Social Networks 14 (1992) 137–161
North-Holland

Building stochastic blockmodels *

Carolyn J. Anderson and Stanley Wasserman

Department of Psychology and Department of Statistics, University of Illinois, Urbana, IL 61801, USA

Katherine Faust

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

The literature devoted to the construction of stochastic blockmodels is relatively rare compared to that of the deterministic variety. In this paper, a general definition of a stochastic blockmodel is given and a number of techniques for building such blockmodels are presented. In the statistical approach, the likelihood ratio statistic provides a natural index to evaluate the fit of the model to the data. The model itself consists of a set of actors partitioned into positions with respect to a definition of equivalence, and a representation based on estimated probabilities. The specific statistical model that is used to illustrate the techniques is p_1, which was first introduced as a method for stochastic blockmodeling by Fienberg and Wasserman (1981), and developed by Holland *et al.* (1983) and Wasserman and Anderson (1987).

1. Introduction

Blockmodels are used to analyze and describe the structure of a group and the positions of individual actors in a group. These tasks, which are standard components of a positional analysis, are achieved through the simplified representations of the patterns in complex social networks that are produced by blockmodels. A blockmodel consists of a mapping of approximately equivalent actors into blocks or positions and a statement regarding the relations between the positions. The

* This research was supported by grants from the Research Board of the University of Illinois. This paper is based, in part, on Chapter 16 of Wasserman and Faust (1993). Electronic mail addresses: canderso@s.psych.uiuc.edu, stanwass@uiuc.edu, n040012@univscvm.bitnet.

data studied by a relational analysis consist of R relational variables measured on g actors where

$$X_{ijr} = \begin{cases} c & \text{if actor } i \text{ relates to actor } j \text{ at level } c \text{ on variable } r,\ i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

These variables are collected into R sets of $(g \times g)$ matrices, $\boldsymbol{X}_1$, $\boldsymbol{X}_2, \ldots, \boldsymbol{X}_R$, referred to as sociomatrices.

The standard approach to blockmodeling seeks to simultaneously permute the rows and columns of sociomatrices to reveal patterns with respect to the entries. Partitions of the g actors into B positions are sought such that actors who are "approximately equivalent" (i.e., actors who exhibit the same patterns of entries in the corresponding rows and columns of the sociomatrices) are assigned to the same position. This approach is deterministic and relies on algorithms to find optimal partitions of actors (e.g. White *et al.* 1976; Boorman and White 1976; Heil and White 1976; and many others). The literature abounds with papers describing this deterministic approach to the construction of blockmodels.

Relative to the standard approach, statistical blockmodels are less well known tools for performing positional analyses. As opposed to deterministic blockmodels, statistical or stochastic ones have the additional advantages of an explicit theoretical model for the relations between actors, a proposed stochastic mechanism, and a natural means of testing the goodness-of-fit of the model to the data. The main purpose of this paper is to illustrate a general approach to building and evaluating blockmodels based on statistical models and theory. The technique will be presented using a specific family of models designed to analyze social network data. Following the work of Holland *et al.* (1983) and Wasserman and Anderson (1987), we employ the specific network model p_1 (Holland and Leinhardt 1981; Fienberg and Wasserman 1981). Various techniques for finding partitions of actors based on the relational data will be described.

In Section 2, additional notation is introduced and a general definition of a stochastic blockmodel is given. To construct a stochastic blockmodel for an observed network, a particular probability distribution and a mapping function need to be designated. Ways of specifying these two basic components are presented in Sections 3 and 4, respectively, along with procedures designed to examine the ade-

quacy of each of them. In Section 5, stochastic blockmodel representation of networks are described. The concepts and techniques reviewed in Sections 2–5 are illustrated in detail in Section 6, where they are used to analyze a network data set.

2. Definition of a stochastic blockmodel

The sociomatrices $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_R$ are matrices of *random* variables. The matrix $\boldsymbol{X}$ represents the entire set of the R random matrices and is referred to as a super-sociomatrix of the "adjacency matrix for a multigraph". The probability distribution for $\boldsymbol{X}$, $p(\boldsymbol{X}) = \Pr(\boldsymbol{X} = \boldsymbol{x})$, gives the probability that various relational linkages between actors across all relations are equal to the specified values in $\boldsymbol{x}$. Random vectors $\mathbf{X}_{ij}$ associated with actors i and j are defined as the set of R relational ties from actor i to actor j: $\mathbf{X}_{ij} = (X_{ij1}, X_{ij2}, \ldots, X_{ijR})$. The basic modeling unit in statistical models for social network data is the dyad. Since the sets of random variables $\mathbf{X}_{ij}$ and $\mathbf{X}_{ji}$ contain all the relational data for the dyad consisting of actors i and j, dyadic random vectors $\mathbf{D}_{ij}$ are defined as $\mathbf{D}_{ij} = (\mathbf{X}_{ij}, \mathbf{X}_{ji})$, $i < j$.

A stochastic blockmodel is based on the probability distribution for $\boldsymbol{X}$, as well as the mapping function that assigns the g actors to the positions $\{\mathscr{B}_s\}$, where $s = 1, 2, \ldots, B$. The assumption of a probability distribution for all the relational ties is the major difference between a stochastic blockmodel and a deterministic one. Specifically,

Definition 1. Let $p(\boldsymbol{x})$ be the probability function for a stochastic multigraph, which is represented by the super-sociomatrix $\boldsymbol{X}$. Further, suppose that $\mathscr{B} = \{\mathscr{B}_1, \mathscr{B}_2, \ldots, \mathscr{B}_B\}$ is a partition of the g actors into the B positions, as specified by a mapping function ϕ. With respect to $\mathscr{B}$, $p(\boldsymbol{x})$ is a stochastic blockmodel if the following two conditions are satisfied:

1. The random dyadic variables $\boldsymbol{D}_{ij}$ are all statistically independent of each other.
2. For any actors $i \neq j$ and $i' \neq j$, if i and k' belong to the same position, then the random dyadic vectors $\mathbf{D}_{ij}$ and $\mathbf{D}_{i'j}$ have the same probability distribution.

This definition states that a stochastic blockmodel consists of a probability distribution and a mapping of actors to positions. If the blockmodel is stochastic, then the relational linkages are assumed to be random variables, and the two stated conditions must be met.

The first condition, which states that the dyads must be independent of each other, places the focus on the dyad, rather than on the individual relational ties $\mathbf{X}_{ij}$. As a result, Holland *et al.* (1983) refer to such stochastic blockmodels as *pair-dependent stochastic blockmodels*. An advantage of modeling dyads is the ability to analyze structural tendencies that occur at the level of the dyad, such as reciprocity. Such tendancies cannot be studied if the sets of random variables $\boldsymbol{X}_{ij}$ and $\mathbf{X}_{ji}$ are assumed to be independent of each other. If actor independence is assumed, then the focus is on the individual relational ties or "choices" $\mathbf{X}_{ij}$. The resulting stochastic blockmodel would be analogous to standard blockmodels where actors are implicitly assumed to be independent entities who do not take into account the choices made of them when deciding which choices to make. In many instances, actor independence is an unrealistic assumption, unlike the assumption of dyadic independence.

