of-fit statistics are a natural by-product of the statistical modeling process.

We begin by introducing several goodness-of-fit indices for the fit of a blockmodel to a network data set. This methodology will be illustrated on Krackhardt's high-tech managers and the world system network.

7.2. Goodness-of-fit statistics for blockmodels

Consider a specific blockmodel, presented in a $B \times B \times R$ matrix, labelled $\boldsymbol{B}$, whose entries (denoted by $\{b_{klr}\}$) tell how the positions are related on the various relations. This blockmodel $\boldsymbol{B}$ presents all the relational linkages among positions of approximately equivalent actors. The quantity b_{klr} equals 0 if there is not (or 1 if there is) a linkage from position $\mathscr{B}_k$ to position $\mathscr{B}_l$ on relation χ_r. A blockmodel also contains a mapping function defined on the actors, ϕ, that tells which position each actor belongs to. We will use this mapping function, as well as the blockmodel $\boldsymbol{B}$ in this section.

One can view a blockmodel as an idealization, or an optimal model, in which actors in a specific position are predicted to be perfectly structurally equivalent. In practice, actors in a position are only approximately structurally equivalent. Nonetheless, we usually like to see how close to the ideal the actors actually are; in other words, how approximate are our approximate structural equivalences. If the optimal model holds, then all actors in a position are exactly structurally equivalent.

There are two ways to evaluate the goodness-of-fit of a hypothesized blockmodel. The first way compares the blockmodel image matrices to the densities of relational ties within and among positions. The second way compares the observed relational ties among actors to the relational ties predicted by the blockmodel.

Suppose we permute the original sociomatrices, so that the order of the actors matches the assignment of actors to positions, and then consider the submatrices that arise due to the partitioning of the g actors into B positions. All of the actors with the same value of the mapping function ϕ will be assigned to the same position. The first g_1 rows and columns of the permuted matrix will contain all the actors i with $\phi(i) = \mathscr{B}_1$, the next g_2 rows and columns will contain all the actors with $\phi(i) = \mathscr{B}_2$, and so forth. Each submatrix, which in general, will be of dimension $g_k \times g_l$, will have a density of ones, equal to the proportion of ties that are actually present between actors. If all actors are perfectly structurally equivalent, this density will be either zero or one, as specified (or predicted) by the blockmodel. The entire set of densities can then be compared to the blocks, or entries in the image matrices, to determine how well a blockmodel fits (that is, how close to optimal the blockmodel really is).

If all the densities are zeros and ones, the blockmodel fits perfectly, since the actors within the positions are exactly structurally equivalent. In this instance, all blockmodel criteria (including fat fits, lean fits, and α blockmodels) yield exactly the same $\boldsymbol{B}$. But rarely is this the case. To evaluate the fit of a blockmodel, we need methods and measures to compare the image matrices with the matrices of densities.

Alternatively, one can compare the original sociomatrices, which generate the image matrices, to their "predictions" under the blockmodel. Let us assume that actor i is in position $\mathscr{B}_k$ and actor j is in position $\mathscr{B}_l$. The predicted value for the relational tie (on χ_r) from actor i to actor j is equal to the link in the reduced graph from position $\mathscr{B}_k$ to position $\mathscr{B}_l$. If this link is present, the predicted value is unity, then all the actors in $\mathscr{B}_k$ are predicted to have relational ties to the actors in position $\mathscr{B}_l$; otherwise, if the arc is not present, the predicted value is zero. In this section, we discuss how to do these comparisons in order to evaluate how well a blockmodel fits a specific network data set.

7.2.1. Comparing observed densities to a target blockmodel

As mentioned earlier, the elements in the observed sociomatrices are usually aggregated across actors to yield the densities of ties within each of the positions. The density of ties within and between positions are given in Equations (1) and (2), above, and can be viewed as elements of a $B \times B \times R$ matrix, $\boldsymbol{\Delta}$. The matrix $\boldsymbol{\Delta}$ is the density table discussed earlier.

The first goodness-of-fit index for a blockmodel simply compares these densities to the blocks, or elements of the blockmodel $\boldsymbol{B}$. Clearly the comparison of these two matrices, $\boldsymbol{B}$ and $\boldsymbol{\Delta}$, ignores which actors are in which positions; indeed, all that matters here is how well the image matrix (without the mapping function) models the average of the relational ties of the actors in the B positions (as reflected by the densities). Only when all the densities are 0 or 1, will a blockmodel fit perfectly.

