Journal of Mathematical Sociology, 1993, Vol. 18(1), pp. 1–26
Reprints available directly from the publisher
Photocopying permitted by license only
© 1993 Gordon and Breach Science Publishers S.A.
Printed in the United States of America

NONPARAMETRIC INFERENCE FOR NETWORK DATA*

KATHLEEN CARLEY

Department of Social and Decision Sciences
Carnegie Mellon University

DAVID BANKS

Department of Statistics
Carnegie Mellon University

May 6, 1992; revised December 28, 1992

Various theoretical concerns often require researchers to answer questions of the form does co-membership in a group predict other ties between those individuals. Data appropriate for answering such a question often is referred to as network data. Network data exhibits row-column dependencies that often invalidate traditional statistical methods for doing group comparisons such as analysis of variance and the Kruskal–Wallis Procedure. Because of the positive dependence within rows/columns the significance probabilities of such traditional methods may be exaggerated. This paper uses restricted-randomization to develop exact permutation tests for network data where co-membership in groups can be specified a priori. This enables the nonparametric estimation of the significance of standard statistics for group-difference tests and ordered-alternative tests where the group orderings have been prespecified. These methods are demonstrated by examining three different data sets: Sampson's Monastery data, Carley's Tutor Selection data, and Humana's Human Rights data.

KEY WORDS: Anova; nonparametric tests; permutation analysis; dyads; co-membership; network data; Kruskal-Wallis; Jonckheere-Terpstra.

Sociologists often are concerned with whether similar behavior depends on similar position in the social structure, such as occurs when two individuals are members of the same group. For example, sociologists have debated whether co-membership in some group has predictable influence on such phenomena as the production of consensus (Blau, 1977; Burt, 1982; Carley, 1986), conflict (Homans, 1950), the rate of information diffusion (Burt, 1980; Carley, 1990; Fararo and Skvoretz, 1986), and the type of information information that diffuses (Granovetter, 1973; Granovetter, 1974). Being concerned with the patterns of relationships among individuals sociol-

*This research was supported, in part, by the NSF under grant No. IST-8607303.

ogists often ask whether having a particular relationship in common predicts other relationships between individuals. For example, for all mid-career individuals, are co-workers more likely to marry than other couples? Or, for all countries, do countries with similar economic levels show similar human rights behavior?

These questions are all concerned with whether the dyad-groups[1] (hereafter, D-groups) are different, hence we refer to them as group-difference questions. Examples of D-groups are co-worker couples/non co-worker couples, pairs of students at the same college/pairs of students at different college, and pairs of countries that are economically similar/pairs of countries that are economically dissimilar. Sometimes the social scientist wants to know not just whether the D-groups are different, but whether a prespecified ordering of those D-groups can predict the ordering of a different relationship. For example, given data on faith and the amount of time spent interacting, is it the case that pairs of individuals of mixed faith interact the least, pairs where both people are Christian interact the next most, and pairs where both are Bhuddist interact the most? Or, are pairs of poor countries more heterogeneous with respect to human rights than are pairs of moderately well off countries which are in turn more heterogeneous than pairs of moderately well off countries which are in turn more heterogeneous than pairs of wealthy countries? These questions are all of the form on relationship y is D-group $a <$ D-group $b <$ D-group c. In each question an ordering of the D-groups are specified, hence we refer to these questions as ordered-alternative questions.

Both group-difference questions and ordered-alternative questions are addressed using network data—where the unit of analysis often is the pair of actors. In fact, these questions require two networks. In the first network, the independent variable, the relation between pairs of actors indicates to which D-group both actors belong. In the second network, the dependent variable, the relation between pairs of actors indicates some linkage that is expected to vary with D-group membership. Without this network structure to the data, as when the unit of analysis is the actor, traditional methods (such as analysis of variance and the Kruskal-Wallis procedure) offer easy tests for group-difference questions. Also, without this network structure when one has an ordered-alternative question traditional methods such as the Jonckheere-Terpstra test (Hollander and Wolfe, 1973) can be used. For network data, however, the observations that pertain to a given actor are not independent. For example, imaging that for each pair of students you know how many of the same classes they have taken. If one of the students has taken four times as many classes as the aver-

[1]Throughout this paper we use the term dyad-group or D-group to refer to a set of dyads such that both members in the dyad share some feature (have a particular type of relationship), such as each dyad is composed of actors who are both Catholic (Catholicism relationship), or each dyad is composed of one actor who is Catholic and one who is Jewish (inter-faith relationship). Because D-groups classify dyads and not individuals the same actor may be in multiple D-groups. In contrast, we use the term actor-group or A-group to refer to a set of actors such that all actors have some feature, such as the actor is Catholic or the actor is Jewish. These A-groups are equivalent to the subgroups identified by Iacobucci and Wasserman (1987). In both cases, we use the term group, rather than subgroup, both for ease of exposition and to clarify that we are not discussing the *subgroup problem*. In clinical trials, the subgroup problem occurs when, for example, having given a drug to a group of individuals the researcher tries to determine whether there exists a specific subgroup that is affected more than others by the drug. The difference between the problem we address and the subgroup problem is that in our case the groups (or subgroups) are predefined whereas in the subgroup problem they are not.

age student then it is likely that the number of classes this student has in common with each other student is much higher than for other pairs of students. This dependency can invalidate the traditional procedures for calculating the significance probability of these measures.

This paper (Section 2) develops a nonparametric analysis that takes proper account of the dependency structure, and shows how the failure of independence can cause traditional statistical procedures for examining group-difference and ordered-alternative questions, when applied to network data to find spuriously significant results. Because the proposed tests are permutation tests they are conditionally distribution-free and therefore impose no requirement for normally distributed random variables. As a consequence, the proposed tests are usually more conservative (not always) than the inappropriate traditional tests. In Section 3, we illustrate the proposed methods by applying them to three data sets: Sampson's Monastery Data (Sampson, 1968), Carley's Tutor Selection Data (Carley, 1984; Carley, 1986), and Humana's Human Rights Data (Humana, 1983). In Section 3, for group-difference questions, we will contrast the traditional statistical approach for calculating the significance probability for the F and Kruskal-Wallace measures of separation with our proposed permutation approach for calculating this probability. For ordered-alternative questions, we will contrast the traditional statistical approach for calculating the significance probability for the Jonckheere-Terpstra measure with our proposed permutation approach for calculating this probability. In Section 4 a discussion of these and related methods is provided. In Section 5 we draw some conclusions.

Our goal is to determine whether a set of D-groups shows significant differences in some dependent network-variable. Consequently, the tests we propose are particularly useful in studying the impact of social structure defined via a set of D-groups. Prior to using our tests the researcher must define a set of D-groups. There are many ways of defining a set of D-groups. For example, a set of D-groups might consist of couples that are both male, both female, or opposite sex; alternatively, a set might consist of couples that are same sex and couples that are opposite sex. Thus the D-groups can consist of pairs of actors who are similar in some relationship such as plays with (Homans, 1950), like (Homans, 1950; Newcomb, 1961), or respects (Sampson, 1968), or in terms of some pattern of relationships (Breiger, Boorman, and Arabie, 1975; Heil, 1976; Burt, 1989), or set of relationships (White, 1976; Lorrain and white, 1971). For our purpose, it is not important how the D-groups are defined, merely that the pairs of actors in the network can be divided into a set of D-groups, and that the data used to define D-groups is different from the dependent variables whose relationship we are trying to study.

METHODS

Permutation tests often enable nonparametric inference for complex problems. As a simple example of the strategy, consider the k-sample location problem. This examines a set of k random samples to decide whether these populations have a common center (or location), or whether there exists evidence that some populations have distributions that are shifted above or below the others. One observes independent samples of size n_i, $i = 1,\ldots,k$. Each of these populations has a center (which can be measured as its mean or median) which we denote as Δ_i. The samples are independent of each other, and each sample consists of independent and identically distributed values having cumulative distribution function $F(x - \Delta_i)$. To make a nonparametric test of the null hypothesis that there exists a common center; i.e., $H_0 : \Delta_1 = \cdots = \Delta_k$, the permutation principle uses the fact that, under the null hypothesis, all possible samples generated by any rearrangement of data values that assigns observations to other populations, while holding the sample size per population fixed, are equally likely. Any such possible sample is called a pseudo-dataset. Since there exists a total of $n = \sum n_i$ observations, the probability under the null of observing any particular set of samples (conditional on the data) is just

$$\binom{n}{n_1,\ldots,n_k}^{-1} = \frac{1}{n!}\prod_{i=1}^{k} n_i!.$$

Thus the significance probability of the test is

$$d\binom{n}{n_1,\ldots,n_k}^{-1},$$

where d is the number of pseudo-datasets that support the alternative hypothesis as or more strongly than the observed dataset, where the strength of support is assessed by the magnitude of some suitably chosen test statistic, such as a standard normal theory F-test value from the analysis of variance. However, because of the equiprobability of the pseudo-datasets under the null hypothesis, this result does not depend on the form of $F(\cdot)$, and thus the procedure is nonparametric.