The second condition states that if two actors are in the same position, then the choices that they "make" and "receive" are governed by the same probability distribution. This implies that the calculated probabilities using $p(\boldsymbol{x})$ are not changed by interchanging actors belonging to the same position. This fact leads to a definition of "stochastic equivalence". Two actors i and i' are said to be *stochastically equivalent* if the probability of i relating to and being related to by every other actor in the group is the same as the probability for actor i'. Formally,

Definition 2. Given a stochastic multigraph, represented by the set of random matrices $\boldsymbol{X}$, actors i and i' are stochastically equivalent if and only if the probability of any event concerning $\boldsymbol{X}$ is unchanged by interchanging actors i and i'.

Stochastic equivalence is a generalization of structural equivalence, a central and important concept in standard blockmodel analyses. Two actors i and i' are structurally equivalent if and only if i relates to and is related to by all the other actors in exactly the same way that i' relates to and is related to by all the other actors. Structural equivalence implies stochastic equivalence, but not vice versa. Empiri-

cally, actors rarely exhibit perfect structural equivalence, and researchers usually need to assume some form of approximate structural equivalence or an alternative equivalence definition (such as regular or automorphic equivalence) to partition actors into positions. With stochastic equivalence, the relational linkages need not be identical for actors to be equivalent.

3. Specific statistical models

The specific probability distribution that is assumed in the paper is p_1 (Fienberg and Wasserman 1981; Holland and Leinhardt 1981; Wasserman and Faust 1993). The blockmodels described here are constructed by constraining the parameters associated with individual actors to be equal for all actors within positions (Fienberg and Wasserman 1981; Holland *et al.* 1983; Wasserman and Anderson 1987). In Section 3.1, the basic blockmodel for one binary variable (i.e., $R = 1$ and $X_{ij1} = X_{ij} = 0, 1$) is described in detail, and is followed in Section 3.2 by a review of some alternatives and extensions. In Section 3.3, the likelihood ratio statistic is proposed as an index for measuring the goodness-of-fit of stochastic blockmodels.

3.1. Stochastic blockmodel based on p_1

Since dyads are the basic units modeled by p_1, it is convenient to define a matrix of indicator variables **Y** as follows:

$$Y_{ijkl} = \begin{cases} 1 & \text{if } X_{ij} = k \text{ and } X_{ji} = l \text{ for } k, l = 0, 1 \\ 0 & \text{otherwise.} \end{cases}$$

Given **Y**, p_1 can be fit using standard loglinear modeling procedures (Fienberg and Wasserman 1981; Wasserman and Faust 1993, Ch. 15).

Assuming p_1, the probability distribution for the dyad consisting of actors i and j is

$$\Pr(Y_{ijkl} = 1) = \exp\{\lambda_{ij} + k\alpha_i + k\beta_j + l\alpha_j + l\beta_i + (k+l)\theta + kl\rho\} \qquad (1)$$

where λ_{ij} ensures that $\Sigma_k \Sigma_l Y_{ijkl} = 1$, and $\Sigma_i \alpha_i = \Sigma_j \beta_j = 0$. The parameters α_i and α_j represent the sending or *expansiveness* effects of actors i and j, respectively, and the parameters β_i and β_j represent the receiving or *popularity* effects of actors i and j, respectively. The parameter θ is an overall choice effect, and the last parameter, ρ, reflects the tendency of relationships in the network to be reciprocal. Dyads are assumed to be statistically independent, which is one of the conditions for a stochastic blockmodel. The full probability distribution $p(\boldsymbol{x})$ is found by multiplying Equation (1) over all $\binom{g}{2}$ dyads.

Assume for now that a function ϕ exists that maps the g actors onto the B positions $\mathscr{B}_1, \mathscr{B}_2, \ldots, \mathscr{B}_B$. (Techniques for obtaining such a function are described in Section 4.) When the α's and β's, which depend only on the individual actors, are equivalent for actors within positions, the second condition for a stochastic blockmodel is met. In other words, for actors i and i' within position $\mathscr{B}_s$,

$$\begin{aligned} \alpha_i &= \alpha_{i'} = \alpha_{[s]} \\ \beta_i &= \beta_{i'} = \beta_{[s]}. \end{aligned} \tag{2}$$

The index s in the bracketed subscripts indicates positions. Interchanging actors i and i' does not change the probability distribution. When the equality condition in (2) is imposed, the number of independent parameters to be estimated is reduced by $2(g - B)$. The parameters α and β are now associated with positions, rather than individual actors.

The model, p_1 with condition (2) imposed, is fit by aggregating the observed y_{ijkl} values within positions and fitting the appropriate loglinear model to the aggregated data. The aggregated data consist of a w-array where $w_{stkl} = \Sigma_{i \in \mathscr{B}_s} \Sigma_{j \in \mathscr{B}_t} y_{ijkl}$. Any one of a number of statistical packages that fit standard loglinear models or programs that use a generalized iterative scaling algorithm will perform this task and yield the maximum likelihood fitted values $\hat{y}_{ijkl}$. For details, see Fienberg *et al.* (1985); Iacobucci and Wasserman (1987); Wang and Wong (1987); Wasserman and Faust (1993). The maximum likelihood estimates of the model parameters can be computed from the fitted values. The next version of *UCINET*, version 4, will contain programs to fit p_1 and compute parameter estimates.

Since $\log \hat{y}_{ijkl}$ is a linear function of the model parameters, the parameters can be estimated by setting up an appropriate design

matrix and using least squares linear regression. The only technical problem that can arise occurs when the sociomatrix has rows and/or columns that contain all zeros or all ones. In this instance, the associated parameter estimates are $-\infty$ or $+\infty$, respectively, and should not be used when centering the remaining parameter estimates to sum to zero. The $\hat{y}_{ijkl}$'s associated with such rows and/or columns should be deleted from the vector of fitted values and the design matrix adjusted accordingly before estimating the remaining parameters. Additional technical discussion is given by Wasserman and Faust (1993).

3.2. *Modifications, extensions and generalizations*

Stochastic blockmodels can be based on variations of the basic model given in Equation (1). For example, some combination of the parameters ρ, $\{\alpha_i\}$, and $\{\beta_i\}$ could be set equal to zero to yield simpler models. Alternatively, if p_1 does not adequately capture the relational ties among individual actors, more complex models can be generated by introducing additional parameters, such as reciprocity parameters that depend on individual actors (i.e., $\{\rho_i\}$, as was proposed by Fienberg and Wasserman (1981) and Wasserman and Galaskiewicz (1984). If either a simpler or more complex model is chosen to represent the relational ties, a stochastic blockmodel can be constructed based on this chosen model by imposing constraints analogous to (2); namely, the parameters that depend on individuals are forced to be equivalent for actors within positions. These variations of p_1 can be fit using programs that fit loglinear models by deleting from or including in the model statement the appropriate margins of the **y**-array (if fitting the model for individual actors) or the **w**-array (if fitting the blockmodel). For details, see Chapters 15 and 16 of Wasserman and Faust (1993).