Thus, one measure of how well a blockmodel fits a data set is based on the differences between the elements of $\boldsymbol{\Delta}$, and the elements of $\boldsymbol{B}$. If the blockmodel is constructed using the lean fit criterion, then a $b_{klr} = 0$ only when the corresponding density $\Delta_{klr} = 0$. Thus, any submatrices with ones become oneblocks. Lean fits are sometimes called "zeroblock" fits, since only zeroblocks are fit perfectly. A "oneblock" fit is the opposite of a lean fit: $b_{klr} = 1$ only when the corresponding density $\Delta_{klr} = 1$. If the blockmodel is a fat fit, then the fit is a combination of a zeroblock and a oneblock fit: a $b_{klr} = 0$ only when the corresponding density $\Delta_{klr} = 0$, *and* a $b_{klr} = 1$ only when the corresponding density $\Delta_{klr} = 1$. Clearly, fat fits are perfect structural equivalence fits. They fit exactly only when all actors within all positions are exactly structurally equivalent. We note that for a lean fit, $b_{klr} = 1$ does not imply that the corresponding density Δ_{klr} is actually 1; in fact, the density only needs to be greater than zero. And for a oneblock fit, $b_{klr} = 0$ does not imply that the corresponding density $\Delta_{klr} = 0$. These very strict definitions usually force most researchers to construct blockmodels using the more realistic α-fit criterion. Regardless of the criterion chosen, it is rare for densities to be only zeros and ones, so that the b's, which, by definition, can only be zeros and ones, will rarely exactly equal the Δ's.

A very simple goodness-of-fit index is the sum of the absolute differences between the elements of $\boldsymbol{\Delta}$ and the elements of $\boldsymbol{B}$. Specifically, we calculate

$$\delta_{b1} = \sum_{r=1}^{R} \sum_{k=1}^{B} \sum_{l=1}^{B} |b_{klr} - \Delta_{klr}|. \tag{4}$$

This index, which varies from 0 to $B \times B \times R = RB^2$ (the number of entries in $\boldsymbol{B}$), attains the maximum if the fitted blockmodel is completely reversed (0's instead of 1's and 1's instead of 0's) from the observed densities. The smaller it is, the better the fit. This measure is a crude indicator of fit.

Another measure originated with Carrington *et al.* (1979/80), and Carrington and Heil (1981). Their index is constructed using a minimum chi-squared argument, and by considering the worst possible fitting blockmodel for a particular data set. Such a fit would have ones where zeros belong, and vice versa. This index is appropriate when the blockmodel is constructed from an α-fit criterion. The worst possible α-fit arises when the observed density for a given block is exactly equal to α, since a slight change in the density necessitates that the "fitted" block value be changed from a 0 to a 1, or vice versa. The "goodness-of-fit" for a particular block then depends simply on how close Δ_{klr} is to α. Details on the construction of this index can be found in Carrington *et al.* (1979/80), Carrington and Heil (1981) and Chapter 18 of Wasserman and Faust (1992).

Evaluating δ_{b1} and related indices is difficult. There is no statistical theory or distribution for the indices. One could use a permutation test, permuting the actors to arrive at a different assignment of actors to positions, and hence, an entire collection of $\boldsymbol{\Delta}$ matrices. There will be one $\boldsymbol{\Delta}$, and hence one δ_{b1}, for every possible permutation of actors to positions. We have found that permutation tests are quite useful. Some of these tests are implemented in UCINET (Borgatti *et al.* 1991).

The methodology just discussed examines blocks, and the properties of blocks, and ignores the relational ties among the actors. We now turn to the second approach for blockmodel goodness-of-fit, comparing the actual observed data to the relational linkages predicted by a blockmodel.

7.2.2. Comparing observed relational linkages to a target blockmodel

Take the observed sociomatrices, with entries x_{ijr}, and use the mapping function for the blockmodel under consideration to arrive at the target, or "predicted" collection of relational linkages for each pair of actors on each relation. The blockmodel classifies actor i into position $\mathscr{B}_{\phi(i)}$, and actor j into position $\mathscr{B}_{\phi(j)}$, and the image matrix tells whether relational ties are present among and between positions. So, the "target" matrix is a hypothesized sociomatrix in which all actors in a position have identical ties to and from actors in other positions.

Thus, the predicted value, $x_{ijr}^{(t)}$ for actors i and j on relation r is

$$x_{ijr}^{(t)} = b_{\phi(i)\phi(j)r}, \tag{5}$$

indicating whether the actors in the same position as i ($\mathscr{B}_{\phi(i)}$) are predicted by the blockmodel to choose the actors in the same position as j ($\mathscr{B}_{\phi(j)}$). We can view the $x^{(t)}$'s as elements of a target array (or hypothesis matrix), to which we compare the x's, the actual sociomatrix entries. The superscript (t) indicates that the matrix (or its entries) is the "target" matrix, calculated from the blockmodel. We should note that this methodology is quite flexible, and can be used to compare $\boldsymbol{x}$ to any "hypothesis" matrix, even if the target or hypothesis matrix $\boldsymbol{x}^{(t)}$ is not generated from any particular blockmodel. Such is also the nature of permutation/randomization tests.