Network data offer a more complicated version of the k-sample location problem, and the simple permutation test does not apply. Consider the following application: a network researcher records for n_1 male students and n_2 female students the number of minutes each pair of students spends interacting during a typical schoolday. This interaction time will be our dependent variable. The research hypothesis is that there are group differences in the length of interaction among male-male pairs, female-female pairs, and mixed-sex pairs. We record our dependent variable in a matrix with deleted diagonal (see the right hand matrix in Figure 1). We note that if the data were counts obtained from cross-classifications of many observations with respect to two or more factors, instead of being non-count numerical measurements pertaining to a pair of individuals, then one would call the diagonal entries structural zeroes. The entry in row i and column j of this matrix is the number of minutes student i is observed as interacting with student j during a typical schoolday. We also record our independent or D-group variable in a matrix with deleted diagonal (see the left hand matrix in Figure 1). The entry in row i and column j of this matrix is a number identifying which D-group the pair of students is in: (1) male-male, (2) female-female, and (3) mixed-sex. For each of these three D-groups we can calculate Δ_i as the median level of interaction for all pairs of students in that D-group. We can now state the null hypothesis as $H_0 : \Delta_1 = \Delta_2 = \Delta_3$. In this example the data matrix for the dependent variable is symmetric, but the strategy we will describe extends to asymmetric matrices as well. Asymmetric matrices allow

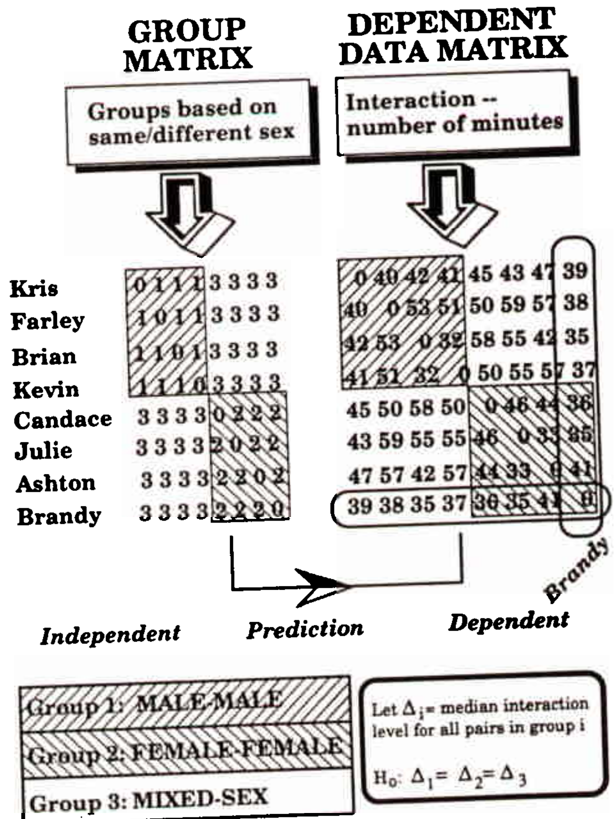


FIGURE 1. Example network data for nonparametric testing.

the possibility of distinguishing such things as female-male pairs from male-female pairs, which is pertinent in some applications.

Traditional location tests (e.g., ANOVA, the Kruskal-Wallis test or the k-sample permutation test) cannot correctly analyze these data; they all assume that observations are independent. Here the data within a given row or column are dependent, since the values pertain (in part) to a common individual. Figure 1 shows this dependence with Brandy; entries in her row and column are remarkably small. Many researchers attempt to model this dependency parametrically, but this forces one to build a fairly complicated model and make possibly unwarranted assumptions about the underlying distributions. Often it is preferable to have a simple nonparametric procedure; however, any conventional nonparametric analysis that assumes all rearrangements of the sample data are equally likely under the null hypothesis necessarily ignores the dependencies within rows and columns of the dependent-variable matrix that characterize network data.

Nonetheless, the permutation principle can apply to network data if the appropriate adjustments are made. In our example, the A-groups are male and female. Whereas, D-groups are groups of dyads defined by the reearcher that separate pairs of actors (i and j) into categories such that actor i is in A-group a and actor j

is in A-group b and a may be equal to b. In our example, we defined three D-groups—male-male, female-female, and mixed-sex. There is no unique link between the number of A-groups and the number of D-groups that applies to all research situations. For example, one researcher might, given the A-groups in our example, look at the three D-groups we defined; another might combine the male-male and female-female pairs into a single D-group of same-sex pairs.) We draw a distinction between A-groups and D-groups since the hypotheses are concerned with what difference in some dependent network can be attributed to differences in pairs' memberships in D-groups, whereas the permutation test depends on rearrangements of the data relative to the A-groups.

To apply the permutation principle to network data we use the fact that under the null hypothesis of no D-group differences, the pseudo-datasets generated by rearrangements of the row and column labels are equiprobable (Hubert, 1985). This paper exploits the equiprobability or row-column rearrangement under the null to estimate the nonparametric significance probability of two types of hypothesis tests—one for group-difference questions and one for ordered-alternative questions. Both tests share the same null hypothesis; i.e., $H_0 : \Delta_1 = \cdots = \Delta_r$, which asserts that all r D-group populations have the same distribution $F(\cdot)$ and the same location parameter. The first test, for group-difference questions, is analogous to the k-sample location problem, with omnibus alternative alternative $H_A : \exists i, j$ such that $\Delta_i \neq \Delta_j$. The second test, for ordered-alternative questions, has the alternative hypothesis that $H_A : \Delta_{\pi_1} \leq \cdots \leq \Delta_{\pi_r}$. Here, $\pi_1, \ldots, \pi_r$ is a prespecified ordering of the r D-groups and at least one of the inequalities is strict.

This second test often is more practical, since it enables the researcher to make an inference on the significance of a particular arrangement of the D-group centers. In passing, we note that the ideas developed in the following subsections are quite generalizable and enable the extension of other sorts of traditional statistical tests to nonparametric inference on network data. For example, one could compare all D-groups to a control D-group, by using Dunnet's test (1964) to measure the degree of conformity with the hypothesis that all populations are centered at the center of the control population. Also, one could extend Friedman's test (1937) to examine a two-way layout structure on the D-groups (in our simple example, one factor in the design might be D-groups defined by sex, and the second factor might be D-groups defined by race). Finally, if one is interested in a specific set of contrasts among the D-group centers, one can use Scheffe's projection procedure (1959) to develop a test statistic for a specific dataset, and then use the fact that under the null, all rearrangements of the A-group labels are equiprobable to generate a permutation test.

Group-Difference Tests

To implement a permutation test for network data, we enumerate all possible rearrangements of the A-group labels, and find the proportion of rearrangements that yield pseudo-data as or more supportive of the alternative than the actual data in the dependent variable matrix. The strength of support can be measured using any of many statistics that reflect group separation. In this paper, we will use the F

statistic from an analysis of variance and the Kruskal-Wallis statistic. Row-column dependency invalidates comparison of these statistics with customary tables; instead, we use the values obtained from all possible rearrangements of the A-group labels to develop the null distribution, and thereby obtain the significance probability.

Each rearrangement of the A-group labels induces a randomization of the data in the D-groups. Since entries in common rows and columns are relocated together, these randomizations preserve the row-column dependency structure. We call these rearrangements of the dependent variable data matrix restricted-randomizations, to distinguish them from the unconstrained rearrangement of this matrix allowed by conventional permutation tests. The permutation principle ensures that these restricted-randomizations are equiprobable under the null hypothesis that all D-groups have common centers, and we denote this common probability by p. Then the exact significance probability is $s = dp$, where d is the number of restricted randomizations that give pseudo-datasets that are as, or more, supportive of the alternative than the data actually observed. This follows the same logic as for the standard permutation test. For example, in Figure 1, the number of possible rearrangements of the dependent variable data matrix that keep elements together within rows and columns is $p^{-1} = \binom{8}{4} = 70$. Without this restriction preserving row/column dependence, the number of rearrangments is 28!; this reflects the fact that there are 28 entries in the upper triangular matrix and any arrangement of these entries can occur. If we had an experiment in which the matrix could be asymmetric, then the number of unrestricted rearrangments increases to $(2 \times 28)!$.

In practice, the combinatorial explosion prevents actual enumeration of all possible rearrangements of the A-group labels. Therefore we use a Monte Carlo procedure to equiprobably sample the set of restricted randomizations. Our programs generate 10000 such samples and calculate a measure of group separation for each.[2] The estimate of s is $d^*/10000$ where d^* is the number of sampled pseudo-data matrices that show group separation at least as great as the original data. The standard error of this estimate is $\sqrt{s(1-s)/10000}$; it is largest when $s = .5$, and decreases monotonically with s. There is no need for accurate estimation of s in the vicinity of .5; these values offer no support for the alternative hypothesis. From a practical standpoint, one wants accuracy near $s = .1$; here the normal approximation to the binomial gives the width of a 90% confidence interval on the significance probability as $2(1.645)\sqrt{.1(1-.1)/10000} = 0.00981$, which seems sufficiently accurate.

There are many sensible measures that capture the separation among the centers of the k populations. Our programs calculate two: the traditional F statistic, as used in the simple one-way ANOVA, and the Kruskal-Wallis statistic, as used in one of the nonparametric analogues to ANOVA. Generally, the F statistic is most powerful when the data derive from a normal or light-tailed distribution. Whereas, the Kruskal-Wallis statistic generally performs well with heavy-tailed distributions. Arbitrarily many alternative measures could be devised. We use these two statistics as they are known to bracket a wide range of behavior. Using our permutation procedure we found that in general both measures are in close agreement on the

[2] These programs are written in Fortran, require IMSL, and have been run on a workstation under the VAX-VMS operating system. Programs are available upon request from the authors.

significance probability. This finding emphasizes the tendency for all variants of the permutation test to agree.

Consider the example in Figure 1. In this case there are 3 D-groups and 28 elements in the upper triangular matrix for collaboration time. The traditional *F* statistic for a one-way ANOVA is 3.084 with 2 and 25 degrees of freedom and a 0.064 significance probability. The traditional Kruskal-Wallis test provides a value of 5.024 with a significance probability of 0.077. In contrast, the restricted-randomization permutation procedure we have described gives the significance probabilities of 0.148 and 0.15 for the *F* and Kruskal-Wallis measures respectively.[3] This simple example illustrates that traditional analyses are not conservative and can exaggerate the conclusions. In fact, in this example, the difference between the traditional and permutation approaches is quite large (approximately a factor of 2). Further, this example illustrates that the significance probability generated by the permutation test is relatively insensitive to the measure of group separation (here an *F* statistic and a Kruskal-Wallis statistic).

The calculated significance probability depends on both the assumptions of the model and the data. The nonparametric tests we have described make no assumptions other than that the rows (or columns) are independent. Other tests, which make additional assumptions, are not valid unless these additional assumptions hold. Their power, however, is an open question. Both the traditional F-test and the Kruskal-Wallis test make assumptions about the data's distribution that our permutation tests do not. Consequently we know that in contrast to their permutation versions, the traditional tests are not more powerful but can be less conservative and exaggerate the significance probabilities of the measures of separation. Further, depending on whether we use an *F* or a Kruskal-Wallis measure of group separation, even under the permutation procedure, there will be some difference in the significance probabilities. These differences reflect prior ideas about alternative senses of group differences as embodied in these measures, but the permutation technique does not enforce distributional assumptions required by the traditional method of analyses.

We note that network data need not yield a symmetric dependent variable data matrix, and the analysis should distinguish the asymmetric and the symmetric cases. When asymmetry occurs, all matrix entries except the diagonal are informative. But if the matrix necessarily is symmetric, then using the full matrix with deleted diagonal spuriously doubles the sample size, affecting the sensitivity of measures of group separation. to avoid this problem, our programs check for symmetry. If the data is symmetric, then the analysis depends only upon the superdiagonals. Otherwise, the full matrix is used.