Rather than constraining the parameters associated with individual actors to be equivalent for actors within positions, Wang and Wong (1987) proposed an extension of p_1 in which special blockmodel parameters are added to Equation (1). Their model retains the parameters for individual actors. The class of stochastic models proposed by Wang and Wong (1987) cannot be fit by standard statistical packages and requires a special algorithm.

Other extensions and generalizations of p_1 involve relaxing the restriction of $R = 1$ binary variable. Such generalizations of p_1 consist of those for multirelational data (Fienberg *et al.* 1985; Wasserman and Galaskiewicz 1984; Wasserman and Faust 1993, Ch. 15) and extensions to more than two response levels (Wasserman and Iacobucci 1986).

3.3 Goodness-of-fit measures for stochastic blockmodels

When a particular distribution is assumed for $p(\boldsymbol{x})$, such as p_1, the likelihood ratio statistic G^2 provides a natural solution to the problem of measuring the adequacy of a blockmodel's representation of the observed data. Let $\hat{y}_{ijkl}^{\mathscr{B}}$ be the predicted value of y_{ijkl} based on the stochastic blockmodel with the partition $\mathscr{B}$ of B positions. The likelihood ratio statistic G_B^2 is

$$G_B^2 = 2 \sum_{i<j} \sum_{k} \sum_{l} y_{ijkl} \log\left(y_{ijkl} / \hat{y}_{ijkl}^{\mathscr{B}} \right) \tag{3}$$

The degrees of freedom for G_B^2 equal the difference between the number of independent cells in $\mathbf{y}$ and the number of independent estimated parameters in $p(\boldsymbol{x})$. Determining the exact degrees of freedom is not simple (see Fienberg and Wasserman 1981; Haberman 1981; Wong and Yi 1989; Iacobucci and Wasserman 1990; Wasserman and Faust 1993).

When statistical packages such as SPSS, BMDP or SYSTAT are used to fit p_1 to individual actors (i.e., to $\mathbf{y}$-array), $G_B^2 = G_g^2 = G^2/2$, where G^2 is the likelihood ratio statistic given in the output. This adjustment is needed because each dyad is included in G^2 twice, rather than just once as in Equation (3) (Fienberg and Wasserman 1981). When a stochastic blockmodel is fit (i.e., a $\mathbf{w}$-array), the correction is more complex, but the goodness-of-fit index (3) is easy to compute given the data and fitted values.

When $B = g$, the likelihood ratio statistic $G_B^2 = G_g^2$ depends on the size of the table (i.e., on g, the number of actors). In this case, the exact distribution of G_g^2 is not known, but is should be close to a chi-squared distribution. However, when $B < g$, G_B^2 depends on B, the number of positions, and its asymptotic distribution is chi-squared. Caution is still warranted when comparing G_B^2 to the chi-squared

distribution, because for relatively large B, the **w**-array may still be large and relatively sparse, in which case, asymptotic theory does not apply. An advantage of the likelihood ratio statistic is that differences between them, say $\Delta G^2 = G^2_{B_1} - G^2_{B_2}$ where model $\mathscr{B}_1$ is a special case of $\mathscr{B}_2$, are conditional likelihood ratio statistics, and they are asymptotically distributed as chi-squared random variables.

The lack of fit of a stochastic blockmodel as measured by G^2_B is decomposable into two parts; namely,

$$G^2_B = 2\sum_{i<j}\sum_k\sum_l y_{ijkl}\log\left(\hat{y}^g_{ijkl}/\hat{y}^{\mathscr{B}}_{ijkl}\right) + 2\sum_{i<j}\sum_k\sum_l y_{ijkl}\log\left(y_{ijkl}/\hat{y}^g_{ijkl}\right)$$

$$= G^2_{(B,g)} + G^2_g \tag{4}$$

where $\hat{y}^g_{ijkl}$ are the fitted values from p_1. The quantity G^2_g reflects the lack of fit of p_1 to the observed relations among individual actors, and the quantity $G^2_{(B,g)}$ reflects the lack of fit due to the assignment of actors to positions. The latter quantity is particularly useful for assessing how closely actors adhere to the definition of stochastic equivalence.

In sum, G^2_B is a useful and natural index for assessing the goodness-of-fit of stochastic blockmodels, and is especially useful for examining the difference in fit between models. The likelihood ratio statistic has all of the desirable characteristics that Carrington, Heil and Berkowitz (1979) list for a goodness-of-fit index: it uses all of the information imposed by a given blocking without sacrificing parsimony, it is sensitive to the nature of the data, and it has a relatively high degree of known precision.

4. Mapping functions

The other major component of a blockmodel is the function ϕ that assigns actors to positions. A number of strategies exist for generating partitions of actors. Recall that actors assigned to the same position should be stochastically equivalent. After presenting specific strategies for generating potential mapping functions, the evaluation of these functions is discussed. The conditional likelihood ratio statistic $G^2_{(B,g)}$

is proposed as an index that measures how closely actors adhere to the definition of stochastic equivalence for a given partition.

4.1. Generating partitions

The assignment of actors to positions can be based on exogenous attribute information about the actors or on the relational data. Examples of exogenous characteristics are age, gender and income. This approach is straightforward and has been used by many (for examples, see Wasserman and Faust 1993). As an alternative to *a priori* classifications, Wasserman and Anderson (1987) explored ways of discovering *a posteriori* partitions based on the relational data, a characteristic of deterministic blockmodeling procedures. The discovery of partitions based on relational data is more difficult task than generating partitions based on attribute data. Techniques for identifying partitions *a posteriori* are reviewed here.

One possible strategy is to examine all possible partitions. Even with increases in computing power, such an approach is not practical, and certainly not efficient. For fixed B, the number of possible partitions for even moderately sized groups is extremely large. Furthermore, researchers will typically want to examine partitions for different values of B.

Another approach is to seek a spatial representation of the actors that reflects the relational ties between them. In such a representation, actors who are (approximately) stochastically equivalent should be close to each other, and those who are not equivalent should be far apart. When $p(\boldsymbol{x}) = p_1$, actors who are stochastically equivalent have the same α's and β's, and the task of finding equivalent actors reduces to that of finding subgroups of actors with (approximately) equivalent parameters. For the simple case of one binary relational variable, Wasserman and Anderson (1987) found that examining plots of $\hat{\beta}_i$ versus $\hat{\alpha}_i$ from fitting p_1 to the $\boldsymbol{y}$-array were extremely useful. A set of potential partitions for different numbers of positions can be suggested by visually examining such plots.

