

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The version of the following full text has not yet been defined or was untraceable and may differ from the publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/26948>

Please be advised that this information was generated on 2021-09-23 and may be subject to change.

**WORKING PAPER SERIES ON
RESEARCH IN RELATIONSHIP MANAGEMENT**

**Sensitivity of MRQAP Tests to Collinearity and
Autocorrelation Conditions**

David Dekker, David Krackhardt and Tom A. B. Snijders

WORKING PAPER SERIES ON RESEARCH IN RELATIONSHIP MANAGEMENT	
Reference number	RRM-2005-11-MKT
ISSN	1572-4255
Publication status / version	May 2005
Email address contact author	d.dekker@fm.ru.nl
URL (electronic version)	http://www.ru.nl/fm
Address	Nijmegen School of Management University of Nijmegen Thomas van Aquinostraat 1 P.O. Box 9108 6500 HK Nijmegen, The Netherlands Phone: +31 (0)24 361 54 62 Fax: +31 (0)24 361 30 79

Sensitivity of MRQAP Tests to Collinearity and Autocorrelation Conditions

David Dekker
Radboud University Nijmegen

David Krackhardt
Carnegie Mellon University

Tom A. B. Snijders
University of Groningen

June 27, 2005

Abstract

MRQAP (Multiple Regression – Quadratic Assignment Procedure) tests also known as Mantel tests are permutation tests for multiple regression coefficients for data organized in square matrices of relatedness among n objects, instead of vectors. Such a data structure is typical in social network studies, where variables indicate some type of relation between a given set of actors. Over the last 15 years, new approaches to permutation tests have been developed. Some of the proposed tests have been found to be substantially more robust against collinearity in the data (e.g., correlation among explanatory variables, correlation between ancillary explanatory variables and the dependent variable). Most studies evaluating the performance of permutation tests in linear models for square matrices do not consider the type of structural autocorrelation that is typical for social network data (e.g., dependence between observations in rows and/or columns). We present a new permutation method that complements the family of extant tests. Performance of various different approaches to MRQAP tests is evaluated under conditions of row and column autocorrelation in the data as well as collinearity between the variables through an extensive series of simulations.

Keywords: MRQAP, Mantel tests, Permutation tests, Social Networks, Structural Autocorrelation, Collinearity

Introduction

The problem of statistical analysis of social network data is an old one (Proctor, 1969). Such data are dyadic in the sense that for a basic set of objects O_1, \dots, O_n , the observed variables refer to pairs of objects and therefore are doubly indexed, e.g., the variable Y_{ij} refers to the way in which object O_i is related to object O_j . It has long been understood that the lack of independence of dyadic observations creates a severe barrier to reasonable interpretations of statistical tests (Laumann and Pappi, 1976; Laumann et al., 1977). Indeed, using standard OLS models on dyadic data with a moderate amount of structural autocorrelation (lack of independence among observations within the rows and columns of network data) biases the estimates of the second moments to such an extent that it is not uncommon for Type I errors of t-statistics to exceed 50% (Krackhardt, 1988), rendering the significance tests unusable for all practical purposes.

Within the past twenty years, two substantially different approaches to solving this problem have been proffered: the use of linear models tested with the aid of the Mantel test (Mantel, 1967), which is often referred to as the quadratic assignment procedure (QAP) in social network studies, and exponential random graph models. The QAP approach (Mantel, 1967; Hubert, 1987; Krackhardt, 1987) provides a specific type of permutation test which keeps intact the dyadic data structure under the permutations. The principle of this test can be applied to many kinds of model, but it is usually applied to linear models for data treated as continuous. Through a series of Monte Carlo simulations, Krackhardt (1988) showed that parameters in an OLS model of network autocorrelated data could be tested using a multiple regression extension of the QAP test, called MRQAP. Exponential random graph models (Holland and Leinhardt, 1981; Frank and Strauss, 1986; Wasserman and Pattison, 1996; Snijders et al., 2004), on the other hand, are a family of models for dichotomous or other discrete network data where the focus is on modeling the specific network-related dependence structure. The present paper is in the line of the first approach.