A number of goodness-of-fit indices exist for quantifying how close the observed $\boldsymbol{x}$ is to the target $\boldsymbol{x}^{(t)}$. Each index measures the similarity (or dissimilarity) of the target and the actual relational data. The index of choice depends on the advantages and disadvantages of each, as we discuss below. We should note that there is no parametric statistical theory for any of them, but all can be evaluated using a nonparametric, randomization test approach.

For example, we can calculate the sum of the absolute differences between the entries of the observed and target matrices:

$$\delta_{x1} = \sum_{r=1}^{R} \sum_{i=1}^{g} \sum_{j=1}^{g} \left| x_{ijr} - x_{ijr}^{(t)} \right|. \tag{6}$$

The value of δ_{x1} is the number of entries in the observed sociomatrix, $\boldsymbol{x}$ that are not identical to their predicted values in the target matrix, $\boldsymbol{x}^{(t)}$. It is thus a measure of dissimilarity between the two matrices. As with δ_{b1}, shown in Equation (4), there is no statistical theory (except through the use of permutation tests) to evaluate the statistical significance of a particular value of δ_{x1}. The measure, δ_{x1}, comparing a sociomatrix to a target matrix, is a simple function of measure δ_{b1}, given in Equation 4:

$$\delta_{x1} = g \times (g-1)\delta_{b1}. \tag{7}$$

The network analysis program UCINET (Borgatti *et al.* 1991) calculates a match coefficient, which we will denote δ_{x2}, which is very closely related to δ_{x1}. The match coeficient is the *proportion* of entries in $\boldsymbol{x}$ that are identical to $\boldsymbol{x}^{(t)}$. Since there are $g \times (g-1)$ entries in the sociomatrix, the value of δ_{x2} is:

$$\delta_{x2} = 1 - \frac{\sum_{r=1}^{R} \sum_{i=1}^{g} \sum_{j=1}^{g} \left| x_{ijr} - x_{ijr}^{(t)} \right|}{g \times (g-1)} = 1 - \frac{\delta_{x1}}{g \times (g-1)}. \tag{8}$$

The measure, δ_{x2} is a similarity measure; larger values indicate a closer fit between the observed sociomatrix and the target matrix.

A third index is the "matrix correlation", calculated using all the elements in $\boldsymbol{x}$ and the elements in $\boldsymbol{x}^{(t)}$ (excluding diagonal elements). This index, which we will call δ_{x3}, has been used by many researchers over the years (see Arabie *et al.* 1978). It was labeled $\Gamma(\boldsymbol{X}, \boldsymbol{X}^{(t)})$ by Hubert and Baker (1978). It is best utilized if the target sociomatrices are based on the lean fit blockmodel criterion since such fits yield targets containing more ones and fewer zeros (correlation coefficients are better measures of association if the data arrays do not contain too many zeros). As noted above, one can calculate δ_{x2} for any target or hypothesis matrix, and use it to evaluate the "hypothesis" that generated that particular $\boldsymbol{x}^{(t)}$.

Panning (1982a) and Noma and Smith (1985) recommend the use of the squared multiple correlation coefficient R^2 as a goodness-of-fit measure to evaluate the fit of a blockmodel to data, and show how to use the statistic to find a "best-fitting" blockmodel. Both authors note that the predicted value for a specific submatrix from the original sociomatrices is either a 0 or a 1, and that the sum of squares of deviations of the entries in the submatrix from this predicted value can be used as a measure of fit. From the sums of squares for all submatrices or blocks, one can calculate both a within-block sum of squares and a sum of squares deviation from the grand mean (analogous to within-subjects and total sums of squares in a one-way analysis of variance). The index R^2 is simply the ratio of these two sums of squares, subtracted from unity. The within-block or numerator sum of squares is the "total unexplained sum of squares", and the measure increases as this quantity becomes small, relative to the total (or denominator) sum of squares. Illustrations of the use of this index can be found in Panning (1982a, 1982b), and Noma and Smith (1985). Calculations are detailed by Panning (1982a).

The utility of this index rests on an argument that blockmodeling is actually a form of an analysis of variance, in which the "independent" variables are the block densities, and the "dependent" variables are the observed entries within the sociomatrices. Hence, an "optimal" blocking should maximize the percent of explained variance, and lead a researcher to focus on R^2, the statistic that has this property. Panning (1982a) gives a strategy for finding blockmodels that have maximal R^2's. We note that one can generate a permutation distribution for this index (see Noma and Smith 1985) simply by considering all possible permutations of the actors to positions, and calculating R^2 for each permutation. This leads to a valid, nonparametric statistical test for the goodness of an observed fit. The index is also easily used for multiple relation network data sets.