Ordered-Alternative Tests

The permutation principle also enables one to perform ordered-alternative tests. Often there is more interest in detecting whether the D-group locations conform to a hypothetical ordering than in simply discovering that an unspecified subset of

[3] For permutation tests, degrees of freedom for significance probabilities are entirely meaningless, since readers are never referred to a distribution table, such as an F table.

D-groups shows separation. For non-network data, one can use a single degree of freedom test of a prespecified orthogonal contrast when the assumptions needed for ANOVA hold, or the Jonckheere-Terpstra test when the Kruskal-Wallis analysis is more appropriate. This subsection develops a nonparametric analogue of such tests for network data.

The two statistics used in this paper do not exhaust the possibilities. Robertson, Wright and Dykstra (1988, p. 206) point out that when classical assumptions hold and the population medians are equally spaced in accordance with the alternative hypothesis, then the Jonckheere-Terpstra test has very good power; however, when only one inequality is strict and the ordering accords with the alternative hypothesis, then the Chacko-Shovack statistic shows greater power than the Jonckheere-Terpstra Test. Of course, the classical assumptions are not satisfied when the data show dyadic dependence, and thus this power comparison may be misleading.

As before, the strategy rests upon Monte Carlo estimation of the significance probability under restricted-randomization. We sample the set of row/column permutations of the dependent variable data matrix, thereby preserving the dependency structure within rows and columns, and calculate a test statistic for each that measures support for the alternative hypothesis. The crucial change is that instead of using measures of group separation, we use two measures that respond to separations which conform to the user-specified ordering. One measure is best for continuous data; the other is better for binary data. At need, one could refine this approach to handle intermediate cases, such as ordinal data with a small number of categories.

For continuous data, the program calculates the Jonckheere-Terpstra statistic for the prespecified ordering, and counts the number of pseudo-data matrices that give values which equal or exceed the actual dependent variable matrix. The Jonckheere-Terpstra test is a commonly used rank procedure; its test statistic is just the sum of all $r(r-1)/2$ one-sided Mann-Whitney test statistics that correspond to one of the prespecified pairwise orderings. Since the procedure is based on the ranks of the observations and not the actual values, the result is resistant to outliers and small perturbations of the data. For example, consider the data in Figure 1. Suppose we suspect that male-male pairs spend the least time interacting. This is an expected ordering of $1 < 2 < 3$. The traditional Jonckheere-Terpstra test finds the significance probability of the prescribed ordering as 0.043. After we restrict the randomization to take account of the row-column dependency, the significance probability of the prescribed ordering decreases to 0.034. In other applications the difference can be much larger.

When the dependent variable is binary, the Jonckheere-Terpstra statistic encounters a problem with excessive ties. In this case our program calculates an alternative measure. Let $\hat{p}_i$, $i = 1, \ldots, r$ be the proportion of ones in each of the r D-groups. Without loss of generality, assume that the prespecified ordering is $1 < 2 < 3$. Then our measure of conformity is $\sum_{i=1}^{r-1}(\hat{p}_{i+1} - \hat{p}_i)^2 \text{signum}(\hat{p}_{i+1} - \hat{p}_i)$.[4] This statistic increases if and only if the observed proportions agree with the prespecified order. As before, there are many alternative measures one might use (a natural candi-

[4]Signum is a standard operator that gives the sign of the argument. It is 1 if the variable is positive, −1 if it is negative, and 0 when the variable is 0.

date is $\sum_{i=1}^{r-1}(\hat{p}_{i+1} - \hat{p}_i)\log(\hat{p}_{i+1} - \hat{p}_i)^+$, which attains its maximum when the p_i are spaced equally and in the correct order) but in most situations all sensible measures agree.

When using ordered-alternative tests there is the temptation to examine more than a single prespecified ordering. This can be useful in exploratory analysis, but it is incorrect to view the significance probabilities generated by multiple applications as exact. This relates to the traditional problem of making repeated tests of significance using the same set of data; one finds that the true significance probability always is greater than the result of the most significant test.

APPLICATIONS TO THE ANALYSIS OF SOCIAL DATA

We illustrate the proposed techniques with applications to three different data sets. Prior to these applications each of the data sets is described briefly.

Sampson's Monastery Data

Our first example concerns the classic Sampson monastery data. Sampson (1968) describes a twelve month study of the relations between monks in a contemporary American monastery. During the study conflict arose, the social group disintegrated, and a large number of members left. Among the data collected by Sampson are answers to a set of sociometric questions on various relations (Affect, Esteem, Influence, and Sanctioning) and information on the order of leaving. One question one might want to address is whether pairs of individuals who are more similar in their pattern of sociometric relationships are more likely to leave at the same time. This is a group-difference question.

Since our interest is in answering this question and not in locating groups given the sociometric data, we shall use the set of A-groups previously located by Breiger, Boorman and Arabie (1975). Breiger, Boorman and Arabie (1975) applied the CONCOR algorithm to Sampson's measurements of the Affect, Esteem, Influence, and Sanctioning relations between monks at time 4. Time 4 immediately precedes the conflict and breakup. They generated a three block partition of the monks that corresponds well with Sampson's division of the monks into the groups "Loyal Opposition", "Young Turk", and "Outcast". Table 1 shows the Breiger, Boorman and Arabie partition of the monks into groups; we use a 1 for pairs of individuals who are both members of the Loyal Opposition, 2 for pairs who are both members of the Young Turks, and 3 for pairs who are both Outcasts. We will use this matrix as our independent or D-group matrix.

Our dependent variable is shared order of leaving. Table 2 contains these data such that pairs of individuals receive a 1 if they both stayed, a 2 if they both left in last stage, a 3 if they both left in the first stage, and a 4 if they both were forced to leave. Thus 1 through 4 gives the reverse order of leaving.

We can determine whether sociometrically similar pairs of monks are more likely to leave the monastery at the same time by applying the group-difference test developed in Section 2. Based on Sampson's analysis of the monastery, however, one might wish to ask a more specific question such as does similarity in pattern of

TABLE 1
Sampson: Group Matrix—Structural Groups

0	2	0	0	0	0	2	0	0	0	0	2	0	2	2	2	0	0
2	0	0	0	0	0	2	0	0	0	0	2	0	2	2	2	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3
0	0	0	0	1	1	0	1	1	1	1	0	1	0	0	0	0	0
0	0	0	1	0	1	0	1	1	1	1	0	1	0	0	0	0	0
0	0	0	1	1	0	0	1	1	1	1	0	1	0	0	0	0	0
2	2	0	0	0	0	0	0	0	0	0	2	0	2	2	2	0	0
0	0	0	1	1	1	0	0	1	1	1	0	1	0	0	0	0	0
0	0	0	1	1	1	0	1	0	1	1	0	1	0	0	0	0	0
0	0	0	1	1	1	0	1	1	0	1	0	1	0	0	0	0	0
0	0	0	1	1	1	0	1	1	1	0	0	1	0	0	0	0	0
2	2	0	0	0	0	2	0	0	0	0	0	0	2	2	2	0	0
0	0	0	1	1	1	0	1	1	1	1	0	0	0	0	0	0	0
2	2	0	0	0	0	2	0	0	0	0	2	0	0	2	2	0	0
2	2	0	0	0	0	2	0	0	0	0	2	0	2	0	2	0	0
2	2	0	0	0	0	2	0	0	0	0	2	0	2	2	0	0	0
0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0

TABLE 2
Sampson: Dependent-Variable Matrix—Shared Order of Leaving

0	0	0	0	0	0	3	0	0	0	0	0	0	3	3	3	0	0
0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	4	4
0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	4
0	0	0	0	0	0	0	2	0	2	0	2	2	0	0	0	0	0
0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	3	3	3	0	0
0	0	0	2	0	0	0	0	0	2	0	2	2	0	0	0	0	0
0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0
0	0	0	2	0	0	0	2	0	0	0	2	2	0	0	0	0	0
0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	2	0	0	0	2	0	2	0	0	2	0	0	0	0	0
0	0	0	2	0	0	0	2	0	2	0	2	0	0	0	0	0	0
3	0	0	0	0	0	3	0	0	0	0	0	0	0	3	3	0	0
3	0	0	0	0	0	3	0	0	0	0	0	0	3	0	3	0	0
3	0	0	0	0	0	3	0	0	0	0	0	0	3	3	0	0	0
0	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
0	4	4	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0

TABLE 3
Sampson: Results of Traditional and Proposed Tests

	Statistic		Significance Levels Traditional	Proposed
Group-Difference				
F	12.404,	*df* = 2, 49	less than 0.001	0.002
Kruskal-Wallis	12.127		less than 0.001	0.003
Ordered-Alternatives				
Jonckheere-Terpstra			0.146	0.001

sociometric relations predict the specific order of shared leaving. Or, more specifically, does shared membership in the Loyal Opposition lead people to leave not only at the same time but first, the Young Turks second, and the Outcasts last? This is an ordered-alternative question. In order to illustrate the proposed methods we will apply both the traditional statistical tests that do not take the network structure into account and the proposed permutation tests that take the network structure into account by restricting the randomization. The results of all tests are shown in Table 3.

For the group-difference test, when the network structure is not taken into account, the significance probabilities are exaggerated whether the F or the Kruskal-Wallis statistic is used. Yet, even our more conservative approach suggests that co-group members are likely to leave at the same time. For the ordered-alternative test, the traditional test would lead the researcher to conclude that co-group members probably did not leave at the same time. Whereas, the proposed test suggests the opposite conclusion. This analysis indicates that when the network structure is taken into account co-group members are not only likely to leave at the same time but that structural similarity correctly predicts the order of leaving.

Carley's Third East Tutor Selection Data

Carley (1984) examined the relationship between undergraduate interaction in an MIT dormitory and the students' group decision-making process in selecting a new tutor (graduate resident). The data included both an interaction matrix and cognitive maps. A cognitive map describes a single student's conception of the tutor's role at a particular time. These maps consist of a series of statements (called facts in Carley, 1986) identifying each student's criteria for determining whether someone would make a good tutor. For every every pair of students, one counts the number of statements in their maps in common. This number is an indication of the amount of information or knowledge that the students shared. Carley calculated the amount of shared knowledge at both the beginning and end of the decision making process. The difference in these counts from the beginning to the end of the process represents the increase in shared knowledge between that pair of students. Given these data a series of questions can be asked. For example, are pairs of individuals that interact frequently more likely to have a greater increase in shared knowledge over time than pairs of individuals who interact infrequently. Or, are pairs of individuals with similar patterns of interactions (i.e., pairs who are structurally equivalent (Lorrain and White, 1971)) more likely to have a greater increase in shared knowledge over time than pairs of individuals with dissimilar patterns. These are both group-difference questions. Carley (1989) used simple t-tests to answer these questions.