Among other possible graphical approaches is one explored by Wasserman and Anderson (1987), and discussed by Wasserman *et al.* (1990) and Wasserman and Faust (1989). These researchers plot row versus column scores from the correspondence analysis of a sociomatrix. Correspondence analysis is a technique that seeks to simultane-

ously scale the rows and columns of a table such that rows that are similar have similar scores, and columns that are similar have similar scores (Greenacre 1984). Other possibilities not considered were biplots (Gabriel 1982; Gabriel and Zamir 1979) and the *RC*-association model (Goodman 1985, 1986), both of which are related to correspondence analysis and also yield row and column scores. All of these methods can be applied to sociomatrices. The application, potential usefulness and limitations of these methods will not be explored further in this paper.

Another approach, which is complementary to graphical methods for discovering equivalent actors, is cluster analysis. Since stochastic equivalence among actors is operationally defined as equivalence of p_1 parameters, estimated parameters can be used in cluster analytic methods to try to find "optimal" partitions of actors. While cluster analysis can be used in addition to parameter plots, clustering techniques can also be used in more complex cases where the examination of parameter plots is difficult. For example, if there are more than two sets of parameters corresponding to individuals, as might be the case with multiple relations, the dimensionality of the parameter space equals the number of different sets of parameters. At most, three sets of parameters can be visually examined at any one time. Cluster analysis is not so limited. Advantages and the complementary nature of cluster analysis and the visual inspection of parameter plots are demonstrated in the example in Section 6.

Numerous cluster analytic methods exist that are potentially useful, but only two are mentioned here to illustrate how cluster analysis can be employed to help identify possible mapping functions for stochastic blockmodels. A promising method is Hartigan's (1975) *K-means* technique, which seeks to split objects into a fixed number of sets by maximizing the variation between sets relative to the variation within sets. This method requires an objects by variables matrix, which for our purposes corresponds to the actors by estimated parameters matrix. The parameters do not need to be re-scaled or standardized. For different numbers of positions, the *K-means* technique will not necessarily yield nested sets of partitions.

If a nested set of partitions is desired, then an hierarchical cluster analysis method could be used. These methods successively join together objects and subgroups until there is only one large cluster. The various methods differ with respect to the criterion used to join

individuals/subgroups at each stage. These methods operate on square symmetric matrices of (dis)similarities, so for our purposes, (dis)similarities between actors in the parameter space need to be computed. A logical choice for dissimilarities is the Euclidean distance between actor i and j,

$$\text{distance}(i, j) = \sqrt{\left(\hat{\alpha}_i - \hat{\alpha}_j\right)^2 + \left(\hat{\beta}_i - \hat{\beta}_j\right)^2}$$

which corresponds to the distances that are examined in parameter plots. Hierarchical methods may not be as useful as *K-means* cluster analysis, because the goal of a *K-means* analysis more closely resembles that of finding subgroups of actors with equivalent parameters.

4.2. Measures of stochastic equivalence

Regardless of whether a mapping function is based on exogenous characteristics of the actors or on the relational data, an index that measures the degree to which actors adhere to the definition of stochastic equivalence is needed. When partitions are based on exogenous information, an assessment is needed of whether actors within positions are actually (or approximately) stochastically equivalent. When partitions are based on the relational data, a means of identifying "optimal" or "good" mappings is needed. "Optimal" and "good" are defined in terms of stochastic equivalence.

As mentioned in Section 3.3, the conditional likelihood ratio statistic $G^2_{(B,g)}$ is a natural index to evaluate the degree to which actors within positions adhere to the definition of stochastic equivalence. As was seen from the decomposition in (4), $G^2_{(B,g)}$ reflects the lack of fit due to the assignment of actors to positions. Since $G^2_{(B,g)}$ is a difference between likelihood ratio statistics, it is an asymptotic chi-squared random variable and can be used to statistically test whether actors assigned to positions by a particular mapping function are consistent with the definition of stochastic equivalence. If actors can be assigned to blocks without "significantly" reducing the fit of the model, then actors and the relation(s) are consistent with the definition of stochastic equivalence.

The statistic $G^2_{(B,g)}$ can also be used to assess which of a number of different mapping functions for various numbers of blocks is the best

in terms of producing partitions of actors who more closely adhere to the definition of stochastic equivalence. Remember that $G^2_{(B,g)}$ reflects lack-of-fit. For fixed B, the mapping function that yields the smallest $G^2_{(B,g)}$ is the "best" one. For fixed B, the difference between the $G^2_{(B,g)}$'s from two different partitions is not a chi-squared random variable, because one model is not a special case of the other. However, the difference does indicate which partition is better and reflects the degree to which the actors within positions in one model are "more" stochastically equivalent than those in the other model.

For different numbers of blocks, mapping functions can be compared by computing the differences between $G^2_{(B,g)}$'s. When one partition is nested within the other, these differences are asymptotic chi-squared random variables with degrees of freedom equal to the difference between in the number of estimated blockmodel parameters. When one partition is not a special case of the other, the difference $G^2_{(B,g)}$'s can still be examined with respect to the difference in degrees of freedom. In this case, the difference between $G^2_{(B,g)}$'s is not distributed as a chi-squared random variable and cannot be used to test whether the difference is statistically significant.

5. Stochastic blockmodel representations

Given a probability distribution $p(\mathbf{x})$ and a mapping function $\phi(\cdot)$ for a stochastic blockmodel, the positions and relational ties between positions need to be represented. These representations are used to substantively interpret the model, which is an important but relatively neglected aspect of blockmodeling (Faust and Wasserman 1992). In deterministic blockmodel analyses, density tables, image matrices, and reduced graphs are three common ways in which the relations between positions are represented. Density tables and reduced graphs are useful in stochastic blockmodel analyses, but image matrices are irrelevant and not necessary.

Of substantive interest are the probabilities that actors relate to and are related to by other actors when actors are in the same or different positions. A density table (or matrix) contains these observed probabilities. Each row and column of the table corresponds to a

position. The observed probabilities equal

$$\Pr(x_{i,j}=1)=\begin{cases}(w_{st10}+w_{st11})/(g_s g_t) \text{ if } s\neq t\\ (w_{ss10}+w_{ss11})/[g_s(g_s-1)] \text{ if } s=t,\end{cases} \tag{5}$$

where $\phi(i)=s$ and $\phi(j)=t$. The counts g_s and g_t are the number of actors in positions $\mathscr{B}_s$ and $\mathscr{B}_t$, respectively, w_{st10} is the frequency of actors in position $\mathscr{B}_s$ who relate to but are not related to by the actors in position $\mathscr{B}_t$, and w_{st11} is the frequency of actors in $\mathscr{B}_s$ who relate to and are related to by those in $\mathscr{B}_t$. Whereas the diagonal entries of the sociomatrix $X_{ii}=0$, this is not the case for positions. Relational ties between a position and itself can exist, and $\Pr(x_{ij}=1)$ is the probability that actors in s relate to each other.