However, as several authors have noted, neither δ_{x3} nor R^2 is well-suited for binary data, and thus, are not recommended. Carrington *et al.* (1979/80) comment on the suitability of the use of correlation coefficients (and hence squared multiple correlation coefficients, which are squares of correlation coefficients themselves) to compare two binary matrices. Faust and Romney (1985b) also comment on the use of correlation coefficients used to compare sociomatrices.

Hubert and Baker (1978) show that for $R = 1$, the expected value of δ_{x3} is zero, where the distribution is taken over all possible permutation assignments of the actors into the prespecified number of positions, B. They also compute the variance of this index. These two results lead nicely to a permutation test which yields the significance level (or p-value) of the observed δ_{x3}. A good illustration of this approach is given by Baker and Hubert (1981). The advantages of this approach are discussed by Arabie *et al.* (1978), and include the fact that the number of actors partitioned into each of the B positions is constrained to match the position sizes actually observed.

One disadvantage of this index is that the results of Hubert and Baker have not been extended to multiple relations, $R > 1$. Further, as noted by Arabie *et al.* (1978), lean fits and their filled-in oneblocks are usually a "poor assumption" about underlying social structure. Thus, indices built around them may not be very accurate.

Other measures of association, comparing $\boldsymbol{x}$ to $\boldsymbol{x}^{(t)}$, can be found in Katz and Powell (1953), Hubert and Baker (1978), Zegers and ten

Berge (1985), and Wasserman (1987). Some of them are implemented in UCINET (Borgatti *et al.* 1991). Carrington *et al.*'s (1979/80) measure also falls into this category, since it can be written as a function of the observed data, rather than the observed densities (see Equation 7). For $R = 1$, one can view this problem as a birelational network analysis, where the two relations are the observed and the target. In this setting, Wasserman (1987) shows how to compare an observed to a target sociomatrix using dyadic interaction statistical models.

Examples. We will now illustrate blockmodel evaluation using a permutation test to compare the sociomatrix $\boldsymbol{x}$ to $\boldsymbol{x}^{(t)}$. Calculations were done using the quadratic assignment program in UCINET 4.0 (Borgatti *et al.* 1991). In each case we ran 1000 permutations of the rows (and columns) of the sociomatrix. Tables 8 and 9 report the calculated values of δ_{x2} the match coefficient (Equation 8), and δ_{x3} the matrix correlation for each of the examples. We also report the number of permutations out of 1000 in which the value of δ_{x2} or δ_{x3} was greater than the calculated value for $\boldsymbol{x}$ and $\boldsymbol{x}^{(t)}$.

The results in Table 8 show that for the World system example, for each relation, the ties predicted by the blockmodel image matrices (including the assignment of actors to positions, and the statement of relational ties among positions) are closer to the observed values of the relational variable than to any other assignment of actors to positions. So, for the World system example, the blockmodel image matrices are good representations of the relational ties among countries. Similarly, the results for Krackhardt's high-tech managers in Table 9 show that the blockmodels are good representations of the

Table 8
Comparison of observed relations to target blockmodels – world system example

Relation	Measure	
	δ_{x3}	δ_{x2}
Manufactured goods	0.687	0.844
	(0.000)	(0.000)
Crude materials	0.593	0.799
	(0.000)	(0.000)
Diplomatic ties	0.588	0.801
	(0.000)	(0.000)

Table 9
Comparison of observed relations to target blockmodels – Krackhardt's high-tech managers example

Relation	Measure	
	δ_{x3}	δ_{x2}
Advice	0.512	0.752
	(0.000)	(0.000)
Friendship	0.238	0.614
	(0.001)	(0.001)

relational ties among the managers. However, for the friendship relation there is one permutation of actors (out of the 1000 random permutations) that would better match the observed sociomatrix than does the proposed assignment of actors from the blockmodel.

It is not surprising that, for both the World system example and Krackhardt's high-tech managers, the relational ties predicted by the blockmodel are extremely similar to the observed relational ties. After all, the blockmodels were constructed from the relational ties in the first place. In an exploratory study, the researcher often seeks the "best" blockmodel of a given data set. In such a case if the permutation test shows that there are one or more assignments of actors to positions that yield a better match between the observed data and the target blockmodel, then the researcher might be interested in studying this better assignment of actors to positions. This strategy of assigning actors to positions in order to optimize an objective function (such as δ_{x2} or δ_{x3}), is an promising way to construct blockmodels. The direct construction of blockmodels has received considerable attention recently (Arabie *et al.* 1990; Batagelj *et al.* 1992a, 1992b).

7.2.3. Comparing observed relational linkages to a theoretical blockmodel

Permutation tests can also be used to compare the observed relational ties in a set of network data to a blockmodel that represents a theoretical structure. Figures 5 and 6 present examples of blockmodel image matrices for some theoretically important structures, such as cohesive subgroups, a transitive structure, and a center-periphery structure. One can evaluate how well a specific theoretical structure represents a given set of network data by constructing the target sociomatrix, $x^{(t)}$, from the hypothesized theoretical structure.