Krackhardt's MRQAP approach had an appeal of simplicity and accessibility and at the time it was said that it "... takes an exciting step forward in the analysis of network relations" (Pattison, 1988). It has been frequently used since then to test models in network research (e.g., Borgatti and Cross, 2003; Gibbons and Olk, 2003; Sorenson and Stuart, 2001; Mizruchi, 1990). However, the simulations used in Krackhardt (1988) to justify this approach were confined to a quite restricted set of conditions, namely that there is no correlation among the independent variables nor a correlation in the popu-

lation between the dependent variable and *any* of the independent variables (that is, he took an overly strict view of the null hypothesis). That the presence of such collinearity¹ might cause problems has been shown in the literature on permutation tests (e.g., Anderson and Robinson, 2001). However, usually in this literature the problem of the typical type of dependence of matrix data is not addressed and no specific attention is given to QAP-like tests; an exception is Legendre (2000).

In this paper we study what the literature on permutation tests in linear models has to offer for MRQAP tests for social network data. We assess the problems that might occur due to collinearity in MRQAP-like tests and whether the different proposed permutation tests are robust against some types of dependence as encountered in social network research. Based on the results derived in the literature on permutation tests in linear models we conjecture about the performance of different MRQAP approaches under various collinearity and autocorrelation conditions. Furthermore, we develop a new approach that to our knowledge has not been described in the permutation test literature, nor the MRQAP literature. Finally, we evaluate these conjectures through an extensive simulation study. We show how various reasonable permutation methods, all consistent with the spirit of the originally proposed MRQAP tests, and including our new method, vary considerably in their ability to provide statistical tests with correct type-I error rates and good power properties.

QAP Permutation Test for Regression Coefficients

Background

The QAP approach to studying dyadic data in general has been widely used by researchers in a broad array of disciplines, such as statistics (Oden and Sokal, 1992), biology (Legendre, 2000), and psychology (Hubert, 1987). The approach was first suggested by the statistician Mantel (1967) to address the epidemiological question of whether the distribution of diseases appeared to be significantly co-located. But it was Hubert (1987) who adopted the term "quadratic assignment" and who found a vast array of applications for this type of test.

Since these early days, many statisticians have uncovered a variety of techniques for performing these permutation tests and concomitant advan-

¹Here we use the term collinearity to indicate any type of correlation between the variables used in an analysis.

tages and disadvantages for different data situations (see Legendre, 2000, for good review on distance data) (see Anderson and Legendre, 1999; Anderson and Robinson, 2001, for good studies on permutation tests in ordinary regression). However, the vast majority of these studies have focused on the type of autocorrelation proposed in the original Mantel study — that is, observations are thought to be autocorrelated to the extent that they are geographically proximal. In Legendre et al. (1994), autocorrelation was phylogenetic and not geographic. The focus of this paper, by contrast, is network data that are commonly considered to have a different kind of autocorrelation structure, wherein dyadic observations are autocorrelated if they are in the same row (emanate from the same sender) or in the same column (go to the same receiver) (e.g., Holland and Leinhardt, 1981). We wish to assess whether these recent theoretical developments and insights into the properties of the QAP permutation tests hold up under conditions of network autocorrelated data.

Bivariate QAP

The MRQAP test was developed as an extension of the bivariate QAP model. Mantel (1967) first proposed the quadratic assignment procedure (QAP) to assess association between spatial and temporal distance data using a regression approach. These data can be represented in square $n \times n$ matrices where the elements reflect the distances in space between n objects (O_j , where $j = 1, \dots, n$) and the diagonal is structurally zero. Application of the QAP method subsequently was proposed for other structurally autocorrelated data such as social network data where square matrices reflect relations between individuals (e.g., Baker and Hubert, 1981; Krackhardt, 1987). Matrices that reflect social networks are specific in that dyadic relations are often asymmetric. This is usually not the case with the distance matrices analyzed in other fields using the Mantel test, although it happens from time to time.

Assume we wish to use QAP to assess statistical significance of some measure of association Γ_{YX} between distance or network variables Y and X , where both Y and X are $n \times n$ matrices referring to a set of n objects O_1, \dots, O_n , and where the (i, j) elements of the matrices refer to aspects of how objects i and j are related. The objects are points in space for distance data, and network vertices (e.g., social actors) for network data. The data matrices could be symmetric or non-symmetric, depending on the situation; this distinction is of no concern to the discussion given here. The association coefficient Γ_{YX} could be, e.g., the elementwise Pearson product

moment correlation between X and Y .