Rather than an observed density matrix, a matrix of expected or predicted probabilities can be computed based on the stochastic blockmodel. The predicted probabilities are computed by replacing the observed frequencies in Equation (5) by the predicted frequences from the stochastic blockmodel. The predicted frequencies, $\hat{w}_{st10}$ and $\hat{w}_{st11}$, are the fitted values computed from fitting the appropriate loglinear model to the **w**-array. Since the predicted probabilities for actors in the same position are equal, these predictions can also be computed as follows

$$\Pr(\hat{x}_{ij}=1)=\Pr(\hat{y}_{ij10}=1)+\Pr(\hat{y}_{ij11}=1) \tag{6}$$

where $i\neq j$; actors i and j are in positions $\mathscr{B}_s$ and $\mathscr{B}_t$, respectively; $\Pr(\hat{y}_{ijl1}=1)=\hat{w}_{st1l}/(g_s g_t)$ for $l=0$, 1 and $s\neq t$; and $\Pr(\hat{y}_{ijl1}=1)=\hat{w}_{ss1l}/(g_s(g_s-1)))$ for $l=0$, 1 and $s=t$. Alternatively, $\Pr(\hat{y}_{ijkl}=1)$ can be computed from Equation (1); that is,

$$\Pr(\hat{y}_{ijkl}=1)=\exp\left\{\hat{\lambda}_{ij}+k\hat{\alpha}_{[s]}+l\hat{\alpha}_{[t]}+k\hat{\beta}_{[t]}+l\hat{\beta}_{[s]}+(k+l)\hat{\theta}+kl\hat{\rho}\right\} \tag{7}$$

While the predicted and observed density matrices can be compared to see how closely the model is reproducing the observed probabilities, the predicted probabilities should be used in substantive interpretations of the model. The predicted density table contains the

stochastic blockmodel based probabilities of relational ties between actors in the same and different positions.

In deterministic blockmodels, image matrices are often used to represent the relational ties between positions. Similar to density tables, image matrices have rows and columns that correspond to positions and the entries carry information regarding relational ties between positions. The entries in an image matrix are ones and zeros and indicate whether or not a relational tie exists. In a stochastic blockmodel, relational ties between actors exist with certain probabilities, which can be anywhere in the range of zero to one; therefore, an image matrix is not useful for representing the relations between positions in a stochastic blockmodel analysis.

A third way of representing relational ties between the positions of a blockmodel is a reduced graph. Reduced graphs consist of nodes that correspond to positions, and arrows are drawn between nodes such that they point away from the sending position and towards the receiving position. In a deterministic blockmodel, arrows represent the existence of a relational tie. Reduced graphs for deterministic blockmodels are pictorial representations of image matrices. In a stochastic blockmodel, a reduced graph is based on the predicted density table. In this case, arrows are only drawn for the relational ties with large probabilities, and the predicted probabilities are written on or next to the arrows to convey the probabilistic information. Reduced graphs based on predicted density tables are pictorial summaries of the information in the corresponding density tables. The tables contain more information, but the reduced graphs provide a visual summary of the information in the tables.

6. Example: World systems data

The data analyzed in this example are from Wasserman and Faust (1993), and Faust and Wasserman (1991, 1992). The actors in this network are 24 countries that are geographically, economically and politically diverse. They represent a range of interesting features and span the categories of existing world system/development topologies. The relation examined here is whether a country exported basic manufactured goods in 1984 to other countries in the network. The data are given in Table 1 in the form of a sociomatrix. The (i, j)th

Table 1
Sociomatrix: Trade of basic manufactured goods

	Nation																							
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	Alg Algeria	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	
2	Arg Argentina	1	0	1	1	0	1	0	0	1	0	1	1	1	0	0	0	1	1	1	0	1	0	1	0
3	Bra Brazil	1	1	0	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
4	Chi China	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
5	Cze Czechoslovakia	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
6	Ecu Ecuador	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
7	Egy Egypt	0	0	0	0	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1	0	0	1	1	1
8	Eth Ethiopia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
9	Fin Finland	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1
10	Hon Honduras	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
11	Ind Indonesia	1	0	0	1	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	0	1	1	1	1
12	Isr Israel	0	1	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	1	1	0	1	1	1	1
13	Jap Japan	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
14	Lib Liberia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	Mad Madagascar	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
16	NZ New Zealand	1	0	0	1	0	0	1	0	0	0	1	0	1	0	0	0	1	1	0	0	1	1	1	1
17	Pak Pakistan	0	0	0	1	1	0	0	0	1	0	1	0	1	1	0	1	0	1	1	1	1	1	1	0
18	Spa Spain	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1
19	Swi Switzerland	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
20	Syr Syria	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	Tai Thailand	0	0	1	1	0	0	0	0	1	0	1	1	1	0	0	1	1	1	1	1	0	1	1	1
22	UK United Kingdom	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
23	US United States	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
24	Yug Yugoslavia	1	1	0	1	1	0	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1	0

entry equals one if country i exported basic manufactured goods to country j and equals zero otherwise. This matrix was generated based on data from the United Nations Commodity Trade Statistics (1984). For further details, see Faust and Wasserman (1991a).

The p_1 model for binary relations given in Equation (1) was fit to the data, using iterative proportional fitting as discussed by Fienberg and Wasserman (1981). We used several *FORTRAN* programs specially written for this purpose. The models and fit statistics are reported in Table 2. The first column lists the parameters included in the model, and the second column lists the margins (i.e., the loglinear model) of the **y**-array that were fit for each of the models. The first model, model (i), is the "full" p_1 model, and models (ii)–(iv) are special cases of it. The special cases were fit to see if a simpler model could be used to represent the relational ties among the countries.

Table 2
Fit statistics for p_1 and special cases: ΔG^2 equals the difference between G^2 associated with models (ii)–(iv) and model (i), and Δdf equals the difference between the associated degrees of freedom

	Model	Margins fit	G^2	ΔG^2	Δdf
(i)	$\theta, \{\alpha_i\}, \{\beta_j\}, \rho$	12 13 14 23 24 34	245.18		
(ii)	$\theta, \{\alpha_i\}, \{\beta_j\}$	12 13 14 23 24	252.56	7.38	1
(iii)	$\theta, \{\alpha_i\}, \rho$	12 13 24 34	298.35	53.17	23
(iv)	$\theta, \{\beta_j\}, \rho$	12 14 23 34	667.20	422.02	23

While the G^2's associated with models (i)–(iv) are not asymptotic chi-squared random variables. Since the ΔG^2's for models (ii)–(iv) are all large, the reciprocity parameter ρ (deleted in model (ii)), the set of "popularity" parameters $\{\beta_j\}$ (deleted in model (iii)), and the set of "expansiveness" parameters $\{\alpha_i\}$ (deleted in model (iv)) should all be included in the model. Therefore, all of the stochastic blockmodels fit will assume $p(\boldsymbol{x}) = \text{model (i)} = p_1$. The G^2 for model (i) corresponds to G_g^2, which was discussed in Section 3.3, and will be the lower bound for the fit of all the stochastic blockmodels (i.e., all blockmodels will have $G_B^2 \geq G_g^2$).

Assume that actors exhibit stochastic equivalence. The next step in constructing a stochastic blockmodel is to choose a mapping function. Since the methods described in Section 4 for generating mapping functions involve estimated model parameters, the maximum likelihood estimates of the parameters of p_1 were computed from the fitted values $\hat{y}_{ijkl}$ (as mentioned, these programs are available from the authors). The estimated θ equals -0.668 and the reciprocity parameter $\hat{\rho}$ equals 2.03. The latter indicates that trade between countries tends to be reciprocated.