Constructing the target sociomatrix, $x^{(t)}$, requires several steps. The first step is to partition actors into positions. This partition could be the result of a positional analysis in which approximately equivalent actors are assigned to the same position (for example, using hierarchical clustering or CONCOR). The second step is to specify, for each pair of positions, whether a relational tie is present or absent. For some structures, such as cohesive subgroups, this decision is straightforward, since in a cohesive subgroup structure, relational ties are hypothesized only within, and not between, positions. However, for other structures, such as a hierarchy, or a center-periphery structure, the hypothesized presence or absence of a tie between positions depends on where the positions are "located" in the structure. For example, in a hierarchy, relational ties are directed from "lower" positions to "higher" positions. Thus, the order of positions in the blockmodel is important. One way to arrive at an ordering is to consider theoretically important attributes of actors in the positions. For example, one could hypothesize that relational ties of advice in a blockmodel of an organization form a transitive system in which ties are directed from each position to all positions whose members have, on average, longer tenure in the organization. Thus, the positions would be ordered in term of the average tenure of members. Finally, the target sociomatrix is constructed, as usual, using equation 5. A relational tie is hypothesized to be present from actor i to j if there is a hypothesized relational tie from position $\phi(i)$ to position $\phi(j)$ in the theoretical structure that is being evaluated.

Let us now turn to some examples to illustrate the evaluation of theoretical structures.

Examples. We will look first at the advice relation for Krackhardt's high-tech managers, and then study the three relations in the world system example.

First, let us investigate the hypothesis that advice seeking among Krackhardt's high-tech managers forms a transitive system in which members of each position seek advice from members of positions that contain managers with longer tenure in the company. Table 3 presented the mean tenure in years for managers in each position. According to these means, members of position $\mathscr{B}_3$ have been with the company longest on average, followed in order by positions $\mathscr{B}_1$, $\mathscr{B}_2$, and $\mathscr{B}_4$. Thus, the blockmodel for the theoretical structure would be:

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

Notice that the oneblocks do not occupy the upper right, or lower left, triangles of this image matrix, since the positions are not ordered from most to least tenure.

The matrix correlation (δ_{x3}) between relational ties predicted by this transitive blockmodel structure, and the observed ties on the advice relation is equal to 0.269, and the match coefficient (δ_{x2}) is equal to 0.460. Both of these are the largest values out of 1000 random permutations of the rows and columns of the sociomatrix. Thus, the model of transitivity among positions, based on the average tenure of actors in the positions, fits the observed data quite well.

Let us turn now to the world system example. As noted above, numerous authors have hypothesized that the world political and economic system is a center-periphery structure, in which more developed nations occupy central positions, and less developed nations occupy peripheral positions. In the world system example we have already ordered the positions from most central, $\mathscr{B}_1$, to least central, $\mathscr{B}_6$, based on the three image matrices presented in Figure 2. This ordering also appears to correspond well to the attributes of the positions, presented in Table 4.

A permutation test can be used to compare the target sociomatrix based on a center-periphery structure with the observed relational ties for each relation in the world system example. The theoretical blockmodel image for a center-periphery structure was constructed with $B = 6$ positions, and oneblocks in the upper left triangle of the image matrix (see Figure 6). The target sociomatrix for this image matrix was then compared to each of the three relations. For the manufactured goods relation, $\delta_3 = 0.536$, and $\delta_2 = 0.772$; both are the largest out of 1000 permutations. For the crude materials relation, $\delta_3 = 0.532$, and $\delta_2 = 0.799$; both are largest out of 1000 permutations. Finally, for the diplomatic ties relation, $\delta_3 = 0.513$, and $\delta_2 = 0.759$; both are the largest out of 1000 permutations. So, this assignment of countries to positions, and this ordering of positions, matches a center-periphery structure quite well.

We turn now to a brief summary and discussion of the results, with a comparison of the various approaches to blockmodel interpretation and evaluation.

8. Summary

In this paper we have discussed and illustrated several approaches to the interpretation and evaluation of blockmodels. Ideally, the several approaches to blockmodel interpretation should yield complementary insights into a network data set. This seems to be the case for the examples discussed in this paper.

For Krackhardt's high-tech managers, the blockmodel of the advice and friendship relations gives a consistent picture of positions within the corporation. Positions can be ordered in a transitive hierarchy, with position $\mathscr{B}_3$ at the top, followed by position $\mathscr{B}_1$, and then the other two positions. This is confirmed by the permutation test. Position $\mathscr{B}_3$, the "top" of the hierarchy, contains the oldest managers, and the managers with the longest tenure in the company. Furthermore, $\mathscr{B}_3$ and was identified as a Primary position on both the friendship and advice relations. Position $\mathscr{B}_1$ contains the managers with the second longest tenure, on average. Position $\mathscr{B}_1$ was identified as a Primary position on the friendship relation.