In the QAP we randomly permute the n objects O_j to obtain a new order of objects $O_{\pi(j)}$, the permutation being denoted by π . For variable Y we calculate the distances or network variables between all objects in the order $O_{\pi(j)}$; the matrix obtained is called $\pi(Y)$. This is equivalent to applying the same permutation to the row and column indices of Y . Since the permuted structure is isomorphic to the original, all the structural features of the permuted matrix are retained, except for those referring to the order of the objects. The result is that we generate a random data set $\pi(Y)$ with the same row-column interdependence of observations as there was in Y .

Under the null hypothesis of no association between Y and X , the association $\Gamma_{\pi(Y)X}$ between $\pi(Y)$ and X is a random association that was drawn from the same underlying distribution as the actually observed association Γ_{YX} . Now we can simulate this distribution by repeating the permutations of the data and after each permutation calculating $\Gamma_{\pi(Y)X}$. This generates a reference distribution of values for Γ under the set of all $n!$ permutations. Comparing the observed Γ_{YX} to this reference distribution provides a distribution-free way of testing the null hypothesis that X and Y are independent, under the model assumption that the distribution of the variables arranged in X and Y is permutationally invariant under permutation of the objects O_1, \dots, O_n . In practice, it suffices to use a relatively small random sample (1000 or 10000) from the set of all $n!$ permutations to approximate this reference distribution (see also Pitman, 1937; Mantel, 1967; Jackson and Somers, 1989).

Hence, the QAP provides a permutation- or randomization-based non-parametric test of dependence between two square matrix variables of the same size, which may represent, e.g., distances or similarities between objects, or a relation in a group of social actors. The QAP procedure also directly provides the p -value, which is the probability of observing a value of the test statistic at least as large as the actually observed value under the null hypothesis (or in short: the probability of the data under H_0). The p -value is calculated as the relative frequency of the values of the statistic, in the null distribution, that are larger than or equal to the observed value.

Validity of randomization tests hinges on the exchangeability of values of variables under permutation. Another way of looking at the QAP test for square data matrices is that a specific subset of all possible permutations of the matrix elements is used. Note that the same value of the evaluated association statistic would be given if we would represent the data in vector format, omitting the structural zeros. However, randomizing such vectors would lead to $(n(n-1))!$ possible permutations in contrast to the $n!$ possible

permutations for the $n \times n$ squared matrices. Hence, QAP only uses a subset of all possible permutations given the total number of $n(n - 1)$ matrix elements. This is precisely the subset of permutations that keep intact the autocorrelation structure among the observations, subject to the exchangeability, or distributional invariance under permutation, of the n objects.

MRQAP Approaches

In Multiple Regression - QAP (MRQAP), the study of the dependence between matrices X and Y is compounded because there are other variables Z for which the association between X and Y must be controlled². Overviews and comparisons for vector data are given by Anderson and Legendre (1999); Anderson and Robinson (2001), for analysis of variance (ANOVA) by Anderson and Ter Braak (2003), and for distance matrices by Legendre (2000). The approach taken is the multiple linear regression approach: it is assumed that Y depends linearly on Z , and it is tested whether there is an additional (linear) dependence on X . It should be noted that in all cases, for vector and matrix data alike, linear dependence between variables is proposed as an approximation that is convenient and, hopefully, a good representation of reality. In the case of real-valued numerical data, the approximation often is quite good, and it may be expected that there is sufficient robustness for small deviations from linearity. For binary and other coarsely grained discrete data, however, linear models are less appropriate and other models for spatial or network data such as, e.g., those proposed by Snijders et al. (2004) or Heagerty and Lele (1998) may be used.

The basic linear model for square matrix data considered here is

$$Y = \beta X + Z\gamma + E \tag{1}$$

where Y , X , and E are $n \times n$ matrices, β is a scalar, Z is a $n \times n \times q$ array, and γ is $q \times 1$. The diagonals of the matrices are always assumed to be structurally zero, and of no concern. The null hypothesis is

$$H_0 : \beta = 0 . \tag{2}$$

The variables Z and X are not assumed to be independent. Specifically, we assume between these variables the linear model

$$X = \delta Z + V . \tag{3}$$

²All variables in this paper are considered to be mean-centred

where V is a $n \times n$ matrix. The situation $\delta \neq 0$ will be called collinearity. The nonparametric approach to square matrix data means here that the residuals associated to the n objects, i.e. the elements of the matrices E and V , respectively, are exchangeable, or equivalently, invariant under permutations of rows and columns simultaneously by the same permutation. Whenever the term *permutation* is used, it will be assumed that this permutation acts on rows and columns simultaneously and in the same way, as describe above for the QAP procedure.

