QAP PARTIALLING AS A TEST OF SPURIOUSNESS *

David KRACKHARDT

Johnson Graduate School of Management, Cornell University

A test of spuriousness for structural data is proposed. Partials are calculated using OLS estimates. The test of significance is based on Hubert's QAP, a nonparametric permutation test. Results of Monte Carlo simulations indicate that statistical bias and efficiency characteristics of this procedure are very reasonable.

1. Introduction

Consider the following problem. You are interested in testing a theory that suggests that people seek out their personal friends for work-related advice and help. You dutifully collect the appropriate network data and discover, to your relief (if not amazement) that, indeed, a significant correlation exists between the friendship network and the advice network. Now, suppose an organizations scholar reviews your hard work and claims that your observed correlation is simply due to the fact that proximity in the organization creates opportunities for both friendship and advice-relationships. That is, people do not seek out their friends per se; proximity constrains their friendship choices just as it constrains their advice choices. The observed correlation, concludes your organizational colleague, is spurious.

In reflection, one can see the potential for spurious correlations in most empirical work in network analysis. Networks are frequently multiplex in nature and are affected by attributes of the actors. One may wonder, then, why few have attempted to uncover key relationships by statistically controlling for such spuriousness. But to test for

such spuriousness is particularly difficult, given the inherently interdependent nature of dyadic data in social networks.

This paper evaluates an approach proposed earlier (Krackhardt 1984; Krackhardt and Breiger 1985) to addressing the question of whether two structural variables are spuriously correlated with each other. This approach combines both traditional least-squares estimating techniques with Hubert's nonparametric test, the quadratic assignment procedure (QAP) [1]. Our intent in this paper is twofold. First, we present the mechanics of the procedure in some detail. Second, we test the procedure's robustness and statistical performance under varying population conditions using a series of Monte Carlo simulations.

2. The proposed test of spuriousness

2.1. Calculation of partial correlation in structural variables

We will assume a dependent variable (Y), an independent variable (X), and a set of possible sources of spuriousness $(Z_1, Z_2, \ldots)$. Moreover, we assume these variables are in a traditional structural form as follows: [2]

$$
Y = \begin{pmatrix}
0 & y_{1,2} & \cdots & y_{1,n-1} & y_{1,n} \\
y_{2,1} & 0 & \cdots & y_{2,n-1} & y_{2,n} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
y_{n,1} & y_{n,2} & \cdots & y_{n,n-1} & 0
\end{pmatrix},
$$

[1] It is worth noting that Hubert (1985) has independently developed a separate method of partial associations using the QAP. In addition, three-variable analyses have been proposed by Hubert and Golledge (1981), and applied by Nakao and Romney (1984) and Dow and Cheverud (1985), although their aim was not partialling, per se. Most recently, Sokal, Smouse and Neel (1986, p. 282) have referred to unpublished techniques that appear to be comparable to those proposed here. However, in none of these studies has the statistical bias of these multivariate methods been systematically explored either analytically or through simulations.

[2] Throughout this paper, we will assume that reflexive relations are not defined. Consequently, diagonal elements are ignored. However, the analytical scheme presented here is not dependent on this assumption; partials and their associated tests can be calculated with the diagonals included, if they are meaningful to the researcher.

$$X = \begin{pmatrix} 0 & x_{1,2} & \cdots & x_{1,n-1} & x_{1,n} \\ x_{2,1} & 0 & \cdots & x_{2,n-1} & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,n-1} & 0 \end{pmatrix},$$

$$Z_1 = \begin{pmatrix} 0 & z1_{1,2} & \cdots & z1_{1,n-1} & z1_{1,n} \\ z1_{2,1} & 0 & \cdots & z1_{2,n-1} & z1_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z1_{n,1} & z1_{n,2} & \cdots & z1_{n,n-1} & 0 \end{pmatrix},$$

$$Z_2 = \begin{pmatrix} 0 & z2_{1,2} & \cdots & z2_{1,n-1} & z2_{1,n} \\ z2_{2,1} & 0 & \cdots & z2_{2,n-1} & z2_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z2_{n,1} & z2_{n,2} & \cdots & z2_{n,n-1} & 0 \end{pmatrix},$$

and so on.