The estimated α's and β's are plotted in Figure 1. The points represent countries (ignore the open circles for now). Since Syria and Liberia did not export manufactured goods to any other country in the network and the United States, Japan and Switzerland exported goods to all of the other countries, the $\hat{\alpha}_i$'s for these nations equal $-\infty$ and $+\infty$, respectively. To represent the countries with $\hat{\alpha}_i = \pm\infty$, these countries were placed at the extreme ends of the horizontal axis. Overall, the countries show more variation with respect to their exporting behavior ($\hat{\alpha}$) than they do with respect to their importing behavior ($\hat{\beta}$), even when disregarding the five nations with $\hat{\alpha}_i = \pm\infty$.

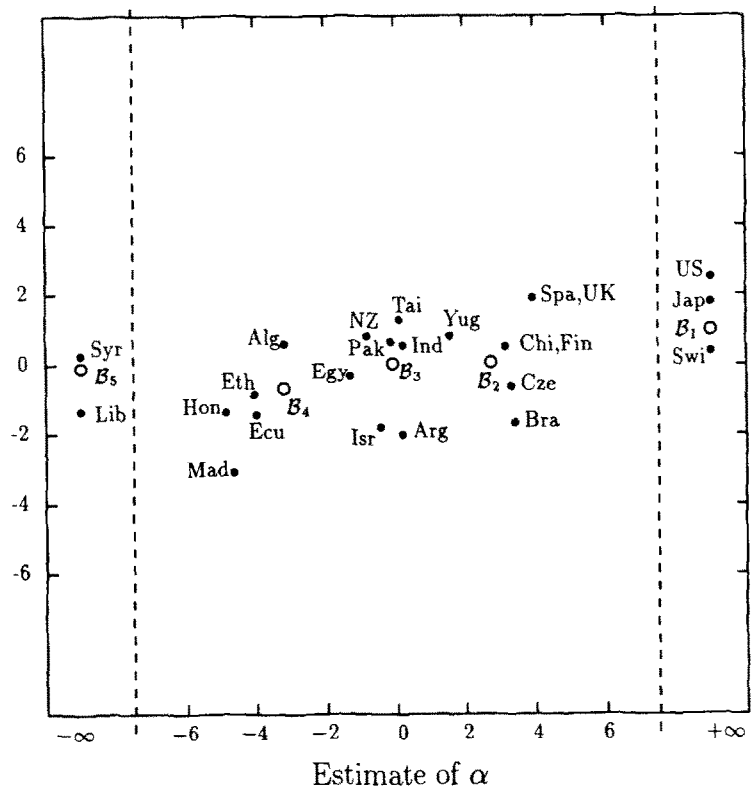

Fig. 1. Plot of $\hat{\alpha}_i$ versus $\hat{\beta}_i$. The points are p_1 parameter estimates $(\hat{\alpha}_i, \hat{\beta}_i)$, $i = 1, 2, \ldots, 24$, and the circles are stochastic blockmodel parameter estimates $(\hat{\alpha}_{[s]}, \hat{\beta}_{[s]})$, $s = 1, 2, \ldots, 5$.

Figure 1 greatly facilitates the search for equivalent countries. Nations that have similar import and export patterns (have similar $\hat{\alpha}$'s and $\hat{\beta}$'s. For example, Spain and the United Kingdom as well as China and Finland) have nearly identical rows and columns in Table 1, and their model parameters are approximately equal. The points corresponding to Spain and the UK, as well as those for China and Finland, are indistinguishable, which indicates that the nations in each of these two pairs are clearly stochasticaly equivalent. To close proximity of Ethiopia, Ecuador and Honduras suggests that these countries can be placed in the same position without significantly decreasing the goodness-of-fit of the stochastic blockmodel. Various possible mappings of countries to blocks for different values of B were

Table 3
Fit statistics for p_1 stochastic blockmodels

	B	$G^2_{(B,g)}$	$df_{(B,g)}$	$G^2_{(B,B-1)}$
Partitions from visual inspection				
{Jap, Swi, US} {Bra, Cze} {Chi, Fin, Yug} {Spa, UK} {Arg, Isr} {Alg} {Egy, Ind, NZ, Pak, Tai} {Ecu, Eth, Hon} {Mad} {Lib, Syr}	10	19.02	28	
{Jap, Swi, US} {Bra, Cze} {Chi, Fin, Yug} {Spa, UK} {Arg, Isr} {Alg} {Egy, Ind, NZ, Pak, Tai} {Ecu, Eth, Hon, Mad} {Lib, Syr}	9	23.45	30	4.43
{Jap, Swi, US} {Bra, Cze} {Chi, Fin, Yug} {Spa, UK} {Arg, Isr} {Alg, Egy, Ind, NZ, Pak, Tai} {Ecu, Eth, Hon, Mad} {Lib, Syr}	8	39.89	32	16.44
{Jap, Swi, US} {Bra, Cze} {Chi, Fin, Spa, UK} {Arg, Isr} {Alg, Egy, Ind, NZ, Pak, Tai, Yug} {Ecu, Eth, Hon, Mad} {Lib, Syr}	7	52.48	34	
{Jap, Swi, US} {Bra, Chi, Cze, Fin, Spa, UK} {Arg, Egy, Isr} {Alg, Ind, NZ, Pak, Tai, Yug} {Ecu, Eth, Hon, Mad} {Lib, Syr}	6	62.02	36	
Partitions from K-means cluster analysis				
{Jap, Swi, US} {Bra, Cze} {Chi, Fin, Spa, UK} {Arg, Isr} {Alg} {Egy, Ind, NZ, Pak, Tai} {Yug} {Ecu, Eth, Hon} {Mad} {Lib, Syr}	10	19.16	28	
{Jap, Swi, US} {Bra, Cze} {Chi, Fin, Spa, UK} {Arg, Isr} {Alg} {Egy, Ind, NZ, Pak, Tai} {Yug} {Ecu, Eth, Hon, Mad} {Lib, Syr}	9	23.68	30	4.52
{Jap, Swi, US} {Bra, Cze} {Chi, Fin, Spa, UK} {Arg, Isr} {Alg} {Egy, Ind, NZ, Pak, Tai, Yug} {Ecu, Eth, Hon, Mad} {Lib, Syr}	8	32.39	32	8.71
{Jap, Swi, US} {Bra, Cze} {Chi, Fin, Spa, UK} {Alg} {Lib, Syr} {Arg, Isr, Egy, Ind, NZ, Pak, Tai, Yug} {Ecu, Eth, Hon, Mad}	7	44.65	34	12.26
{Jap, Swi, US} {Bra, Cze} {Chi, Fin, Spa, UK} {Lib, Syr} {Arg, Isr, Egy, NZ, Pak, Tai, Yug} {Alg, Ecu, Eth, Hon, Mad}	6	53.68	36	8.03
{Jap, Swi, US} {Bra, Cze, Chi, Fin, Spa, UK} {Lib, Syr} {Arg, Isr, Egy, Ind, NZ, Pak, Tai, Yug} {Alg, Ecu, Eth, Hon, Mad}	5	64.09	38	10.41
{Jap, Swi, US} {Alg, Ecu, Eth, Hon, Mad} {Lib, Syr} {Bra, Cze, Chi, Fin, Spa, UK, Arg, Isr, Egy, Ind, NZ, Pak, Tai, Yug}	4	135.68	40	71.59
{Jap, Swi, US} {Alg, Ecu, Eth, Hon, Mad, Lib, Syr} {Bra, Cze, Chi, Fin, Spa, UK, Arg, Isr, Egy, Ind, NZ, Pak, Tai, Yug}	3	143.88	42	8.20
{Bra, Cze, Chi, Fin, Spa, UK, Arg, Isr, Egy, Ind, NZ, Pak, Tai, Yug, Jap, Swi, US} {Alg, Ecu, Eth, Hon, Mad, Lib, Syr}	2	191.35	44	47.47