The world system example seems to be consistent with a center-periphery structure, as confirmed by the permutation test. The ordering of positions from center to periphery is also related to the characteristics of the countries. Core positions (for example $\mathscr{B}_1$, $\mathscr{B}_2$ and $\mathscr{B}_3$) have lower rates of population growth, higher secondary school enrollment, and higher energy consumption per capita, whereas peripheral positions (for example, $\mathscr{B}_4$, $\mathscr{B}_5$, and $\mathscr{B}_6$) have higher rates of population growth, lower rates of secondary school enrollment, and lower energy consumption per capita.

Hopefully, a greater emphasis on interpretation and evaluation of blockmodels, as one part of a complete positional analysis, will allow researchers to employ blockmodelling more effectively in substantive and theoretical investigations, and to use blockmodels to evaluate structural theories.

References

Anderson, C., S. Wasserman and K. Faust
1992 "Building stochastic blockmodels." *Social Networks 14*: 137–159.

Arabie, P.
1984 "Validation of sociometric structure by data on individuals' attributes". *Social Networks 6*: 373–403.

Arabie, P. and S.A. Boorman
1982 "Blockmodels: Developments and prospects", in: H.C. Hudson (ed.) *Classifying Social Data.* San Francisco: Jossey-Bass.

Arabie, P., S.A. Boorman and P.R. Levitt
1978 "Constructing blockmodels: How and why". *Journal of Mathematical Psychology 17*: 21–63.

Arabie, P., L.J. Hubert and S. Schleutermann
1990 "Blockmodels from the bond energy approach". *Social Networks 12*: 99–126.

Baker, F.B. and L.J. Hubert
1981 "The analysis of social interaction data". *Sociological Methods and Research 9*: 339–361.

Batagelj, V., A. Ferligoj and P. Doreian
1992a "Direct and indirect methods for structural equivalence". *Social Networks 14*: 63–90.

Batagelj, V., P. Doreian and A. Ferligoj
1992b "An optimizational approach to regular equivalence". *Social Networks 14*: 121–135.

Boorman, S.A. and H.C. White
1976 "Social structure from multiple networks II. Role structures". *American Journal of Sociology 81*: 1384–1446.

Borgatti, S.P., M.G. Everett, B. MacEvoy and L. Freeman
1991 "UCINET 4.0". Unpublished manuscript.

Breedlove, W.L. and P.D. Nolan
1988 "International stratification and inequality 1960–1980". *International Journal of Contemporary Sociology 25*: 105–123.

Breiger, R.L.
1976 "Career attributes and network structure: A blockmodel study of a biomedical research specialty". *American Sociological Review 41*: 117–135.
1981a "Structures of economic interdependence among nations", in: P.M. Blau and R.K. Merton (eds.) *Continuities in Structural Inquiry*, pp. 353–380. Newbury Park: Sage.
1981b "Comment on Holland and Leinhardt, 'An exponential family of probability distributions for directed graphs'". *Journal of the American Statistical Association 76*: 51–53.

Breiger, R.L. and P.E. Pattison
1986 "Cumulated social roles: The duality of persons and their algebras". *Social Networks 8*: 215–256.

Breiger, R.L., S.A. Boorman and P. Arabie
1975 "An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling". *Journal of Mathematical Psychology 12*: 328–383.