It will be convenient to use the following notation. By

$$\hat{\beta} = \hat{\beta}(Y, X \mid Z) \quad (4)$$

is denoted the OLS estimator of β -the partial regression coefficient- for the vectorized data. We use the notation, $\pi(Y)$, for the data matrix Y where both rows and columns are rearranged by the permutation π .

Various permutation procedures intended to test the null hypothesis H_0 will be considered, with the objective to indicate the test or tests that can best be used in view of having type-I error rates close to the nominal level of significance (i.e., robustness to collinearity in the sense that $\delta \neq 0$, and robustness to autocorrelation in the sense of within-row and within-column dependence) and having a high power.

The permutation principle can be applied to various different test statistics. The QAP was explained above using as the test statistic an association coefficient such as the elementwise correlation coefficient between the two matrices X and Y .

Consider first the bivariate case where the third variable Z is absent from the model. Then the dependence between X and Y can be tested by the correlation coefficient r in a permutation distribution. The following statistics can be used:

1. the correlation coefficient $r = r(X, Y)$,
2. the t transformation $t = \sqrt{n(n-1)-2}r/\sqrt{1-r^2}$ — which has an exact t null distribution only if X and Y are normally distributed and there is no network autocorrelation, i.e., all elements of the matrices X and Y are independent —,
3. or the estimated regression coefficient

$$\hat{\beta} = \frac{s_Y}{s_X} r \quad (5)$$

where the standard deviations refer to the vectorized data.

In the bivariate QAP permutation test the estimated regression coefficient is a constant multiple of the correlation coefficient since the two standard deviations remain constant under the permutations. Therefore, these three test statistics will lead to the same results for the bivariate QAP procedure. For the MRQAP procedures where more complicated statistics must be used to control for the dependence on the third variable Z , the equivalence between these three possibilities no longer holds.

Different methods of data permutation have been proposed. They can be classified as 'raw-data' permutation methods and 'residual' methods, as will be explained below. However, not all of these methods can be applied to all statistics. For example, Kennedy and Cade (1996) show that the method of permuting the raw Y -matrix requires the use of a pivotal statistic. A statistic is defined to be pivotal when the distribution of the statistic under the null hypothesis is independent of the "nuisance" parameters; in our case, the nuisance parameters may be taken to be the residual variance and the regression coefficient δ in (3). The partial regression coefficient is not pivotal, so that the result of Kennedy and Cade (1996) rules out the use of the regression coefficient under the permutation of the raw Y -matrix. Furthermore, Anderson and Robinson (2001) show that although most 'residual' methods asymptotically give the same results, they differ in the ability to approximate an exact test in small samples. It can be concluded from the literature that in some procedures it is clearly better to use a pivotal statistic and otherwise there is no important difference. Therefore we pay most attention to permutation procedures based on the partial correlation coefficient or the associated t -statistic, which indeed are pivotal statistics. We present a new asymptotically exact test and compare its performance on accuracy and power with the other methods we discuss. First, however, we show some problems with 'raw data permutation' methods that use non-pivotal statistics, and which unfortunately have been used extensively in empirical social network studies and other fields.

X-Permutation and Y-Permutation

For the three matrix case (equation 1), intuitively attractive approaches to MRQAP are two simple extensions of the bivariate case: the X -permutation and Y -permutation procedures (e.g., Oja, 1987; Smouse et al., 1986). In the former, we permute the X matrix associated to the tested multiple regression coefficient and subsequently estimate the multiple regression coefficients for the permuted data sets to generate its reference distribution. In the latter we permute Y .

The H_0 for the permutation of X is that Y is not related to X although it could be related to Z . Now, permuting X destroys the relationship between X and Z and thereby violates the ancillarity principle (Welch, 1990; ter Braak, 1992; Anderson and Legendre, 1999), which states that the dependence between these independent variables should be kept intact in the permutational procedure. It could be expected that this lack of respect for the dependence between X and Z will lead to incorrect type-I error rates in cases of collinearity.