We further assume that the model of interest here is that X and Y are theoretically related, but that $Z_1, Z_2,\ldots$ are potential sources of spuriousness to be controlled for. One way to control for Z_1, $Z_2,\ldots$ is to extract the variance from X and Y due to $Z_1, Z_2,\ldots$, and then correlate the residuals of X with the residuals of Y. This is equivalent, in ordinary linear analysis, to taking the partial correlation between X and Y controlling for Z_1, $Z_2,\ldots$ (Kenny 1979).

To specify how the residuals are calculated, it will be convenient first to transform the above structural representations into vector form:

$$Y = \begin{pmatrix} y_{1,2} \\ y_{1,3} \\ \vdots \\ y_{n,n-1} \end{pmatrix}, \quad X = \begin{pmatrix} x_{1,2} \\ x_{1,3} \\ \vdots \\ x_{n,n-1} \end{pmatrix}, \quad Z_1 = \begin{pmatrix} z1_{1,2} \\ z1_{1,3} \\ \vdots \\ z1_{n,n-1} \end{pmatrix}, \quad Z_2 = \begin{pmatrix} z2_{1,2} \\ z2_{1,3} \\ \vdots \\ z2_{n,n-1} \end{pmatrix}.$$

Finally, in keeping with traditional forms of OLS analysis, define $\tilde{Z}$ for k controlling variables as follows:

$$
\tilde{Z} = \begin{pmatrix}
1 & z_{1_{1,2}} & z_{2_{1,2}} & \cdots & z_{k_{1,2}} \\
1 & z_{1_{1,3}} & z_{2_{1,3}} & \cdots & z_{k_{1,3}} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & z_{1_{n,n-1}} & z_{2_{n,n-1}} & \cdots & z_{k_{n,n-1}}
\end{pmatrix}.
$$

Using the fundamental OLS solution, the residuals for X and Y are then calculated as follows:

$$
Y^* = \begin{pmatrix}
y_{1,2}^* \\
y_{1,3}^* \\
\vdots \\
y_{n,n-1}^*
\end{pmatrix} = Y - \tilde{Z}\beta = Y - \tilde{Z}\big((\tilde{Z}'\tilde{Z})^{-1}\tilde{Z}'Y\big),
$$

$$
X^* = \begin{pmatrix}
x_{1,2}^* \\
x_{1,3}^* \\
\vdots \\
x_{n,n-1}^*
\end{pmatrix} = X - \tilde{Z}\beta = X - \tilde{Z}\big((\tilde{Z}'\tilde{Z})^{-1}\tilde{Z}'X\big).
$$

Correlating vectors Y^* and X^* would give us the partial correlation we are seeking. All that is necessary at this point is an appropriate inferential test to ascertain whether this correlation is significant. Unfortunately, the traditional OLS approach *cannot* be used for such a test (see Proctor (1969) and Laumann and Pappi (1976: 150) for a discussion of the problems of significance testing of structural data). The major difficulty lies in the fact that OLS inference tests assume that the observations being tested are independent of one another. In any structural analysis of N actors, one has $N(N-1)$ dyadic observations that are usually assumed to be quite *interdependent*, rendering the OLS approach inapplicable.

As an alternative, it has been proposed that the quadratic assignment procedure (QAP) be used to test the significance of an observed correlation. The QAP is a nonparametric, permutation-based test that preserves the integrity of the observed structures (i.e., explicitly retains the interdependency among the dyads). In final preparation for this test, then, the elements in the residual vectors Y^* and X^* must be rearranged into their structural form:

$$Y^* = \begin{pmatrix} 0 & y^*_{1,2} & \cdots & y^*_{1,n-1} & y^*_{1,n} \\ y^*_{2,1} & 0 & \cdots & y^*_{2,n-1} & y^*_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y^*_{n,1} & y^*_{n,2} & \cdots & y^*_{n,n-1} & 0 \end{pmatrix},$$

$$X^* = \begin{pmatrix} 0 & x^*_{1,2} & \cdots & x^*_{1,n-1} & x^*_{1,n} \\ x^*_{2,1} & 0 & \cdots & x^*_{2,n-1} & x^*_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x^*_{n,1} & x^*_{n,2} & \cdots & x^*_{n,n-1} & 0 \end{pmatrix}.$$

At this juncture, we will digress to a discussion of QAP as a hypothesis test.