Burt, R.S.
1976 "Positions in networks". *Social Forces 55*: 93–122.
1986 "A cautionary note". *Social Networks 8*: 205–211.
Carrington, P. J. and G.H. Heil
1981 "COBLOC: A hierarchical method for blocking network data". *Journal of Mathematical Sociology 8*: 103–131.
Carrington, P.J., G.H. Heil and S.D. Berkowitz
1979/80 "A goodness-of-fit index for blockmodels". *Social Networks 2*: 219–234.
Coxon, A.P.M.
1982 *The User's Guide to Multidimensional Scaling.* Exeter, NH: Heinemann.
Cronbach, L.J. and G.C. Gleser
1953 "Assessing similarity between profiles". *Psychological Bulletin 50*: 456–473.
Doreian, P.
1988 "Equivalence in a social network". *Journal of Mathematical Sociology 13*: 243–282.
Doreian, P. and T.J. Fararo
1985 "Structural equivalence in a journal network". *Journal of the American Society for Information Science 36*: 28–37.
Ennis, J.G.
1982 "Blockmodels and spatial representations of group structure: Some comparisons", in: H.C. Hudson (ed.) *Classifying Social Data,* pp. 199–214. San Francisco: Jossey-Bass.
Europa Publications
1984 *Europa Year Book.* London: Europa Publications.
Everett M.G. and S.P. Borgatti
1992 "Regular colouring of diagraphs, networks and hypergraphs". *Social Networks 14*: 91–120.
Faust, K.
1988 "Comparison of methods for positional analysis: Structural and general equivalences". *Social Networks 10*: 313–341.
Faust, K. and A.K. Romney
1985a "Does STRUCTURE find structure?: A critique of Burt's use of distance as a measure of structural equivalence". *Social Networks 7*: 77–103.
1985b "The effect of skewed distributions on matrix permutation tests". *British Journal of Mathematical and Statistical Psychology 38*: 152–160.
Fox, J.
1982 "Selective aspects of measuring resemblance for taxonomy", in: H.C. Hudson (ed.) *Classifying Social Data.* San Francisco: Jossey-Bass.
Frank, O., H. Komanska and K.F. Widaman
1985a "Cluster analysis of dyad distributions in networks". *Journal of Classification 2*: 219–238.
Frank, O., M. Hallinan and K. Nowicki
1985b "Clustering of dyad distributions as a tool in network modeling". *Journal of Mathematical Sociology 11*: 47–64.
Galaskiewicz, J. and K.R. Krohn
1984 "Positions, roles and dependencies in a community interorganizational system". *The Sociological Quarterly 25*: 527–550.
Harary, F., R.Z. Norman and D. Cartwright
1965 *Structural Models: An Introduction to the Theory of Directed Graphs.* New York: Wiley.
Heil, G.H. and H.C. White
1976 "An algorithm for finding simultaneous homomorphic correspondences between graphs and their image graphs". *Behavioral Science 21*: 26–35.

Holland, P.W., K.B. Laskey and S. Leinhardt
1983 "Stochastic blockmodels: Some first steps". *Social Networks 5*: 109–137.
Hubert, L.J.
1983 "Inference procedures for the evaluation and comparison of proximity matrices", in: J. Felsenstein (ed.) *Numerical Taxonomy.* New York: Springer.
1985 "Combinatorial data analysis: Association and partial association". *Psychometrika 50*: 449–467.
1987 *Assignment Methods in Combinatorial Data Analysis.* New York: Marcel Dekker.
Hubert, L.J. and P. Arabie
1989 "Combinatorial data analysis: Confirmatory comparisons between sets of matrices". *Applied Stochastic Models and Data Analysis 5*: 273–325.
Hubert, L.J. and F.B. Baker
1978 "Evaluating the conformity of sociometric measurements". *Psychometrika 43*: 31–41.
Hubert, L.J. and L. Schultz
1976 "Quadratic assignment as a general data analysis strategy". *British Journal of Mathematical and Statistical Psychology 29*: 190–241.
Katz, L. and J.H. Powell
1953 "A proposed index of the conformity of one sociometric measurement to another". *Psychometrika 18*: 249–256.
Kick, Edward L.
n.d. "World-system structure, national development, and the prospects for a socialist world order". Unpublished manuscript.
Knoke, D. and D.L. Rogers
1979 "A blockmodel analysis of interorganizational networks". *Sociology and Social Research 64*: 28–52.
Krackhardt, D.
1987a "Cognitive social structures". *Social Networks 9*: 109–134.
1987b "QAP partialling as a test of spuriousness". *Social Networks 9*: 171–186.
1988 "Predicting with networks: Nonparametric multiple regression analyses of dyadic data". *Social Networks 10*: 359–382.
Laumann, E.O., P.V. Marsden and J. Galaskiewicz
1977 "Community-elite influence structures: Extension of a network approach". *American Journal of Sociology 83*: 594–631.
Laumann, E.O., L.M. Verbrugge and F.V. Pappi
1974 "A causal modelling approach to the study of a community elite's influence structure". *American Sociological Review 39*: 164–178.
Lenski, G. and P.D. Nolan
1984 "Trajectories of development: A test of ecological-evolutionary theory". *Social Forces 63*: 1–23.
Light, J.M. and N.C. Mullins
1979 "A primer on blockmodeling procedure", in: P.W. Holland and S. Leinhardt (eds.) *Perspectives on Social Network Research*, pp. 85–118. New York: Academic Press.
Lorrain, F. and H.C. White
1971 "Structural equivalence of individuals in social networks". *Journal of Mathematical Sociology 1*: 49–80.
MacEvoy, B. and L. Freeman
n.d. UCINET, Version 3.0: *A Microcomputer Package for Network Analysis.* Mathematical Social Science Group, School of Social Sciences, University of California, Irvine.
Mandel, M.J.
1983 "Local roles and social networks". *American Sociological Review 48*: 376–386.