First, Carley argued that change in shared knowledge increases with increasing levels of interaction. We refer to this as the first group-difference hypothesis. In making this argument Carley defined two D-groups based on level of interaction—D-group 1, pairs of students who are infrequent interactors (a raw interaction level of 1 or 2) and D-group 2 pairs of students who are frequent interactors (a raw interaction level of 3, 4 or 5). In Table 4 the raw interaction data is shown on the

TABLE 4
Carley: Group Matirx—Raw & Collapsed Interaction

	Raw Interaction	Low/High Interaction
Johann	0 3 2 4 3 1 3 3 5 4 2 2 3 1 4	0 2 1 2 2 1 2 2 2 2 1 1 2 1 2
Lorenzo	3 0 3 2 3 4 4 5 2 4 3 2 4 4 3	2 0 2 1 2 2 2 2 1 2 2 1 2 2 2
Ian	2 3 0 2 3 3 2 3 2 2 2 2 5 2 2	1 2 0 1 2 2 1 2 1 1 1 1 2 1 1
Zebediah	4 2 2 0 1 2 3 3 3 2 5 1 1 3 2	2 1 1 0 1 1 2 2 2 1 2 1 1 2 1
Hilda	3 3 3 1 0 3 4 3 3 3 2 2 3 4 3	2 2 2 1 0 2 2 2 2 2 1 1 2 2 2
Ted	1 4 3 2 3 0 2 4 2 4 3 1 3 3 3	1 2 2 1 2 0 1 2 1 2 2 1 2 2 2
Lowell	3 4 2 3 4 2 0 5 2 3 2 3 3 5 2	2 2 1 2 2 1 0 2 1 2 1 2 2 2 1
Jacques	3 5 3 3 3 4 5 0 2 3 2 2 3 5 2	2 2 2 2 2 2 2 0 1 2 1 1 2 2 1
Eunice	5 2 2 3 3 2 2 2 0 4 2 2 2 1 4	2 1 1 2 2 1 1 1 0 2 1 1 1 1 2
Jubal	4 4 2 2 3 4 3 3 4 0 2 3 4 4 5	2 2 1 1 2 2 2 2 2 0 1 2 2 2 2
Deety	2 3 2 5 2 3 2 2 2 2 0 3 2 3 2	1 2 1 2 1 2 1 1 1 1 0 2 1 2 1
Woodie	2 2 2 1 2 1 3 2 2 3 3 0 2 2 2	1 1 1 1 1 1 2 1 1 2 2 0 1 1 1
Maureen	3 4 5 1 3 3 3 3 2 4 2 2 0 3 3	2 2 2 1 2 2 2 2 1 2 1 1 0 2 2
Hazel	1 4 2 3 4 3 5 5 1 4 3 2 3 0 2	1 2 1 2 2 2 2 2 1 2 2 1 2 0 1
Mannie	4 3 2 2 3 3 2 2 4 5 2 2 3 2 0	2 2 1 1 2 2 1 1 2 2 1 1 2 1 0

TABLE 5
Carley: Dependent-Variable Matrix—Change in Shared Knowledge

Johann	0	98	−13	44	75	73	126	116	32	66	89	71	57	63	69
Lorenzo	98	0	−31	−36	77	89	139	100	2	41	37	54	19	57	33
Ian	−13	−31	0	−147	3	11	9	−12	−38	−2	−115	−54	−70	−60	−81
Zebediah	44	−36	−147	0	27	94	68	41	−17	28	−17	−29	−38	−48	36
Hilda	75	77	3	27	0	59	103	46	11	24	51	51	25	60	17
Ted	73	89	11	94	59	0	119	92	20	4	99	68	62	77	42
Lowell	126	139	9	68	103	119	0	132	19	49	134	171	89	102	110
Jacques	116	100	−12	41	46	92	132	0	7	26	72	84	51	49	24
Eunice	32	2	−38	−17	11	20	19	7	0	0	−24	−7	−8	−24	−54
Jubal	66	41	−2	28	24	4	49	26	0	0	−2	29	11	12	5
Deety	89	37	−115	−17	51	99	134	72	−24	−2	0	12	−6	−56	13
Woodie	71	54	−54	−29	51	68	171	84	−7	29	12	0	−5	25	33
Maureen	57	19	−70	−38	25	62	89	51	−8	11	−6	−5	0	8	36
Hazel	63	57	−60	−48	60	77	102	49	−24	12	−56	25	8	0	−42
Mannie	69	33	−81	36	17	42	110	24	−54	5	13	33	36	−42	0

left, and the collapsed interaction data on the right.[5] For the dependent variable, Carley used the increase in shared knowledge over time for that pair of students. This data is contained in Table 5. Using a *t*-test Carley then contrasted the mean increases in shared knowledge for pairs of students who infrequently interacted with pairs of students who infrequently interacted with pairs of students who frequently interacted. The apparent significance probability of the *t*-test was less than 0.010.

Second, Carley argued that structurally equivalent dyads are as likely to show an increase in shared knowledge as structurally different dyads. We refer to this as the second group-difference hypothesis. In making this argument, Carley first applied CONCOR (Breiger, Boorman, and Arabie, 1975) to the interaction data to locate structural groups (see the lefthand matrix in Table 6) and then combined these pairs into two D-groups—pairs of students who are not structurally equivalent in their

[5]Data on pairwise student interactions were gathered halfway through the decision making process.

TABLE 6
Carley: Group Matrices—Structural Equivalence Groups

	Equivalence Groups	Equivalent or Not
Johann	0 0 0 1 0 0 0 0 1 0 0 1 0 0 0	0 1 1 2 1 1 1 1 2 1 1 2 1 1 1
Lorenzo	0 0 0 0 0 4 0 4 0 4 0 0 0 0 0	1 0 1 1 1 2 1 2 1 2 1 1 1 1 1
Ian	0 0 0 0 0 0 0 0 0 0 3 0 3 0 3	1 1 0 1 1 1 1 1 1 1 2 1 2 1 2
Zebediah	1 0 0 0 0 0 0 0 1 0 0 1 0 0 0	2 1 1 0 1 1 1 1 2 1 1 2 1 1 1
Hilda	0 0 0 0 0 0 2 0 0 0 0 0 0 2 0	1 1 1 1 0 1 2 1 1 1 1 1 1 2 1
Ted	0 4 0 0 0 0 0 4 0 4 0 0 0 0 0	1 2 1 1 1 0 1 2 1 2 1 1 1 1 1
Lowell	0 0 0 0 2 0 0 0 0 0 0 0 0 2 0	1 1 1 1 2 1 0 1 1 1 1 1 1 2 1
Jacques	0 4 0 0 0 4 0 0 0 4 0 0 0 0 0	1 2 1 1 1 2 1 0 1 2 1 1 1 1 1
Eunice	1 0 0 1 0 0 0 0 0 0 0 1 0 0 0	2 1 1 2 1 1 1 1 0 1 1 2 1 1 1
Jubal	0 4 0 0 0 4 0 4 0 0 0 0 0 0 0	1 2 1 1 1 2 1 2 1 0 1 1 1 1 1
Deety	0 0 3 0 0 0 0 0 0 0 0 0 3 0 3	1 1 2 1 1 1 1 1 1 1 0 1 2 1 2
Woodie	1 0 0 1 0 0 0 0 1 0 0 0 0 0 0	2 1 1 2 1 1 1 1 2 1 1 0 1 1 1
Maureen	0 0 3 0 0 0 0 0 0 0 3 0 0 0 3	1 1 2 1 1 1 1 1 1 1 2 1 0 1 2
Hazel	0 0 0 0 2 0 2 0 0 0 0 0 0 0 0	1 1 1 1 2 1 2 1 1 1 1 1 1 0 1
Mannie	0 0 3 0 0 0 0 0 0 0 3 0 3 0 0	1 1 2 1 1 1 1 1 1 1 2 1 2 1 0

pattern of interaction (1) and pairs of students who are structurally equivalent in their pattern of interaction (2) (see the righthand matrix in Table 6).[6] For the dependent variable, Carley again used the increase in shared knowledge over time for that pair of students. Using a t-test Carley then contrasted the mean increases in shared knowledge for pairs of students who are structurally equivalent with pairs of students who are not. The apparent significance probability of the t-test was greater than .25.

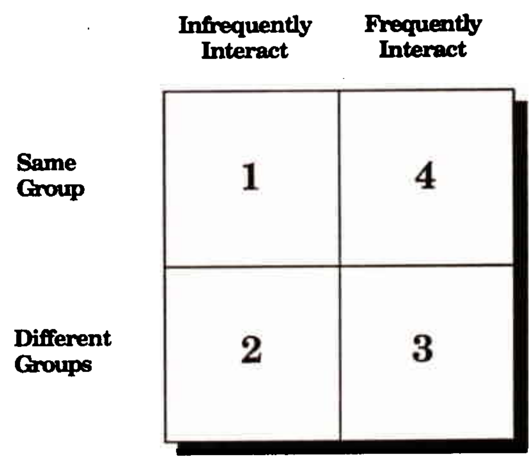
Beyond these two original group-difference questions, Carley (1989) also addressed an ordered-alternative question. Based on the constructural model, she predicted that higher levels of interaction cause larger increases in shared knowledge. Defining five D-groups by interaction level as in the lefthand matrix of Table 4, this implies that group centers for the data in Table 6 are ordered as $1 < 2 < 3 < 4 < 5$. We refer to this as the first ordered-alternative hypothesis.

In a somewhat more complicated argument, Carley suggested that structural equivalence tends to enhance the impact of interaction. Thus infrequent interactors who are structurally equivalent are less likely than structurally different infrequent interactors to increase their shared knowledge. Conversely, frequent interactors who are structurally equivalent are more likely than frequent interactors who are not structurally equivalent to increase shared knowledge. This too, is an ordered-alternative question. Figure 2 shows the predicted relationship between interaction and structure under Carley's constructural model. This assumes that the students are culturally heterogeneous and none holds a major fraction of the possible statements. Labeling the cells in Figure 2 counter-clockwise as 1, 2, 3, 4 the constructural argument predicts that the order of medians for the data in Table 7 where the pairs of student are devided into D-groups as in Figure 2 are ordered as $1 < 2 < 3 < 4$. We refer to this as the second ordered-alternative hypothesis.