2.2. The QAP hypothesis test

A growing literature has focused attention on the QAP as a general test to answer the question of whether two structures are significantly related to each other (e.g., Mantel 1967; Hubert and Schultz 1976; Sokal 1979; Baker and Hubert 1981; Hubert 1983; Dietz, 1983). As an illustration of how QAP works, recall the research question that opened this paper. The researcher was interested in discovering the relationship between friendship structures and advice structures. For example, suppose you have a set of five actors for which you have measured both friendship choices and advice choices. Figure 1 depicts the results of our measures in graph form (assume both relations are symmetric). The question is, are these structural patterns represented in Figure 1 similar to each other? One might be tempted to say that they look similar, as they differ only by a line connecting the bottom two points in Figure
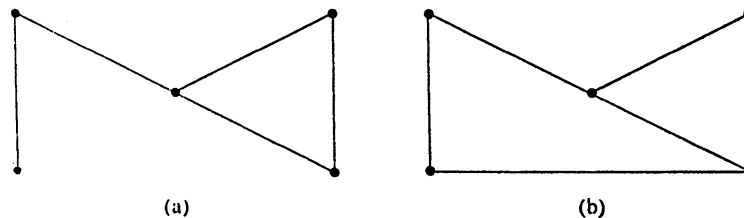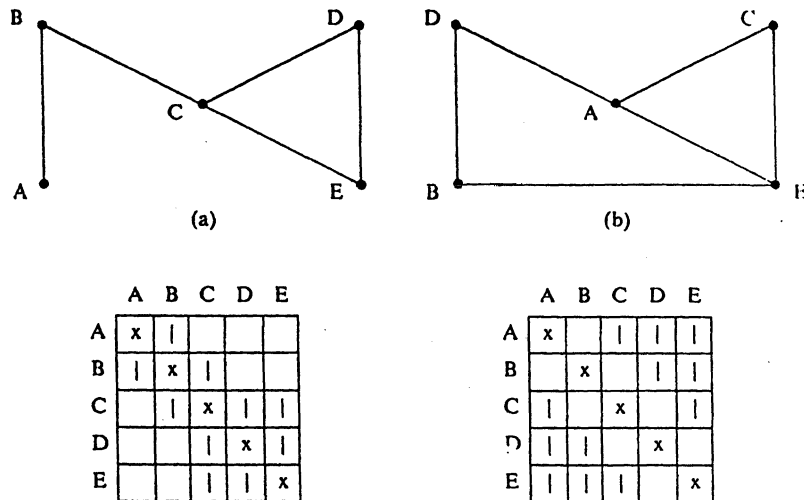

(a)   (b)

Figure 1. Sociograms for two relations: (a) friendship, (b) advice.

Friendship matrix (a):

	A	B	C	D	E
A	x	l			
B	l	x	l		
C		l	x	l	l
D			l	x	l
E			l	l	x

Advice matrix (b):

	A	B	C	D	E
A	x		l	l	l
B		x		l	l
C	l		x		l
D	l	l		x	
E	l	l	l		x

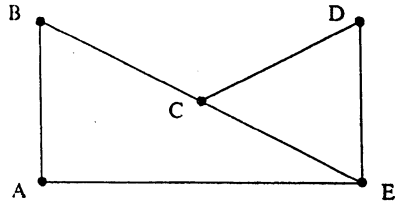r = −0.8165

Figure 2. Assignment of actors and nodes in sociograms: (a) friendship, (b) advice (permutation 1).

1b. With thought, it is apparent that, while they look fairly similar, one cannot answer this question without knowing which node in Figure 1a is associated with (maps onto) which node in Figure 1b.

Figure 2 shows one possible mapping, where the letters A through E are used to identify the actors in our hypothetical study. Below each graph is the corresponding matrix representation of the graph. To measure the degree to which these two structures are similar, we simply transform the matrices into vector form (Y and X) and correlate the two vectors. The result is −0.8165, indicating a relative dissimilarity between the two structures (i.e., one tends to seek advice from those who are not friends).

Suppose that our labels remain fixed for the "Friends" graph, and we permute the assignment of labels on the "Advice" graph. There are 120 ($= N!$) possible mappings of the five actors to the five points on the graph. The permutation in Figure 3 results in a positive correlation between the two matrices ($r = 0.8165$). This case would provide stronger evidence for the theory. A third permutation, in Figure 4, results in a correlation with friendship choices of $r = 0.0$, i.e., the seeking of advice is independent of those whom one chooses as friends.