Marsden, P.V.
1989 "Methods for the characterization of role structures in network analysis", in: L.C. Freeman, D.R. White and A.K. Romney (eds.) *Research Methods in Social Network Analysis,* pp. 489–530. Fairfax, VA: George Mason University Press.
Mullins, N.C., L.L. Hargens, P.K. Hecht and E.L. Kick
1977 "The group structure of cocitation clusters: A comparative study". *American Sociological Review 42*: 552–562.
Nemeth, R.J. and D.A. Smith
1985 "International trade and world-system structure, A multiple network analysis". *Review 8*: 517–560.
Nolan, P.D.
1983 "Status in the world economy and national structure and development". *International Journal of Comparative Sociology 24*: 109–120.
1987 "World system status, income inequality and economic growth: A criticism of recent criticism". *International Journal of Comparative Sociology 28*: 69–76.
1988 "World system status, techno-economic heritage and fertility". *Sociological Focus 21*: 9–33.
Noma, E. and D.R. Smith
1985 "Benchmark for the blocking of sociometric data". *Psychological Bulletin 97*: 583–591.
Panning, W.H.
1982a "Fitting blockmodels to data". *Social Networks 4*: 81–101.
1982b "Blockmodels: From relations to configurations". *American Journal of Political Science 26*: 585–608.
Pattison, P.E.
1982 "The analysis of semigroups of multirelational systems". *Journal of Mathematical Psychology 25*: 87–118.
1988 "Network models: Some comments on papers in this special issue". *Social Networks 10*: 383–411.
Richards, W.D.
1989 "The NEGOPY Analysis Program". Unpublished manuscript, Department of Communication, Simon Fraser University.
Rohlf, F.J. and R.R. Sokal
1965 "Coefficients of correlation and distance in numerical taxonomy". *The University of Kansas Science Bulletin 45*: 3–27.
Sailer, L.D.
1978 "Structural equivalence: Meaning and definition, computation and application". *Social Networks 1*: 73–90.
Smith, D. and D. White
1988 "Structure and dynamics of the global economy: Network analysis of international trade 1965–1980". Unpublished manuscript.
Sneath, P.H.A. and R.R. Sokal
1973 *Numerical Taxonomy: The Principles and Practice of Numerical Classification.* San Francisco: Freeman.
Snyder, D. and E. Kick
1979 "Structural position in the world system and economic growth 1955–70: A multiple network analysis of transnational interactions". *American Journal of Sociology 84*: 1096–1126.
Sokal, R.R. and P.H.A. Sneath
1963 *Principles of Numerical Taxonomy.* San Francisco: Freeman.
United Nations
1984 *Statistical Papers: Commodity Trade Statistics 1984. Series D. 34* Nos. 1-1 through 1-24.

Wang, Y.J. and G.Y. Wong
1987 "Stochastic blockmodels for directed graphs". *Journal of the American Statistical Association 82*: 8–19.
Wasserman, S.
1987 "Conformity of two sociometric relations". *Psychometrika 52*: 3–18.
Wasserman, S. and C. Anderson
1987 "Stochastic a posteriori blockmodels: Construction and assessment". *Social Networks 9*: 1–36.
Wasserman, S. and K. Faust
1992 *Social Network Analysis: Methods and Applications.* New York and Cambridge, England: Cambridge University Press.
White, D.R. and K.P. Reitz
1989 "Re-thinking the role concept: Homomorphisms on social networks", in: L.C. Freeman, D.R. White and A.K. Romney (eds.) *Research Methods in Social Network Analysis,* pp. 429–488. Fairfax, VA: George Mason University Press.
White, H.C.
1977 "Probabilities of homomorphic mappings from multiple graphs". *Journal of Mathematical Psychology 16*: 121–134.
White, H.C., S.A. Boorman and R.L. Breiger
1976 "Social structure from multiple networks. I. Blockmodels of roles and positions". *American Journal of Sociology 81*: 730–779.
Wilkinson, L.
1987 *SYSTAT: The System for Statistics.* Evanston, IL: SYSTAT.
Winship, C.
1988 "Thoughts about roles and relations: An old document revisited". *Social Networks 10*: 209–231.
Winship, C. and M. Mandel
1983 "Roles and positions, A critique and extension of the blockmodeling approach", in: S. Leinhardt (ed.) *Sociological Methodology 1983–1984,* pp. 314–344. San Francisco: Jossey-Bass.
World Bank
1983 *World Bank World Tables.* Volumes I and II. Baltimore: The Johns Hopkins University Press.
Wu, L.
1983 "Local blockmodel algebras for analyzing social networks", in: N.B. Tuma (ed.) *Sociological Methodology 1983–84,* pp. 272–313. San Francisco: Jossey-Bass.
Zegers, F. and J. ten Berge
1985 "A family of association coefficients for metric scales". *Psychometrika 50*: 17–24.