[6]Carley applied CONCOR to interaction data from all 42 students involved in the selection locating 4 structural groups. We display data only for the 15 students for whom shared knowledge data is available.

Constructural Prediction for Dyadic Behavior

Each cell contains the average increase in shared knowledge for dyads with those characteristics



Order of averages for dependent variable:

1 < 2 < 3 < 4

FIGURE 2. Structure magnifies interaction.

TABLE 7
Carley: Group Matrix—Combined Structural Equivalence and Interaction

Johann	0	3	2	1	3	2	3	3	1	3	2	1	3	2	3
Lorenzo	3	0	3	2	3	4	3	4	2	4	3	2	3	3	3
Ian	2	3	0	2	3	3	2	3	2	2	1	2	1	2	1
Zebediah	1	2	2	0	2	2	3	3	1	2	3	1	2	3	2
Hilda	3	3	3	2	0	3	4	3	3	3	2	2	3	4	3
Ted	2	4	3	2	3	0	2	4	2	4	3	2	3	3	3
Lowell	3	3	2	3	4	2	0	3	2	3	2	3	3	4	2
Jacques	3	4	3	3	3	4	3	0	2	4	2	2	3	3	2
Eunice	1	2	2	1	3	2	2	2	0	3	2	1	2	2	3
Jubal	3	4	2	2	3	4	3	4	3	0	2	3	3	3	3
Deety	2	3	1	3	2	3	2	2	2	2	0	3	1	3	1
Woodie	1	2	2	1	2	2	3	2	1	3	3	0	2	2	2
Maureen	3	3	1	2	3	3	3	3	2	3	1	2	0	3	1
Hazel	2	3	2	3	4	3	4	3	2	3	3	2	3	0	2
Mannie	3	3	1	2	3	3	2	2	3	3	1	2	1	2	0

D-Group STRUT1801 consists of D-groups 1 and 3 from Table 6, which are the structurally equivalent infrequent interaction groups. D-Group 2 consists of dyads that are structurally different and have either a level 1 or 2 for interaction. D-Group 3 consists of dyads who are structurally different and have either a level 3, 4 or 5 for interaction. D-Group 4 consists of D-groups 2 and 4 from Table 6, which are the structurally equivalent frequent interaction groups.

TABLE 8
Carley: Results of Traditional and Proposed Tests

	Statistic		Significance Levels Traditional	Proposed
Group-Difference				
Hypothesis 1				
F	7.385,	*df* = 1, 103	0.008	0.053
Kruskal-Wallis	6.460		0.011	0.061
Hypothesis 2				
F	0.491,	*df* = 1, 103	0.485	0.209
Kruskal-Wallis	0.121		0.730	0.525
Ordered-Alternatives				
Hypothesis 1				
Jonckheere-Terpstra	$1 < 2 < 3 < 4 < 5$		0.010	0.036
Hypothesis 2				
Jonckheere-Terpstra	$1 < 2 < 3 < 4$		less than 0.001	0.019

We examine Carley's group-difference and ordered-alternative hypotheses using both traditional tests and those we proposed. The results are shown in Table 8. In this examination we will use those D-groups defined by Carley. For the group-difference tests, for both hypotheses 1 and 2, the results conform with Carley's previous conclusion, but our reanalysis respects dependencies in the data. The reader should in this case notice the relatively large discrepancy between the significance probabilities obtained under the restricted-randomization tests for the *F* and Kruskal-Wallis tests. This discrepancy occurs because the Kruskal-Wallis test tends to lose power in the presence of many tied observations, as happens in this dataset. Additionally, the reader will note that the proposed tests provide more conservative estimates than the traditional tests.

Moving on to the ordered-alternative hypotheses we see that the results support Carley's model, but unlike the traditional test, the permutation version by taking the network structure into account does not exaggerate the significance probability. In the Sampson data, applying restricted-randomization to the ordered-alternative test yielded a less conservative inference. In contrast, with the Carley data, this test yielded a more conservative inference. The restricted-randomization analysis, however, offers a more accurate assessment of the degree of support than is available from traditional procedures.

Humana's Human Rights Data

Humana (1983) rates 74 different countries on four point scales that reflect their practices for 40 different human rights. These rights include such items as *Right of peaceful assembly and association*, *Freedom to leave own country*, and *Right to use contraceptive pills and devices*. A score of 4 indicates that the country respects that right, and a score of 1 indicates that the right is broadly violated. Also, Humana indicates the per capita income of all 74 countries in the sample.

This application extends the proposed method beyond the scope of the two previous studies. First, the units of analysis are countries rather than people. Second, the

TABLE 9
Humana: Group Data—Economic Levels

Country ID, Countries and Per Capita Income								
Level 1								
29	Bang	110	3	Ethi	120	31	Indi	190
49	Viet	200	6	Moza	250	45	SriL	250
11	Tanz	300	13	Zair	300	40	Paki	340
30	Chin	350	10	Suda	370	4	Keny	380
32	Indo	431	2	Egyp	450	8	Sene	500
9	SAfr	500	14	Zamb	550	48	Thai	650
15	Zimb	700	41	Papu	750	42	Phil	800
26	Peru	820	5	Moro	850	7	Nige	1000
24	Colo	1200	25	Ecua	1200	46	Syri	1200
36	NKor	1300	12	Tuni	1400	17	Cuba	1400
71	Turk	1450	19	Pana	1550	38	Mals	1586
37	SKor	1650	1	Alg	1935	22	Braz	1995
23	Chil	2000	67	Roma	2000	66	Port	2375
33	Iraq	2410	47	Taiw	2500	21	Arge	2500
18	Mexi	2720	52	Bulg	2750			
Level 2								
44	Sing	3500	74	Yugo	3500	65	Pola	3500
60	Hung	4000	73	USSR	4110	59	Gree	4250
27	Vene	4700	34	Isra	4900	61	Irel	5000
53	Czec	5510	68	Spai	5578	57	GrDR	6430
62	Ital	6914	39	NewZ	7000	72	UKin	7500
Level 3								
35	Japa	9200	63	Neth	10175	50	Aust	10255
16	Cana	10400	28	Astl	10500	55	Finl	10500
51	Belg	10890	43	Saud	11000	56	Fran	11500
64	Norw	11800	69	Swed	11920	20	USA	12000
58	GrFR	12500	54	Denm	12761	70	Swit	16500

data set is much larger. Third, these data have not previously been analyzed from a social network perspective.

Our examination of Humana's data looks at pairwise rights agreement between countries grouped according to three income levels. The three groups consist of all pairs of countries within the same income category. Those three income categories are: (1) impoverished—less than $3000/person/year; (2) mid-income—between $3000 and $9000; and (3) wealthy—greater than $9000. These levels were determined as natural breakpoints in the histogram of national per capita incomes in 1983 (Table 9). There are 15 countries in category 3, 15 in catory 2, and 44 in category 1. Our analysis excludes pairs of countries with different income levels.

The 40 rights variables are used to generate five different dependent variables that capture different aspects of pairwise national behavior. These are common behavior, shared rights, shared military rights, shared free speech rights, and shared strong rights. In all cases we use a very restrictive definition of "same"; i.e., both nations have to score the same number on the right in question. We use this restrictive definition in order to maintain the categorical nature of the data.

Common Behavior: This counts the number of rights for which the two nations have the same level of protection. Thus nation pairs that both protect a given human right (score 4 on Humana's scale) gain a point, as do pairs of nations that fail to respect that right at the same level (i.e., both score 3, or score 2, or score 1).

Shared Rights: This counts the number of protected rights (scores of 4) common to any pair of countries.

Shared Military Rights: Banks (1989) uses variable clustering to identify five variables among the 40 that shown strong internal covariation. For every pair of nations, Shared Military Rights counts the number of these variables that are both scored at level 4. These five variables are: (1) lethality of weapons normally carried by civil police (this is measured on a four point scale, with 4 representing low routine levels of force); (2) severity of punishment for refusing compulsory national service; (3) freedom from compulsory military service; (4) ratio of police and military to citizens, on a four point scale (a 4 indicates a low ratio); (5) proportion of national income spent on above (this is measured on a four point scale with low proportions scored as 4).

Shared Free Speech Rights: This is similar to Shared Military Rights, except that the five freedoms considered represent a cohesive group of free speech variables identified in Banks (1989). The five freedoms are: (1) freedom from censorship of mail; (2) right of peaceful assembly and association; (3) right of peaceful political opposition; (4) severity of punishment for non-violent antigovernment activities; (5) severity of punishment for possession of banned literature. Pairs of nations receive an increase of one point for each of these freedoms that they both protect at level 4.

Strong Rights: This is a binary variable. A pair of countries receives a 1 if both nations respect 15 common rights at level 4. Otherwise, the nation-pair receives a 0.

The full definition of all rights are presented in Table 10. The raw data used to generate the matrices previously described are presented in Table 11. The rights in Table 11 are in the same order as in Table 10. The concern in this study is to discover how strongly D-groups determined by shared per capita income levels predicts common rights behavior (in terms of the dependent variables defined above). We address these questions in the context of network comparisons, but recognize that alternative statistical methods enable one to consider very similar questions.

Our first hypothesis is that pairs of countries with different economic levels show different patterns of human rights behavior. This is a group-difference hypothesis. To test this hypothesis we run the restricted-randomization permutation test for group-differences on each of the five dependent variables. The significance probabilities for both the F and Kruskal-Wallis measure are reported in Table 12. Our second hypothesis is that the higher the economic level the more similar the pair of countries are in their guarantee of rights. In other words, poor countries are more heterogeneous with respect to human rights than are moderately well off countries which are in turn more heterogeneous than wealthy countries. These are ordered-alternatives tests. To test this set of hypotheses we run the restricted-randomization permutation test for ordered-alternatives on each of the five dependent variables. The significance probabilities for the Jonckheere-Terpstra measure are also reported in Table 12.