	A	B	C	D	E
A	x	I			I
B	I	x	I		
C		I	x	I	I
D			I	x	I
E	I		I	I	x

r = 0.8165

Figure 3. Permutation two of advice relation.
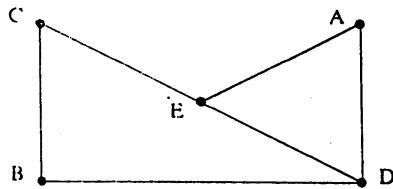
	A	B	C	D	E
A	x			I	I
B		x	I	I	
C		I	x		I
D	I	I		x	I
E	I		I	I	x

$r = 0$

Figure 4. Permutation three of advice relation.

Table 1
Distribution of all 120 permutations

	Measure of similarity (r)	Number of permutations	Percent
	0.8165	4	3.3
	0.4082	32	26.7
	0.0000	48	40.0
	−0.4082	32	26.7
	−0.8165	4	3.3
Weighted Average	0.0	Total: 120	100.0

Table 1 summarizes the results of all 120 possible permutations (mappings) of actors to nodes in the "Advice" structure. Four of the possible permutations (3.3% of the total) resulted in the strongest correlation with friendship choices ($r = 0.8165$), 32 (26.7%) resulted in a moderate correlation (0.4082), 48 (40%) resulted in a zero correlation, and the remaining 30% showed a negative correlation. If our observed data were those represented in the permutation in Figure 3, then we would record that our observed results had only a 3.3% chance of occurring by random assignment of actors to nodes. Our conclusion would be that our observed results are significant at the 0.033 level against the null hypothesis of random assignment.

Calculating all 120 possible permutations is tractable in this hypothetical problem. However, the number of possible permutations in larger networks becomes unmanageable very quickly. Fortunately, an analytical solution to the problem has been developed (Mantel 1967; Hubert and Schultz 1976). Given any two $N \times N$ matrices, the exact mean and standard deviation of the distribution of correlations between the one matrix and all possible permutations of the other is a straightforward function of the cell values in the matrices. The observed correlation, then, can be expressed as a standardized Z-score. Assuming an approximately normal distribution of correlations under all permutations, the significance level of the observed correlation then simply is determined by the appropriate area beyond the Z-score under the normal curve. [3]

[3] The assumption of normality in the distribution of the r under all permutations has been the subject of considerable attention in the literature (Mielke 1979; Costanzo, Hubert and Golledge 1983; Faust and Romney 1985). Deviations from normality can be severe when the data themselves are badly skewed. In the present case, however, we will concern ourselves with observations that do not present this problem.

2.3. QAP as a test of partials

This same procedure for testing the significance of a pair of structures can be used to test nonparametrically the significance of a partial correlation. Recall that the correlation between the vector forms of Y^* and X^* is the partial correlation holding $Z_1, Z_2, \ldots$ constant. By repeated permutations of the rows and columns of one of the two matrices, say X^*, followed by recalculation of the correlation between them after each permutation, we can generate a distribution of values of partial r's against which the observed r can be compared and from which a significance level can be inferred. Thus, while the descriptive statistic is derived from traditional OLS calculations, the inferential probability is derived nonparametrically. If the QAP test of similarity between Y^* and X^* yields a significantly small probability, then one would conclude that the original correlation between them is not spurious due to that set of control variables, $Z_1, Z_2, \ldots$ Conversely, if the QAP test indicates an insignificant correlation, then one would conclude that the original observed correlation between Y and X is spurious due to $Z_1, Z_2, \ldots$

3. Testing the QAP partialling procedure: A Monte Carlo simulation

A serious question of potential statistical bias arises from this proposed procedure (cf., Hubert and Golledge 1981: 220; Krackhardt in press). Is the technique positively biased, that is, is it too "easy" to get a significant result, such that completely random data could appear to be significant more often than 5% of the time (assuming α is selected to be 0.05)? Or, perhaps, it is negatively biased, such that random data would appear significant less than 5% of the time. Krackhardt (in press) demonstrated, for example, that the QAP test, under some restricted conditions, can be extremely positively or negatively biased.

To assess this possible bias, a series of Monte Carlo simulations was conducted. By generating random data with known parameters, one can determine how sensitive this test of spuriousness is. For example, under the null hypothesis that X and Y are uncorrelated by themselves in the population but are spuriously correlated through Z, we should observe a significant ($\alpha = 0.05$) partial correlation in only 5% of the

samples drawn from that population. This should be true no matter how strong the correlation is between Z and either X or Y.