generated by visually examining Figure 1. Some of these mappings are reported in the top half of Table 3.

Other mappings for 2–10 positions were generated by performing *K-means* cluster analyses of the countries using $\hat{\alpha}$ and $\hat{\beta}$. (A large number, 9, was substituted in for ∞.) The clusters, which are listed in the lower half of Table 3, are nested. The cluster analyses confirmed

many of the aspects seen in Figure 1. The partitions for $B = 5$, 6, and 8 from the cluster analyses were also identified as possible partitions from the visual examination of Figure 1. The duplicate partitions are reported only once in Table 3, under the *K-means* section. The major differences between the partitions generated from the figure and those from the cluster analyses involve Yugoslavia. Based on Figure 1, Yugoslavia was generally assigned to the same position as Spain and UK, but in the cluster analyses, it was assigned to the cluster containing Indonesia, New Zealand, Pakistan, and Thailand.

Fit statistics for the various stochastic blockmodels are also given in Table 3. The first column shows the actual mapping of actors onto positions and the second indicates the number of positions. The third and fourth columns contain the conditional likelihood ratio statistics $G^2_{(B,g)}$ and their degrees of freedom $\mathrm{df}_{(B,g)}$, respectively. These quantities are used to assess the degree to which actors within positions adhere to the definition of stochastic equivalence. The last column contains the conditional likelihood ratio statistics $G^2_{(B,B-1)}$ for nested models in which the more restrictive model has one less position. There are two degrees of freedom associated with each $G^2_{(B,B-1)}$.

When the number of positions is fixed at either 8, 9, or 10, the models in the upper and lower halves of Table 3 have approximately the same fit statistics; however, when $B = 6$ or 7, the models in the lower half fit noticeably better than those in the upper half. The *K-means* cluster analyses produced partitions at least as good as those generated from Figure 1. Since the partitions in the lower half of Table 3 tend to have better fit statistics, are all nested, and cover a larger range of models for different numbers of blocks, the models in the top half were eliminated from further consideration.

When the stastistic $G^2_{(B,g)}$ for different numbers of positions are compared with the appropriate chi-squared distributions, the statistics for $B \geqslant 7$ are not statistically "large" (p-values > 0.10). The statistic $G^2_{(6,g)} = 0.029$ is marginally "large' (p-value $= 0.029$), and the statistic $G^2_{(5,g)} = 0.005$ is statistically "large" (p-value $= 0.005$). These fit statistics suggest that the 7 and possibly the 6 position blockmodels are the simplest ones that provide an adequate fit. Since the applicability of asymptotic theory in this example is questionable, other criteria must also be considered.

The fit statistics $G^2_{(B,B-1)}$ indicate the decrease in fit from reducing the number of positions from B to $(B - 1)$ where two positions from

the more general model are combined into one position in the more restrictive model. For models with 5–9 positions, the values for these statistics are relatively constant and range form 4.52 to 12.26. A large decrease in the fit occurs at $B = 4$ where $G^2_{(4,5)} = 71.59$. Given this fact, models with $B \leq 4$ were eliminated from further consideration.

Since the 7 position model contains a position with just one country (i.e., Algeria) and the 6 position model provides a reasonably good fit to the data, the 7 position model was also eliminated. The 5 and 6 position blockmodels differ in that Brazil and Czechoslovakia form a separate position in the 6 position model, but they are included in the cluster with China, Finland, Spain, and the United Kingdom in the 5 position blockmodel. The representations of each of these models was examined. The 5 position model was chosen, because the basic substantive interpretation is the same as the 6 position model, except for one minor difference, which will be noted later. Based on a balance of parsimony and goodness-of-fit, our favorite solution is the 5 position blockmodel from the *K-means* cluster analysis. A substantive interpretation of this model follows.

The nations were mapped onto positions as follows:

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

The estimated values for the overall choice effect and the reciprocity parameter are -0.803 and 2.133, respectively, which are similar to those from p_1. The estimated values for $\alpha_{[s]}$ and $\beta_{[s]}$ correspond to the open circles labeled $\mathscr{B}_1$–$\mathscr{B}_5$ in Figure 1. The positions differ mostly with respect to exports ($\hat{\alpha}_{[s]}$), but show some slight differences with respect to imports ($\hat{\beta}_{[s]}$). To explictly represent and substantively interpret the relations between the positions, the predicted density matrix was computed and a reduced graph based on this matrix was drawn.

The predicted probabilities are given in Table 4. The countries in $\mathscr{B}_1$ exported goods to all of the other countries (i.e., the entries in the first row of Table 4 all equal 1.00), and the countries in $\mathscr{B}_5$ did not export any goods to any of the other countries (i.e., the entries in the

Table 4
Predicted density matrix: Entries equal the predicted probability that a country in a row (column) position exports (imports) manufactured goods to (from) a country in a column (row) position

	$\mathscr{B}_1$	$\mathscr{B}_2$	$\mathscr{B}_3$	$\mathscr{B}_4$	$\mathscr{B}_5$
$\mathscr{B}_1$	1.000	1.000	1.000	1.000	1.000
$\mathscr{B}_2$	0.994	0.983	0.956	0.804	0.868
$\mathscr{B}_3$	0.904	0.770	0.576	0.192	0.276
$\mathscr{B}_4$	0.295	0.119	0.041	0.010	0.017
$\mathscr{B}_5$	0.000	0.000	0.000	0.000	0.000

last row all equal 0.00). The relational ties exhibit a "center-periphery" pattern; that is, the larger probabilities are in the upper left triangle, while the smaller probabilities are in the lower right triangle. Countries in the positions $\mathscr{B}_1$, $\mathscr{B}_2$, $\mathscr{B}_3$ have large probabilities of exporting and importing goods from each other. The nations in positions $\mathscr{B}_1$ and $\mathscr{B}_2$ export goods to countries in $\mathscr{B}_4$ and $\mathscr{B}_5$ with large probabilities, but the nations in $\mathscr{B}_3$ export to $\mathscr{B}_4$ and $\mathscr{B}_5$ with small probabilities.