TABLE 10
Humana: Human Rights

Right
Right to early abortion.
Right to purchase and drink alcohol.
Severity of punishment for non-violent antigovernment activities.
Freedom to practice any religion.
Right of peaceful assembly and association.
Severity of punishment for possession of banned literature.
Freedom of book publishing.
Right to use contraceptive pills and devices.
Prevalence of capital punishment by the state.
Severity of punishment for refusing compulsory national service.
Proportion of national income spent on above.
Freedom from police detention without charge.
Right of divorce (for men and women equally).
Right to publish and educate in ethnic languages.
Ratio of police and military to citizens, on a four point scale.
Freedom from state policies to control artistic works.
Right to seek information and teach ideas.
Freedom from censorship of mail.
Freedom from political press censorship.
Freedom from directed employment or work permits.
Right to practice homosexuality between consenting adults.
Right of all courts to total independence.
Freedom from deprivation of nationality.
Freedom from compulsory military service.
Right of inter-racial, inter-religious and civil marriage.
Freedom of movement within own country.
Freedom to leave own country.
Freedom from compulsory religion or state ideology in schools.
Freedom from police searches of home without warrant.
Right of peaceful political opposition.
Right of assumption of innocence until guilt proved.
Freedom from civilian trials in secret.
Freedom from corporal punishment by state.
Right of accused to be promptly brought before judge or court.
Freedom from serfdom, slavery or forced child labor.
Freedom from torture or coercion by state.
Freedom of radio and television broadcasts from state control.
Right of independent trade unions.
Lethality of weapons normally carried by civil police.
Right of women to equality.

We find support for the hypotheses (though the least with shared rights). For both the group-difference tests and the ordered-alternative tests the restricted-randomization permutation tests yield more conservative inferences. Of particular interest are the results for strong rights. The strong rights measure was a binary network. With such data there are no standard tests for either a group-difference test or an ordered-alternative test. However, the restricted-permutation test is directly applicable. As can be seen, even when a strong notion of shared rights is chosen, we find evidence for both group differences and differences ordered by economic level.

TABLE 11
Humana—Dependent Variable—Countries by Rights

1	Alg	4	4	4	3	3	1	2	3	4	4	3	4	2	3	3	3	4	3	3	4	1	3	2	4	2	4	2	4	2	3	4	1	1	2	4	3	2	2	2	3
2	Egyp	4	4	4	2	3	2	3	3	4	2	4	2	2	3	3	3	4	3	3	3	2	3	4	4	2	4	2	4	2	4	4	2	4	2	2	3	1	2	3	2
3	Ethi	2	2	1	1	1	1	1	2	1	4	1	1	1	1	1	1	1	1	1	1	1	1	2	1	4	2	3	2	3	1	1	1	2	2	2	1	1	1	1	1
4	Keny	3	3	4	2	3	1	2	2	4	4	4	4	2	2	3	3	2	2	2	1	2	3	4	2	4	4	1	4	2	4	2	2	3	4	4	4	1	1	2	4
5	Moro	4	2	4	3	2	2	2	3	4	4	4	2	2	1	2	3	1	2	4	3	3	2	4	3	1	4	1	1	2	4	4	3	3	3	3	2	2	2	3	3
6	Moza	2	1	4	1	2	1	1	3	1	4	3	1	1	1	2	2	2	1	3	2	1	2	4	2	1	4	3	4	2	1	2	1	1	4	2	3	1	1	3	2
7	Nige	4	4	4	4	2	4	4	1	4	4	4	4	3	2	4	1	3	1	3	3	3	2	4	4	4	3	1	3	2	3	3	2	4	4	4	4	1	2	3	4
8	Sene	4	4	4	4	3	3	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4	2	3	2	4	3	3	4	4	4	4	1	4	4	4
9	SAfr	1	1	1	1	2	1	1	1	1	1	3	4	2	1	1	1	1	1	1	1	1	1	3	3	1	4	3	4	3	4	3	1	2	3	3	3	1	2	2	2
10	Suda	3	3	4	2	3	1	2	2	4	3	4	2	3	1	3	2	2	3	1	2	2	3	4	4	4	4	1	3	2	4	3	2	3	4	3	3	1	1	2	4
11	Tanz	4	2	4	3	3	2	1	3	4	4	4	4	1	1	2	3	2	2	3	4	1	2	4	4	4	4	2	4	2	4	3	1	2	4	3	4	1	2	3	4
12	Tuni	4	3	4	3	3	2	2	3	4	4	4	2	1	1	4	3	3	1	3	3	3	2	4	3	2	4	2	4	4	4	4	1	2	4	4	3	2	2	2	3
13	Zair	2	2	1	1	2	1	2	2	4	4	4	4	1	2	1	1	1	1	2	1	1	1	3	3	3	4	1	4	2	2	3	1	3	4	4	3	1	1	2	4
14	Zamb	2	1	4	3	3	1	1	3	4	4	3	4	2	1	2	3	3	3	2	3	2	3	4	4	4	4	1	4	3	3	3	2	2	4	2	3	1	2	3	4
15	Zimb	3	3	4	4	3	3	3	3	4	4	4	4	3	1	2	2	2	3	2	2	3	3	4	4	4	4	4	4	2	4	3	3	3	3	2	4	2	2	3	4
16	Cana	4	4	4	4	4	4	4	3	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4	4	4	4	4	4	2	3	4	4	4	3	4	3	4	4	4	4
17	Cuba	4	1	1	1	1	1	1	3	1	4	3	1	1	1	2	2	1	1	1	1	1	1	4	2	1	4	1	4	4	4	3	1	1	1	2	3	1	2	2	2
18	Mexi	4	4	4	3	3	2	3	3	4	4	3	4	3	1	2	2	2	2	2	3	2	3	4	4	3	4	1	4	2	4	3	3	3	4	4	3	4	2	2	3
19	Pana	4	4	4	4	3	3	3	3	4	4	4	4	3	4	4	3	3	4	3	4	4	4	4	4	4	4	1	4	1	4	4	3	3	3	4	3	4	3	4	4
20	USA	4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	3	4	4	4	4	4	4	4	4	4	4	3	4	3	4	4	4	4	2	3	3	3	4	4	4
21	Arge	4	3	2	2	3	1	1	3	4	4	4	4	2	1	2	1	1	1	1	2	2	1	4	4	1	4	4	3	2	1	2	3	2	2	3	3	1	1	2	2
22	Braz	4	4	4	4	2	2	3	3	4	4	4	4	2	3	3	3	3	3	3	4	2	3	4	4	2	4	4	3	2	1	3	3	3	4	4	3	4	2	2	2
23	Chil	2	3	2	1	3	1	1	2	4	4	4	4	2	1	1	1	1	1	1	1	2	1	4	4	1	4	1	4	2	1	1	2	2	2	3	3	1	1	2	3
24	Colo	2	4	4	3	3	3	3	3	4	4	4	4	3	2	2	2	2	2	3	3	3	3	4	4	1	4	1	4	2	3	3	2	3	3	3	2	3	2	3	2
25	Ecua	4	3	4	4	3	4	3	3	4	4	4	4	4	4	4	3	4	2	4	4	4	4	4	4	1	4	1	4	2	4	3	4	4	3	4	4	4	4	4	3
26	Peru	4	4	4	4	3	3	3	3	4	4	4	4	3	3	3	3	3	3	4	4	3	4	4	4	1	4	4	3	2	4	4	4	4	3	4	3	1	4	4	3
27	Venz	4	4	4	4	4	4	4	3	4	4	4	4	3	3	4	3	4	4	4	4	4	4	4	4	1	4	4	4	2	3	4	3	4	4	4	3	4	4	4	3
28	Astl	4	4	4	4	3	4	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	3	4	4	4	4	3	4	3	3	4	4	4
29	Bang	3	3	4	4	1	2	3	2	4	3	4	2	3	1	3	3	2	3	4	3	2	3	4	4	4	2	1	4	2	3	4	2	3	4	4	4	1	2	3	4
30	Chin	2	2	4	1	3	1	1	3	1	4	3	1	1	1	1	3	1	1	2	1	1	1	3	1	1	4	1	4	4	4	2	2	1	2	2	3	1	1	1	3
31	Indi	4	4	4	4	2	3	3	3	4	4	4	4	3	1	2	2	1	3	3	2	2	3	4	4	4	4	1	4	3	3	4	3	4	4	3	4	1	2	3	4
32	Indo	3	3	4	2	3	2	2	3	3	3	4	1	2	1	2	2	2	1	2	2	2	3	1	3	4	4	2	4	2	4	4	3	3	3	3	3	1	2	2	4
33	Iraq	1	3	1	1	3	1	1	3	3	4	3	2	1	1	2	1	1	1	1	1	1	2	3	3	1	3	1	3	2	2	2	1	1	2	2	3	1	1	2	2
34	Isra	4	4	3	4	4	3	3	4	4	1	4	2	3	4	3	3	4	4	4	3	4	3	4	4	1	4	2	4	3	3	4	3	3	1	1	3	3	3	4	2
35	Jap	4	4	4	4	3	4	4	2	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	3	3	4	4	4	4	3	4	3	1	4	4	4
36	NKor	1	1	4	1	3	1	1	3	1	4	2	1	1	1	1	1	1	1	1	1	1	1	4	1	1	3	1	3	3	3	3	1	1	1	2	2	1	1	2	1
37	SKor	3	3	4	2	3	1	1	3	4	4	4	4	2	2	3	1	2	1	2	1	1	2	4	4	1	4	1	4	3	3	3	2	2	2	3	3	1	2	2	2
38	Mala	4	3	3	2	3	1	2	2	4	3	4	2	2	1	1	3	1	2	2	1	3	2	4	4	4	4	1	4	2	3	3	2	3	3	3	3	1	3	3	4
39	NewZ	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	1	4	3	4	4	4	4	3	4	4	4	4	4	4
40	Paki	3	3	3	1	1	1	1	1	4	1	3	1	1	1	3	3	3	3	3	3	2	1	4	3	4	1	2	4	2	2	2	1	2	3	3	2	1	2	2	4
41	Papu	4	4	4	4	3	4	4	3	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4	3	4	3	3	4	3	4	4	4	4	4	3	4	4
42	Phil	1	3	4	2	3	2	2	3	4	4	4	4	2	1	2	1	1	2	1	1	3	1	4	4	3	4	3	4	2	1	2	2	3	3	3	3	1	2	2	3
43	Saud	3	2	4	1	3	1	1	1	1	1	1	1	2	1	3	2	3	1	2	2	1	2	4	1	4	1	2	3	2	2	1	1	1	3	2	2	1	1	1	4
44	Sing	4	3	3	3	4	2	2	3	4	4	4	4	2	2	3	2	1	1	4	1	3	4	4	4	1	4	2	4	4	4	2	2	3	1	2	3	1	3	3	3
45	SriL	4	4	4	4	3	3	3	4	4	4	4	2	2	1	2	3	2	3	4	3	3	3	4	4	4	4	2	4	2	4	3	3	3	4	4	4	2	3	3	4
46	Syri	3	3	2	1	3	1	1	2	4	4	3	2	2	1	1	1	1	1	2	1	1	1	4	3	1	1	1	4	3	3	3	1	2	1	1	3	1	1	1	2
47	Taiw	4	3	3	2	3	1	2	3	4	4	3	4	1	2	3	2	3	3	2	2	2	3	3	4	1	4	1	4	3	3	4	2	2	1	3	3	1	2	2	2
48	Thai	3	4	4	3	2	2	3	3	4	4	4	4	3	2	3	3	3	2	3	2	3	4	4	4	1	4	3	4	3	4	3	1	2	3	3	3	1	2	2	2
49	Viet	2	2	4	1	2	1	1	4	1	4	3	1	1	1	1	3	1	1	1	1	1	1	3	1	1	4	2	4	4	4	3	1	1	1	2	1	1	1	1	2
50	Aust	4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4	2	4	4	4	4	3	4	2	4	4	4	4	4	4	4	4	3	4	3	4	4	4	2
51	Belg	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	2	4	4	4	4	3	4	3	3	3	4	2
52	Bulg	1	1	4	1	4	1	1	3	1	4	3	1	1	2	2	2	2	2	2	1	1	2	2	1	1	4	1	4	4	4	4	1	1	3	3	2	1	2	2	2
53	Czec	3	1	1	1	3	1	1	4	1	4	2	1	1	1	1	2	2	1	2	2	1	2	2	3	1	4	4	4	3	4	3	1	1	2	3	3	1	2	2	2
54	Denm	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	4	4	4	4	4	4	3	4	3	4	4	4	3
55	Finl	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4	4	2	4	4	4	4	4	4	4	4	3	4	4	4	4	4	3
56	Fran	4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	4	3	3	4	4	4	4	4	4	1	4	4	4	4	4	4	3	3	2	3	3	4	3	4	3
57	GDR	3	1	1	1	3	1	1	4	1	4	3	1	1	1	1	2	1	1	1	2	1	1	4	2	1	4	4	4	4	4	3	1	1	2	3	2	1	2	2	2
58	FRG	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	1	4	4	4	3	4	4	4	4	3	4	3	4	4	3	2
59	Gree	4	4	4	4	4	4	3	3	4	2	4	3	3	4	4	4	4	4	4	3	4	4	4	4	1	4	4	3	3	4	4	3	3	2	3	3	3	3	3	2
60	Hung	4	3	3	2	4	1	1	3	1	4	4	1	2	3	3	2	3	3	3	3	1	2	4	3	1	4	2	4	4	4	4	1	2	2	4	3	1	2	2	2
61	Irel	4	4	4	4	3	4	4	3	4	4	4	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	1	2	2	1	4	4	3	3	4	4	3	4	4	4
62	Ital	4	4	4	4	3	3	3	4	4	4	4	2	4	3	4	4	4	4	4	4	4	3	4	4	1	4	4	4	4	4	4	4	3	3	4	3	3	4	4	3
63	Neth	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	1	4	4	4	3	4	4	4	4	3	4	3	4	4	4	3
64	Norw	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	1	4	4	4	4	4	4	4	4	3	3	4	4	4	4	3
65	Pola	3	1	1	1	4	1	1	4	1	4	4	1	1	1	1	3	1	1	1	1	1	1	4	2	1	4	4	4	3	4	3	1	1	2	3	2	1	2	2	2
66	Port	4	4	4	4	3	3	4	3	4	4	4	4	3	4	4	4	4	3	4	4	4	4	4	4	1	4	4	4	2	4	4	4	4	3	4	3	4	3	3	2
67	Roma	3	1	1	1	2	1	2	4	1	4	3	1	1	1	1	1	1	2	1	1	1	2	2	1	1	4	1	4	4	4	3	1	1	2	4	3	1	2	2	2
68	Spai	4	4	4	4	3	4	4	3	4	4	4	3	3	3	3	2	3	4	4	4	4	3	4	4	1	4	4	3	2	4	4	3	3	2	4	3	4	2	2	2
69	Swed	4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4	4	4	1	4	4	4	4	4	4	4	4	3	4	3	4	4	4	3
70	Swit	4	4	3	4	4	4	4	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	2	4	4	3	3	3	4	4	4	4	4	3	4	4	4	3
71	Turk	3	2	1	1	3	1	1	2	4	4	4	4	1	2	1	1	2	2	2	2	1	1	4	4	1	4	3	4	2	4	3	1	1	2	3	3	1	1	2	2
72	UKin	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4	3	3	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4	3	3	4	4	4	4	4	4
73	USSR	1	1	1	1	2	1	1	4	1	4	3	1	1	1	1	2	2	1	1	1	1	1	3	1	1	4	1	4	4	4	3	1	1	2	2	1	1	2	2	2
74	Yugo	4	3	4	1	4	2	2	4	2	4	4	1	2	2	2	3	3	3	1	1	2	2	4	4	1	4	3	4	4	4	4	1	3	1	3	3	1	2	2	2