This is the minimum we should expect of any test of spuriousness. We can ask more of the test: it should also minimize Type II errors. In a population where X and Y are only slightly correlated (say, $\rho = 0.1$), we would find "significant" simple correlations more frequently than in 5% of the samples, with the exact percentage (call it $\breve{p}$) depending on the sample size and power of the test. If Z is introduced into such populations, such that the observed simple correlation between X and Y is increased, then we would prefer our test of spuriousness to result in a partial correlation that is significant *in no less than $\breve{p}$* percent of the samples drawn from the population. That is, the test should be unbiased, so that under the null hypothesis only about α proportion of the samples appears significant; yet, at the same time, it should be powerful enough to detect underlying correlations in populations where they do occur.

Both of these criteria, statistical bias and power, were assessed by drawing samples from the following population model:

$$Z = \epsilon_1,$$

$$X = k_z Z + \epsilon_2,$$

$$Y = k_x X + k_z Z + \epsilon_3,$$

where:

$$Z = \{Z_{ij}\}, \quad i = 1, 2, \ldots, 10, \quad j = 1, 2, \ldots, 10, \quad i \neq j,$$

$$X = \{X_{ij}\}, \quad i = 1, 2, \ldots, 10, \quad j = 1, 2, \ldots, 10, \quad i \neq j,$$

$$Y = \{Y_{ij}\}, \quad i = 1, 2, \ldots, 10, \quad j = 1, 2, \ldots, 10, \quad i \neq j,$$

$$\text{each } \epsilon = \{\epsilon_{ij}\}, \quad i = 1, 2, \ldots, 10, \quad j = 1, 2, \ldots, 10, \quad i \neq j,$$

and the values of ϵ_1, ϵ_2, and ϵ_3 are all independently and uniformly distributed from 0 to 1.

Data were randomly generated from this model. The size of "pure" correlations among Z, X and Y in the population can be derived from

the k-weights, as demonstrated below:

Given:

$$y = kx + \epsilon,$$

$$\text{cov}(x, \epsilon) = 0,$$

$$\text{var}(x) = \text{var}(\epsilon),$$

then:

$$\rho_{x,\ y} = \frac{\text{cov}(x,\ y)}{\sqrt{\text{var}(x) \cdot \text{var}(y)}},$$

$$\text{cov}(x,\ y) = \text{cov}(x,\ kx + \epsilon) = k \cdot \text{var}(x),$$

$$\text{var}(y) = \text{var}(kx + \epsilon) = k^2 \cdot \text{var}(x) + \text{var}(\epsilon) = \text{var}(x) \cdot (k^2 + 1),$$

$$\rho_{x,\ y} = \frac{k \cdot \text{var}(x)}{\sqrt{\text{var}(x)(\text{var}(x) \cdot (k^2 + 1))}} = \frac{k}{\sqrt{k^2 + 1}},$$

$$k = \frac{\rho_{x,\ y}}{\sqrt{1 - \rho_{x,\ y}^2}}.$$

The k-weights were chosen such that the "pure" correlations [4] between the matrices varied in discrete steps in the populations as follows:

k-weight	ρ
0.0000	0.0
0.1005	0.1
0.2041	0.2
0.3145	0.3
0.4364	0.4
0.5774	0.5
0.7500	0.6
0.9802	0.7
1.3333	0.8
2.0647	0.9

[4] By "pure" correlation we mean the correlation between the two variables that results if we exclude all other sources of systematic covariation. Thus, assuming $k_z = 0$, k_x will give rise to "pure" values of $\rho_{X,Y}$. The actual population value of $\rho_{X,Y}$, of course, will be a function of the combined "pure" contributions of k_z *and* k_x.

Table 2
Percent significant: $\rho_{x,y} = 0$.

$\rho_{z,x}$	Proportion of 1,000 samples where significant correlation was found
0	0.047
0.1	0.044
0.2	0.049
0.3	˛ 0.061
0.4	0.056
0.5	0.054
0.6	0.051
0.7	0.045
0.8	0.047
0.9	0.044

Thus, k_z took on ten values that resulted in population correlations between Z and X of 0, 0.1, 0.2, ... 0.9; k_x took on six values (higher values were not necessary, as will be demonstrated later) that resulted in "pure" population correlations between X and Y of 0, 0.1, ... 0.5. By combining these values of k_x and k_z, 60 populations were defined. One thousand samples were drawn from each of these populations.