The H_0 for the permutation of Y is that Y is not related to X and Z taken together. It hence is a test for significance of explained variance rather than a significance test for individual parameters. However, researchers have been using the test for this purpose. This may explain that the literature on permutation tests has over the years presented contradictory advice on the use of the Y -permutation approach to MRQAP. Although Smouse et al. (1986) suggest this method is not the best approach, it has been applied in many studies. In social network analysis it has been applied perhaps unwittingly by many researchers, because it was the only approach to MRQAP implemented in the main network analysis software package UCINET, versions 3 to 6.1 (Borgatti et al., 2002).

Manly (1997) argued for the Y -permutation, but this approach was criticized by Kennedy and Cade (1996), especially if a non-pivotal statistic is used. Anderson and Robinson (2001) found that the asymptotic significance level of the Y -permutation test is indeed equal to the nominal level when an asymptotically pivotal statistic is used. An asymptotically pivotal statistic does not depend on any unknown parameters and thereby adjusts for any nuisance parameters that are not of interest to the test (Anderson and Robinson, 2001). The partial correlation coefficient and the corresponding t -statistics are examples of asymptotically pivotal statistics, independent of first and second moment population parameters. They are monotonic to each other, and can therefore be used equivalently for a significance test of individual multiple regression coefficients.

The following shows the effects of collinearity (a linear relation between X and Z corresponding to $\delta \neq 0$ in (3)) when one uses the non-pivotal partial regression coefficient as reference statistic. The following relation holds between the partial regression coefficient $b_{yx.z}$ and the partial correlation coefficient $r_{yx.z}$:

$$b_{yx.z} = r_{yx.z} \frac{\sqrt{1 - r_{yz}^2} S.D.(y)}{\sqrt{1 - r_{zx}^2} S.D.(x)}. \quad (6)$$

Under Y -permutation the variable Y is replaced with $\pi(Y)$. The partial regression coefficient under permutation is:

$$b_{\pi(y)x.z} = r_{\pi(y)x.z} \frac{\sqrt{1 - r_{\pi(y)z}^2}}{\sqrt{1 - r_{zx}^2}} \frac{S.D.(y)}{S.D.(x)}. \quad (7)$$

Under random permutations, $E(r_{\pi(y)z})$ will be approximately 0, but this is not necessarily true for the non-permuted data. Now, if $\rho_{yz} \neq 0$ the values $r_{\pi(y)z}$ under random permutations will tend to be lower, and consequently the reference distribution consisting of the values of (7) will tend to be larger in absolute value. This will lead to a conservative test with an unnecessarily low power.

Similarly, when we use the multiple regression coefficient under X -permutation the reference values are

$$b_{y\pi(x).z} = r_{y\pi(x).z} \frac{\sqrt{1 - r_{yz}^2}}{\sqrt{1 - r_{z\pi(x)}^2}} \frac{S.D.(y)}{S.D.(x)}. \quad (8)$$

Under permutation, $E(r_{z\pi(x)})$ is approximately 0, which again is not necessarily true under the null-hypothesis ($\beta_{yx.z} = 0$). If $\rho_{zx} \neq 0$ the reference values will tend to be too small in absolute value, leading to a liberal test which is undesirable.

It was proved by Anderson and Robinson (2001) that for permutation tests on i.i.d. data, the asymptotic permutational distribution of the partial correlation coefficient, which is the correlation between the residuals of Y and X , under permutations of Y , is standard normal under the null hypothesis. This implies that the Y -permutation approach gives asymptotically a test with the correct type-I error rate when a pivotal statistic is used.

The Freedman-Lane Approach to Permutations

Next to the X -permutation and Y -permutation approaches, various approaches have been proposed in the literature that are based on fitting a regression model, calculating the residuals, and permuting these residuals. These residuals are estimates of E . If the values of E would be known exactly then an exact test could be based on these, but the fact that only approximations to E are available leads to a diversity of solutions, none of which are exact (Anderson and Robinson, 2001).

We follow the procedure proposed for the usual linear model by Freedman and Lane (1983) and discussed further by various other authors (cf.

Anderson and Robinson, 2001). In our case the structure is more complicated because we are dealing with square arrays expressing relational data, but the basic approach can be similar.