TABLE 12
Humana: Results of Traditional and Proposed Group-Difference Tests

	Statistic		Significance Levels	
			Traditional	Proposed
Group-Difference				
(1) common behavior				
F	220.301	*df* = 2, 1153*	less than 0.001	less than 0.001
Kruskal-Wallis	152.956		less than 0.001	less than 0.001
Jonckheere-Terpstra	81953		less than 0.001	0.035
$1 < 2 < 3$				
(2) shared rights				
F	756.189	*df* = 2, 1153	less than 0.001	less than 0.001
Kruskal-Wallis	206.595		less than 0.001	less than 0.001
Jonckheere-Terpstra	123113		less than 0.001	0.059
$1 < 2 < 3$				
(3) shared military rights				
F	86.735	*df* = 2, 1153	less than 0.001	0.003
Kruskal-Wallis	121.849		less than 0.001	0.010
Jonckheere-Terpstra	80097		0.323	0.005
$1 < 2 < 3$				
(4) shared free speech rights				
F	425.604	*df* = 2, 1153	less than 0.001	less than 0.001
Kruskal-Wallis	291.280		less than 0.001	less than 0.001
Jonckheere-Terpstra	134071		less than 0.001	0.001
$1 < 2 < 3$				
(5) strong rights				
group-difference			no standard test	less than 0.001
ordered-alternative			no standard test	less than 0.001

*Degrees of freedom for the denominator is calculated as $44*43/2 + 15*14/2 + 15*14/2 - 3 = 1153$.

DISCUSSION

Previous methods for determining whether D-groups show significant differences in some dependent variable include, among other methods, inspection, means-difference tests, Bradley-Terry paired comparison designs (Bradley, 1984), and the quadratic assignment procedure (QAP) (see Hubert, 1987, or for an earlier treatment of the combinatorial assignment strategy which underlies QAP, see also Hubert and Baker, 1978). Unfortunately, none of these is entirely satisfactory, although QAP comes the closest (as will be discussed in more detail). Inspection offers no measure of significance. Means-difference tests (analysis of variance, Kruskal-Wallis) ignore dependencies arising from the fact that each actor influences observations from many pairs. Bradley-Terry designs use binary ranking data, make restrictive assumptions in modeling D-group effect, and require asymptotic theory for inference; this limits their range of application.

Kraemer and Jacklin (1979) develop a normal theory model for examining dyadic data which consist of pairs of measurements for each dyad, one for each partner. This differs from the framework we consider, which is nonparametric and allows for the possibility that only a single measurement is obtained on each dyad observed, although all possible pairs of actors are observed. Inference in their framework requires approximately normally distributed data; probably the most delicate assumption is that dyads formed by partners from different A-groups have variances that are simple functions of the A-group characteristics. The implementation of the analysis can become cumbersome when there are many A-groups, and one must do substantial additional work to make inferences about significance of average differences between dyads according to the partners' A-group memberships.

Kenny and La Voie (1984) also consider models with two observations on each member of the dyad. Their analysis parallels normal theory methods, but uses the jackknife procedure (cf. Efron (1982)) to obtain asymptotically nonparametric inferences. The work is similar in spirit to much of Bradley's research on round-robin tournament designs (reviewed in Bradley (1984)); it develops models with actor, partner, relationship, and error terms. In order to assess the effect of dyadic groupings, one must extend the model to include terms corresponding to the D-group membership. As with much previous work, there exists no immediate test for significant differences between measurements from dyads belonging to different D-groups.

Wasserman and Iacobucci (1986) and Iacobucci and Wasserman (1987) describe log-linear models for dyadic interactions of all possible pairs in a social group. The methods are parametric, and correspond to maximum likelihood inference when the data arise from multinomial, Poisson or product-multinomial distributions. Those models are focused on determining A-group differences, whereas the proposed methods focus on D-group differences. Nevertheless, as Iacobucci and Wasserman (1987, p. 220) note, one can use their models for examining D-group differences by computing the intra-class correlation coefficients and the product-moment correlation coefficient for the inter-class dyad. As the number of levels of the relational (and in the parlance of this paper, dependent) variable increases there is a corresponding increase in the number of parameters estimated. Consequently when the range of the dependent variable is large, the researcher may be forced to collapse the data into coarser units of measurement.

In summary, previous research in this area has focused on cases in which (1) for the dependent variable there are exactly two measurements per dyad, (2) specific model assumptions are made about the data, or (3) the primary interest of the researcher is in predicting behavior from A-groups rather than D-groups, which can necessitate a more complex analysis that attempts to fit multi-way interactions. In contrast to the first point, our analysis addresses the case in which the dependent variable has either one (e.g., actual time spent interacting) or two measures (e.g., perception of time spent interacting) per pair of actors. In contrast to the second point, our approach is classically nonparametric; this enables straightforward implementation when traditional statistical models are inappropriate. The issue underlying the third point is more complex than the first two. The difference is, in part, in what type of groups are being focused on. But an analogy may clarify the issue. Our method is analogous to one-way analysis of variance, whereas the methods proposed by others such as Wasserman and Iacobucci (1986) and, even more so, that proposed by Kenny and La Voie (1984) are analogous to multi-way analysis of variance. (Submodels of both Wasserman and Iacobucci and Kenny and La Voie exist, when one of the parameters is set equal to zero, that are analogous to one-way analysis of variance.) Such a simplification, both as present in the proposed method and as present in the submodels of Wasserman and Iacobucci and Kenny and La Voie, sometimes permits a more straightforward analysis, and for many questions is all that is needed.

Our work in this area suggests the following taxonomy for a large class of network-data problems, based on the nature of the two matrices for which one wants to access correspondence.

Numerical-Numerical: One has two matrices of network data, each of which contains numerical information, and one wants a measure of association that indicates how strongly the entry in one cell predicts the corresponding entry in the other matrix. This is the case for which QAP is designed.