Samples from the null hypothesis model are drawn when $k_x = 0$, that is, when there is no "pure" contribution of X to Y and any observed correlation is spurious through Z. If the QAP partialling test of spuriousness is unbiased, then about 5% of the samples in each of the ten populations where $k_x = 0$ should appear significant. The results of this test are reported in Table 2.

As can be seen from Table 2, very little bias is evident from this procedure. The proportions of "significant" results ranged from 0.044 to 0.061. Thus, independent of how much variance is contributed by Z, approximately 5% of the 1,000 samples in each population appeared significant at the 0.05 level.

The power of QAP partialling is demonstrated by comparing the proportion of "significant" simple correlations for varying levels of $\rho_{X,Y}$ with the proportions found when Z is contributing spuriously and then partialled out. The results of this comparison are provided in Table 3.

As evident in Table 3, the QAP partial sacrifices virtually no statistical power in the process of uncovering the "truth" about the underlying relationship between X and Y. For example, when the

Table 3
Proportion of 1.000 samples where significant correlation was found.

| | | "True" ρ_{xy} | | | | | |
		0	0.1	0.2	0.3	0.4	0.5
No Z in population:		0.047	0.223	0.588	0.903	0.992	1.00
Strength of	0	*	0.229	0.585	0.901	0.990	1.00
contribution	0.1		0.266	0.615	0.908	0.996	1.00
of Z	0.2		0.246	0.638	0.906	0.995	1.00
$(= \rho_{z,x} = \rho_{z,y})$	0.3		0.232	0.587	0.907	0.992	1.00
	0.4		0.254	0.601	0.901	0.994	1.00
	0.5		0.242	0.604	0.895	0.987	1.00
	0.6		0.258	0.600	0.891	0.992	1.00
	0.7		0.246	0.572	0.906	0.992	1.00
	0.8		0.233	0.595	0.899	0.992	1.00
	0.9		0.226	0.583	0.898	0.991	1.00

* These proportions in the first column can be found in Table 1.

population is characterized by a weak relationship between X and Y ($\rho_{X,Y} = 0.1$), and there is no presence of a confounding Z, then 223 out of a thousand samples are significant at the 0.05 level. When Z is introduced at increasing strengths ($\rho_{Z,X} = \rho_{Z,Y} = 0, 0.1,\ldots,0.9$), the probability of finding a significant partial correlation varies little from this base (range = 0.226 to 0.266). When the underlying correlation between X and Y in the population is 0.2 (with no confounding Z), the simple correlations are significant in almost 60% of the samples; the range in probability that the partials are significant is 0.572 to 0.638. As $\rho_{X,Y}$ is increased beyond 0.2, the range of significant partials becomes even narrower, converging on the proportion found significant in the simple, unconfounded population.

To summarize, the QAP partialling test of spuriousness proposed here has been demonstrated to have two desirable properties. First, in the case of these simple populations of size 10 × 10, the probability of finding a significant correlation when none exists is approximately $\alpha = 0.05$. Second, when a true correlation does exist in the population, the partialling procedure identifies it as significant with approximately the same probability as the simple correlation test. This leads us to the conclusion that the QAP partial test is both relatively unbiased – in the narrow statistical sense that α equals the true probability of committing a Type I error – and as powerful as the simple QAP test.

3.1. Limitations of simulation results

These results suggest that the QAP partialling procedure is robust and appropriate under the conditions described by the model presented. However, it would be a mistake to assume that the partialling technique would behave so admirably under all other models as well. In particular, the model used here is characterized by a uniform distribution of values. If a normal distribution were used, the occasional "extreme" value observed in samples might result in more instability in the QAP test. Or, perhaps highly skewed distributions would result in unstable or otherwise perverse conclusions (see Faust and Romney 1985).

Another issue is that the QAP test is nonparametric. It is based on a permutation of values rather than on some theoretical distribution. A clear advantage attributed to such a procedure is that the results should be unbiased if the data were subject to nonlinear transformations, such as monotonic or binary transformations. Currently, however, we have only tested results for continuous data.

And thirdly, all the samples drawn from these models are of order 10×10. This was considered a reasonably small network, and larger ones should behave at least as well (i.e., their standard errors should be smaller). Smaller networks, such as the 5×5 example presented at the start of this paper, may be more unstable. Again, simulations could be performed to test how sensitive the bias and statistical power of the QAP partialling procedure is to smaller sample sizes.