Now consider the general case that we do have the third variable Z in the model. The null hypothesis is that there is no effect of X on Y ; all variation in Y can be explained by Z , or is modelled by E . The proposal by Freedman and Lane (1983) is, first, to estimate the effect of Z alone on Y , yielding the residuals

$$\hat{\epsilon}_{YZ} = Y - \hat{\gamma}Z \quad (9)$$

where $\hat{\gamma}$ is the OLS regression parameter estimate estimate for the reduced model

$$Y = \gamma Z + E . \quad (10)$$

The residuals are equal to

$$\hat{\epsilon}_{YZ} = (\gamma - \hat{\gamma})Z + E , \quad (11)$$

which, under H_0 , will be close to E if n is large and the model is well-specified. Second, they propose to permute the residuals and calculate a new hypothetical Y -value made of the fitted portion of the regression equation plus the permuted residuals:

$$Y^\pi = \hat{\gamma}Z + \pi(\hat{\epsilon}_{YZ}) . \quad (12)$$

which is different from $\pi(Y)$ introduced earlier. For each of the $n!$ permutations π , the pivotal F -statistic is calculated where Y^π is regressed on X controlling for Z , i.e., the model

$$Y^\pi = \beta X + \gamma Z + E . \quad (13)$$

is applied. This statistic may be called $F(Y^\pi, X \mid Z)$. For a one-dimensional X -variable, this is a constant multiple of the squared t -statistic. The result proved by Freedman and Lane in the usual linear model is that for large n and normal error, the permutation distribution of this statistic is approximately a multiple of chi squared, and the same as the large sample distribution of the usual parametric F test. For permutation testing, the important thing is that this asymptotic distribution does not depend on the unknown δ or γ , and that under $H_0 : \beta = 0$, it is the same as the asymptotic distribution that would be obtained if the true E would be used instead of $\hat{\epsilon}_{YZ}$. This implies that the permutation test is a valid test of the null hypothesis defined by (1), (2), given permutational invariance. This result

is subject to some assumptions which say basically that there are no outliers in the Z or Y variables.

If Z is absent from the model, then the two procedures presented in the preceding subsection are also equivalent to this procedure, because permuting X with constant Y or permuting Y with constant X is equivalent if there is no Z variable.

Comparing the F test on the residuals with formula (13), and using (12) and (11), we have

$$\begin{aligned}\hat{\beta}(Y^\pi, X | Z) &= \hat{\beta}(\hat{\gamma}Z + \pi(\hat{\epsilon}_{YZ}), X | Z) \\ &= \hat{\beta}(\pi(\hat{\epsilon}_{YZ}), X | Z) \\ &= \hat{\beta}(\pi((\gamma - \hat{\gamma})Z), X | Z) + \hat{\beta}(\pi(E), X | Z)\end{aligned}\quad (14)$$

and in the last line, the second term has a permutation invariant distribution while the first term is negligible when n is large and therefore $(\gamma - \hat{\gamma})$ is small. This suggests that the Freedman-Lane approach of permuting residuals can be used not only for pivotal statistics but also when using the partial regression coefficient as a test statistic. The same procedure for pivotal statistics was proposed by Anderson and Legendre (1999) for vector data and Legendre (2000) for matrix data.

This approach may be called a *Y*-conditional-semi-partialing regression (FLSP) approach, since the effect of Z is partialled out from Y and then the resulting Y residuals are permuted and regressed on X and Z . Note that in the process Z is entered in the regression twice. The effect of Z need not be also partialled out from X , since Z is used in the regression anyway and X and Z are not permuted with respect to each other. Furthermore, putting in $\hat{\gamma}Z$ as is done in (12) does not change the F statistic because Z is being controlled for in model (13). This suggests that we could also use the partial correlation $r(\hat{\epsilon}_{YZ}, X | Z)$ as a test statistic, and test it in the permutational distribution $r(\pi(\hat{\epsilon}_{YZ}), X | Z)$. This will also be approximately the same as using the F test as proposed by Freedman and Lane (cf. the ‘first idea’ in their proof, which implies that asymptotically, $F(Y^\pi, X | Z)$ and $r^2(Y^\pi, X | Z)$ are functions of each other).

Here, we introduce a new approximate exact test that follows a similar logic as the *FLSP* and hence complements the set of similar existing permutation tests. Whereas the *FLSP* approach conditions the test statistic under permutation on the relationship between Y and Z , we could follow an analogous reasoning and condition it on the relationship between X and Z . Although this is a logical analogue of the Freedman & Lane-approach, it seems not to have been proposed before.