Categorical-Numerical: One matrix contains categorical information on pairs of actors; this represents D-group membership. The other matrix contains numerical information on each pair. The researcher wants to discover the extent to which D-group membership accounts for differences in the numerical data. This is the case for which our procedures are designed. QAP is not appropriate when the explanatory variable (group membership) is categorical.

Categorical-Categorical: In this case, one has two matrices of categorical data, and one wants to determine whether categories in one matrix tend to correspond to categories in the other. There does not seem to be a full treatment of this problem that honors the network dependencies in the data. The closest work here is that by Hubert and Baker (1978) who use bonary data and the group-difference procedure we have outlined when applied to binary data. For binary data, as was the case with the Strong Rights index built from Humana's data, we can use our procedure for categorical-numerical data to treat what is really a categorical-numerical data to treat what is really a categorical-categorical situation. However, as soon as the number of categories in the dependent-variable matrix exceeds two, this generalization fails. Future research should aim at addressing the categorical-categorical case.

When the matrices for both the dependent and independent variable contain only binary data procedures for numerical-numerical, categorical-numerical, and categorical-categorical essentially are equally applicable. More importantly, when there are only two D-groups, a non-asymptotic version of QAP using dummy variables to code for D-group membership is equivalent exactly to the method we describe. Further, when there are only two D-groups and the independent matrix is also binary then Hubert and Baker's (1978) procedure (which is really a precursor to QAP) is equivalent to the group-difference test we have proposed.

These equivalencies fail, however, when there are more than two D-groups. If there are more than two D-groups, neither QAP (nor its predecessor) enable group-difference tests. QAP measures the strength of association between two matrices; whereas, our procedure involves measuring group separation.[7] For the group-

[7] Essentially QAP measure the strength of association between two matrices by referring a normalized measure of association to a z-score. The normalization is obtained by an asymptotic approximation to a null distribution in which any rearrangement of the rows and columns is equally likely. The QAP strategy, however, is broader than any particular implementation. One could vary the measure of association and replace the asymptotic approximation by Monte Carlo simulation without substantively changing

difference test we proposed, when there are two D-groups, strong separation occurs when A-group labels distinguish large data values from small ones in the dependent variable data matrix. Similarly, in QAP, the D-group matrix and the dependent variable matrix are most strongly associated when one group label picks out large values on the dependent variable and the other group label flags small values. Our procedure and QAP essentially are equivalent, when there are only two D-groups, as in both procedures the null distribution is based on row-column permutation.

When there are three or more groups, measures of association (as used by QAP) and separation (as we use) respond to different features of the data. With QAP, changing the group labels will affect the conclusion strongly. If the group labels and corresponding entries in the data matrix tend to increase together, QAP detects strong association. QAP finds less evidence of association if the group labels are shuffled so that the data trend is not monotonic with respect to the group labels. In our method, separation doesn't depend upon codes for the group labels; we can renumber the groups and obtain exactly the same result. Therefore our procedure and QAP address different problems, but the problems happen to coincide when there are exactly two groups.

As a final point, let us consider the ordered-alternative test. Our procedure imposes a strict standard for rejecting the null. Unless the data strongly support exactly the prespecified ordering, say $1 < 2 < 3 < 4 < 5 < 6$, then our test cannot reject the null hypothesis. This contrasts with less stringent tests that depend on association, such as those used in QAP. For example, if the true population ordering were $1 < 2 < 4 < 3 < 5 < 6$, QAP ultimately would detect the generally concordant trend in the data, declare a significant result, but overlook the discrepant cases in groups 3 and 4.

CONCLUSION

We have presented two very general nonparametric tests for network data that complement each other. The first of these, the group-difference test, enables the researcher to determine whether group membership is a reasonable predictor of other shared traits. The second test, the ordered-alternative test, allows the researcher to determine whether a particular ordering of D-groups corresponds to a particular ordering of some measure of centrality for another shared trait. Both tests take into account of row-column dependencies and are nonparametric. We show by several examples that group-difference tests that do not respect the dependencies in the data tend not to be conservative, and exaggerate the significance of their conclusions. For ordered-alternative tests, depending on the particular dataset involved, taking the dependencies into account may move the significance probabilities in either direction from those obtained with a more traditional analysis. Nonetheless,

---

the idea (Hubert, 1985). In contrast to the QAP strategy, we observe a measure of group separation, and determine the null distribution by permuting the rows and columns of the dependent variable data matrix. The specific implementation used in this paper employs an F statistic and Kruskal-Wallis statistic to capture separation. Also we prefer Monte Carlo methods as they are more accurate for small samples. Exactly which test statistics are used and whether an asymptotic approximation or Monte Carlo technique is used are minor matters. They do not speak to the fundamental difference between QAP and our procedure, which is the focus on association as opposed to separation.

our method is more accurate, in that it properly incorporates known dependencies in the data. The difference in the significance probabilities calculated by traditional methods and the permutation methods we propose can be quite large. This is especially true for the ordered-alternative test. Moreover, these tests, depending on the data and the significance level used can change the interpretation of the results from significant to non-significant (or vice-versa in the Jonckheere-Terpstra test).

Both tests can be used in a wide range of applications. However, proper use of the ordered-alternative test requires one to prespecify a hypothesized ordering of the group means (or medians), and then the test evaluates the extent to which the data actually conforms to that ordering. When such an ordering cannot be prespecified, the group-difference test still permits discovery of group effects. Regardless of the test, if the significance probability is small, then the data corroborates the researcher's hypothesis. These tests provide social scientists with tools for doing D-group comparisons on network data in an analytic rather than observational fashion. To use these tools the researcher must have some basis for identifying D-group membership using data other than the dependent variable. These tools increase the researcher's ability to compare D-groups and determine whether the extracted structure, or set of D-groups, can predict other variables. The group-difference test and the ordered-alternative test allow the researcher to determine the explanatory value of a particular set of D-groups, and even to compare the relative value of two different ways of grouping the data.

REFERENCES

Banks, D. L. (1989) Patterns of oppression: An exploratory analysis of human rights data, *Journal of the American Statistical Association* **84:** 674–681.

Blau, P. M. (1977) *Inequality and Heterogeneity*. New York, NY: The Free Press of Macmillan Co.

Bradley, R. A. (1984) Paired comparisons: Some basic procedures and examples. In P. R. Krishnaiah and P. K. Sen (Eds.), *Handbook of Statistics* **4**. Amsterdam: North-Holland.

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

Burt, R. S. (1980) Innovation as a structural interest: Rethinking the impact of network position innovation adoption, *Social Networks* **4:** 337–355.

Burt, R. S. (1982) *Toward a Structural Theory of Action*. New York, NY: Academic Press.

Burt, R. (1989) *Structure* (Tech. Rep.). Columbia University's Center for the Social Sciences, New York, NY.

Carley, K. M. (1984) *Constructing Consensus*. Doctoral dissertation, Harvard.

Carley, K. M. (1986) An approach for relating social structure to cognitive structure, *Journal of Mathematical Sociology* **12(2):** 137–189.

Carley, K. M. (1989) Integration and isolation: The role of structure andinteraction in the acquisition of ideas. Earlier version presented at Sunbelt Social Network Meeting—1987.

Carley, K. M. (1990) Structural constraints on communication: The diffusion of the homomorphic signal analysis technique through scientific fields, *Journal of Mathematical Sociology* **15(3–4):** 207–246.

Dunnet, C. W. (1964) New tables for multiple comparisons with a control, *Biometrics* **20:** 482–491.

Efron, B. (1982) *The Jackknife, the Bootstrap and Other Resampling Methods*. Philadelphia, PA: Siam Monograph No. 38, CBMS-NSF.

Fararo, T. J. and Skvoretz, J. (1986) E-state structuralism, *American Sociological Review* **51:** 591–602.

Friedman, M. (1937), The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *Journal of the American Statistical Association* **32:** 675–701.

Granovetter, M. S. (1973) The strength of weak ties, *American Journal of Sociology* **68:** 1360–1380.

Granovetter, M. S. (1974) *Getting a Job: A Study of Contacts and Careers*. Cambridge, MA: Harvard University Press.

Heil, G. H. and White, H. C. (1976) An algorithm form finding simultaneous homomorphic correspondences between graphs and their image graphs, *Behavioral Science* **21:** 26–35.

Hollander, M. and Wolfe, D. A. (1973) *Nonparametric Statistical Methods*. New York, NY: John Wiley and Sons, Inc.

Homans, G. C. (1950) *The Human Group*. New York, NY: Harcourt, Brace and Co.

Hubert, L. J. (1985) Combinatorial data analysis: Association and partial association, *Psychometrika* **50:** 449–467.

Hubert, L. J. (1987) *Assignment Methods in Combinatorial Data Analysis*. New York, NY: Dekker.

Hubert, L. J. and Baker, F. B. (1978) Evaluating the conformity of sociometric measurements, *Psychometrika* **43(1):** 31–41.

Humana, C. (1983) *World Human Rights Guide*. London, England: Hutchinson.

Iacobucci, D. and Wasserman, S. (1987) Dyadic social interactions, *Psychological Bulletin* **102(2):** 293–306.

Kenny, D. A. and La Voie, L. (1984) The social relations model. In L. Berkowitz V18 (Eds.), *Advances in Experimental Social Psychology*. New York, NY: Academic Press.

Kraemer, H. C. and Jacklin, C. N. (1979) Statistical analysis of dyadic social behavior, *Psychological Bulletin* **86:** 217–224.

Lorrain, F. and White, H. C. (1971) Structural equivalence of individuals in social networks, *Journal of Mathematical Sociology* **1:** 49–80.

Newcomb, T. (1961) *The Acquaintance Process*. New York, NY: Holt, Reinhart, and Winston.

Robertson, T., Wright, F. T., and Dykstra, R. L. (1988) *Order Restricted Statistical Inference*. New York, NY: Wiley.

Sampson, S. F. (1968) A novitiate in a period of change: An experimental and case study of relationships. Doctoral dissertation, Cornell University.

Scheffe, H. (1959) *The Analysis of Variance*. New York, NY: Wiley.

Wasserman, S. and Iacobucci, D. (1986) Statistical analysis of discrete relational data, *British Journal of Mathematical and Statistical Psychology* **39:** 41–64.

White, H. C., Boorman, S. A., and Breiger, R. L. (1976) Social structure from multiple networks. I. Blockmodels of roles and positions, *American Journal of Sociology* **81:** 730–780.