All of these points are fertile ground to be explored in attempts to discover the boundaries around which the QAP test can be applied. Given these limitations, however, we are confident that the QAP test of spuriousness is appropriate for many applications that can be approximated by this simple model.

4. Conclusion

We began this paper by suggesting that structural analysis may frequently be subject to the criticism that within the multiplexity of networks lie potential sources of spurious correlations. We have suggested one way in which this criticism may be empirically addressed. Moreover, we have demonstrated, at least within a narrow but reason-

able class of models, that the QAP partialling technique is a powerful and unbiased method of answering simple questions about the spuriousness of observed relationships.

We have left unanswered important questions, such as how does this technique compare to alternatives, such as that proposed by Hubert (1985), and how sensitive is the technique to changes in the assumptions of the models. But our purpose is not to dictate a universal solution to a statistical problem. Rather, we wish to draw attention to the need to consider potential sources of spurious explanations in the empirical research of network analysts. At a minimum, we hope to inspire further consideration of these issues and appropriate methods for addressing them.

References

Baker, Frank B. and J. Lawrence Hubert
  1981 "The analysis of social interaction data: A nonparametric technique." *Sociological Methods and Research* 9: 339–361.
Costanzo, C.M., L.J. Hubert and R.G. Golledge
  1983 "A higher moment for spatial statistics." *Geographical Analysis* 15: 347–351.
Dietz, E.J.
  1983 "Permutation tests for association between two distance matrices." *Systematic Zoology* 32: 21–26.
Dow, M. and J. Cheverud
  1985 "Comparison of distance matrices in studies of population structure and genetic micro differentiation: Quadratic assignment." *American Journal of Physical Anthropology* 68: 367–373.
Faust, K. and A.K. Romney
  1985 "The effect of skewed distributions on matrix permutation tests." *The British Journal of Mathematical and Statistical Psychology* 38: 152–160.
Hubert, L.J.
  1983 "Inference procedures for the evaluation and comparison of proximity matrices." In J. Felsenstein (Ed.), *Numerical Taxonomy*. New York: Springer-Verlag.
Hubert, J.
  1985 "Combinatorial data analysis: Association and partial association." *Psychometrika* 50: 449–467.
Hubert, L.J. and R.G. Golledge
  1981 "A heuristic method for the comparison of related structures." *Journal of Mathematical Psychology* 23: 214–226.
Hubert, L.J. and J. Schultz
  1976 "Quadratic assignment as a general data analysis strategy." *British Journal of Mathematical and Statistical Psychology* 29: 190–241.
Kenny, David A.
  1979 *Correlation and Causality*. New York: Wiley.

Krackhardt, D.
  1984 "Partialling networks: An application in organizational behavior." Paper presented at the Social Network Conference, Phoenix, Arizona.
Krackhardt, David
  in press "A caveat on the use of the quadratic assignment procedure." *Psychometrika*.
Krackhardt, D. and R. Breiger
  1985 "Comparative advantages of QAP partialling and log linear analysis of multivariate network data." Paper given at Fifth Annual Social Network Conference, Palm Beach, Florida.
Laumann, Edward O. and Franz Pappi
  1976 *Networks of Collective Action: A Perspective on Community Influence Systems*. New York: academic press.
Mantel, N.
  1967 "The detection of disease clustering and a generalized regression approach." *Cancer Research* 27: 209–220.
Mielke, P.W.
  1979 "On asymptotic non-normality of null distributions of MRPP statistics." *Communications in Statistics – Theory and Methods* 8: 1541–1550.
Nakao, K. and A.K. Romney
  1984 "A method for testing alternative theories: An example from english kinship." *American Anthropologist* 86: 668–673.
Proctor, C.H.
  1969 "Analyzing pair data and point data on social relationships, attitudes and background characteristics of Costa Rican Census Bureau employees." Social Statistics Section, Proceedings of the American Statistical Association, 457–465.
Sokal, R.R.
  1979 "Testing statistical significance of geographic variation patterns." *Systematic Zoology* 28: 227–232.
Sokal, R.R., P.E. Smouse and J.V. Neel
  1986 "The genetic structure of a tribal population, the Yanomama Indians: XV: Patterns inferred by autocorrelation analysis." *Genetics* 114: 259–278.