As noted earlier, the predicted density matrices for both the 5 and 6 position blockmodels were examined. The basic difference between the 5 and 6 position blockmodels was that in the $B = 6$ model, the predicted probability that countries in the cluster {China, Finland, Spain, United Kingdom} imported goods from $\mathscr{B}_3$ was 0.88, while the same probability for the countries in the cluster {Brazil, Czechoslovakia} was only 0.51. In the 5 position model, the corresponding

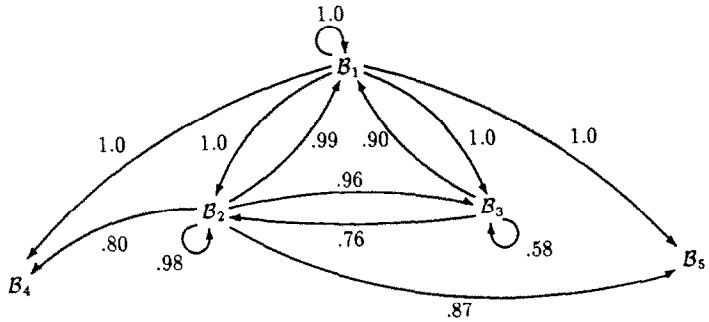

Fig. 2. Reduced graph based on predicted probabilities > 0.30.

predicted probability is 0.77, which is intermediate between these two values.

Figure 2 is a the reduced graph based on Table 4. It is pictorial representation of the probabilities that goods are exported/imported between countries in the five positions. The nodes (positions) are labeled $\mathscr{B}_1$–$\mathscr{B}_5$, and arrows are draw from one position to another positions for probabilities greater than 0.30. The central-periphrey pattern is well illustrated in this figure. The positions $\mathscr{B}_1$ and $\mathscr{B}_2$ export to countries in all of the other positions, but differ with respect to probabilities. Nations in $\mathscr{B}_4$ and $\mathscr{B}_5$ appear quite similar with respect to importing, but referring to Table 4, we see that the nations in $\mathscr{B}_4$ export goods to countries in other positions with small probabilities, while those in $\mathscr{B}_5$ do not export to any of the other countries.

7. Summary

Stochastic blockmodels consist of a probability distribution for data and a function that maps stochastically equivalent actors onto positions. In this paper, the specific model assumed for $p(\boldsymbol{x})$ was the model p_1. Two complementary techniques were described for generating mapping functions: examining plots of the estimated parameters α_i and β_j, and cluster analyses of these parameters. Both of these techniques depend on the data and the stochastic blockmodel. In a detailed example, a stochastic blockmodel analysis was performed on a network consisting of 24 countries and the relation of whether countries exported/imported basic manufactured goods from each other.

Faust and Wasserman (1992) state four basic tasks that must be performed in a complete positional analysis. These are

1. define "equivalence" among actors,
2. measure how closely the actors adhere to this definition,
3. represent the equivalences of the actors, and
4. measure the adequacy of this representation.

Each of these component tasks were fulfilled in the stochastic blockmodeling approach presented in this paper. Equivalence was defined as *stochastic equivalence*; that is, actors are equivalent if they have the

same probability distributions. The likelihood ratio statistic G_B^2 was shown to be a natural index for measuring the fit of the blockmodel to the data, and the conditional likelihood ratio statistic $G_{(B,g)}^2$ was shown to be an index for measuring how closely the actors adhere to the definition of stochastic equivalence. Predicted density tables and reduced graphs were used to represent the relational ties between actors within positions and interpret the results of stochastic blockmodel analyses.

References

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

Breiger, R.L.
1981 "Comment on Holland and Leinhardt, 'an exponential family of probability distributions for directed graphs'". *Journal of the American Statistical Association 76*: 51–53.

Faust, K. and S. Wasserman
1991 "Centrality and prestige: A review and synthesis". Unpublished manuscript.
1992 Blockmodels: Interpretation and evaluation. *Social Networks, 14*: 5–61.

Fienberg, S.E. and S. Wasserman
1981 "Categorical data analysis of single sociometric relations", in: S. Leinhardt (ed.) *Sociological Methodology 1981*, pp. 156–192. San Francisco: Jossey-Bass.

Fienberg, S.E., M.M. Meyer and S. Wasserman
1985 "Statistical analysis of multiple sociometric relations". *American Statistical Association 80*: 51–67.

Gabriel, K.R.
1982 "Biplot", in: S. Kotz, N.L. Johnson and C.B. Reed (eds.) *Encyclopedia of Statistical Sciences*, Vol. 1, pp. 263–271.

Gabriel, K.R. and S. Zamir
1979 "Lower rank approximation of matrices by least squares with any choice of weight". *Technometrics 21*: 489–498.

Goodman, L.A.
1985 "The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries". *The Annals of Statistics 13*: 10–69.
1986 "Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables". *International Statistical Review 54*: 243–309.

Greenacre, M.J.
1984 *Theory and Application of Correspondence Analysis*. Orlando, Florida: Academic Press.

Haberman, S.J.
1981 "Comment on Holland and Leinhardt, 'An exponential family of probability distributions for directed graphs'". *Journal of the American Statistical Association 76*: 60–62.

Hartigan, J.A.
1976 *Clustering Algorithms*. 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. and S. Leinhardt
1981 "An exponential family of probability distributions for directed graphs". *Journal of the American Statistical Association 76*: 33–65.
Holland, P.W., K.B. Laskey and S. Leinhardt
1983 "Stochastic blockmodels: Some first steps". *Social Networks 5*: 109–137.
Iacobucci, D. and S. Wasserman
1990 "Social networks with two sets of actors". *Psychometrika 55*: 707–720.
United Nations
1984 *Statistical Papers: Commodity Trade Statistics*. 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. and C. Anderson
1987 "Stochastic *a posteriori* blockmodels: Construction and assessment". *Social Networks 9*: 1–36.
Wasserman, S. and K. Faust
1989 "Canonical analysis of composition and structure of social networks", in: C.C. Clogg (ed.) *Sociological Methodology 1989*, pp. 1–42. Cambridge, MA: Blackwell.
Wasserman, S. and K. Faust
1993 *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.
Wasserman, S. and J. Galaskiewicz
1984 "Some generalizations of p_1: External constraints, interactions, and non-binary relations". *Social Networks 6*: 177–192.
Wasserman, S. and D. Iacobucci
1986 "Statistical analysis of discrete relational data". *British Journal of Mathematical and Statistical Psychology 39*: 41–64.
Wasserman, S., K. Faust and J. Galaskiewicz
1990 "Correspondence and canonical analysis of relational data". *Journal of Mathematical Sociology 15*: 11–62.
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.
Wong, G.Y. and Q.Q. Yi
1989 "Computation and asymptotic normality of maximum likelihood estimates of exponential parameters of the p_1 model". Unpublished manuscript.