Similar to (9) we define

$$\hat{\epsilon}_{XZ} = X - \hat{\delta}Z \quad (15)$$

where $\hat{\delta}$ is the OLS estimate for the model

$$X = \delta Z + V . \quad (16)$$

In the new method these residuals are permuted and the model

$$Y = \beta \pi(\hat{\epsilon}_{XZ}) + \gamma Z + E \quad (17)$$

is used to obtain reference values for the test statistic. The rationale is here that under the null hypothesis $\beta = 0$, the reference model (17) for Y is the same as the original model (1), and if the estimation error $\hat{\delta} - \delta$ is negligible, the permutational invariance assumption for V implies that

$$\pi(\hat{\epsilon}_{XZ}) = \pi((\delta - \hat{\delta})Z + V) \quad (18)$$

has the same distribution as V .

This approach has some parallels to the Freedman & Lane approach that focuses on generating a permuted Y . More, specifically Y^π , which as is shown in(14), is equal to using $\pi(\hat{\epsilon}_{YZ})$ (see also Anderson and Legendre, 1999). Instead of focusing on generating a permuted Y^π this procedure focusses on creating an X^π , or rather, similar to what Anderson and Legendre (1999) propose to simplify things, focus on $\pi(\hat{\epsilon}_{XZ})$.

This approach is coined *Double-Semi-Partialing regression (DSP)* since the effect of Z is partialled out of X and then the resulting residuals $\hat{\epsilon}_{XZ}$ are permuted and entered in a regression of Y on $\hat{\epsilon}_{XZ}$ and Z . As such Z enters the regression twice, hence the "double". Where the *FLSP*-approach minimizes the effect of ancillary variables and the dependent variable under permutation, the *DSP*-approach minimizes the correlation between the focal variable and the ancillary variables under permutation. Both methods therefore respect the ancillarity principle as they condition on the nuisance statistics. Simulations will have to determine which of these two methods minimizes the effect of the "nuisance" parameters under permutation.

Many other approaches could be considered with respect to what is partialled out and what is permuted. One of the possibilities is to estimate residuals E under the full model (1) rather than the reduced model (10), as proposed by ter Braak (1992). However, Anderson and Robinson (2001) obtained for this method power properties that were inferior to some other

methods, which was confirmed by some of our own simulations (not further reported below). Therefore this approach is not further elaborated here.

Another approach proposed in the MRQAP-literature is called partialing (Smouse et al., 1986; Krackhardt, 1988; Kennedy, 1995). This approach tests $r(\epsilon_{YZ}, \epsilon_{XZ})$ in the distribution of reference values $r(\epsilon_{YZ}, \pi(\hat{\epsilon}_{XZ}) | Z)$. The partialing approach was found by Anderson and Legendre (1999) to be incorrect in the case of multiple regression, and also found to be incorrect by Legendre (2000) for the partial Mantel test.

Anderson and Legendre (1999) review many different permutation proposals for the usual linear model (with vector data). These authors conclude that the Freedman-Lane proposal (here called *FLSP*) applied to the partial correlation coefficient corresponds best to the test based on the unobserved true residuals E ; where 'best' refers to the actual significance level as well as the power of the test. For the partialing approach, which they call the Kennedy approach in reference to Kennedy (1995), they show that this test will always result in a higher level of type-I errors than the Freedman-Lane approach.

In the remainder we analyze how different MRQAP tests on partial regression coefficients perform in terms of accuracy and power when we impose different collinearity and autocorrelation conditions. Also, we compare the consequences of using non-pivotal statistics in "raw data permutation" methods. This amplifies the simulation study of MRQAP tests for matrix data carried out by Legendre (2000), who considered distributions representing distance matrices, whereas we specifically focus on dependence structures that are typical for network data.

Method

It is common practice to use MRQAP approaches on network data that are autocorrelated in the form consistent with the models we explore here (e.g., Nelson, 1989; Krackhardt and Kilduff, 1999; Borgatti and Cross, 2003). Following Krackhardt (1988), we generate random square matrices with varying levels of structural autocorrelation on a dependent variable Y_{ij} and two explanatory variables X_{ij} and Z_{ij} , as will be described in the following paragraphs. We impose dependence between the observations in rows and columns, which we refer to as structural row and column autocorrelation. Furthermore, we impose different levels of spurious correlation between Y and X . This implies collinearity between X and Z , and an association between Y and Z . More specifically, we have

$$r_{YX} = r_{YZ}r_{ZX} \quad (19)$$

where we enforce that $r_{YZ} = r_{ZX}$, and hence $\sqrt{r_{YX}} = r_{YZ} = r_{ZX}$. The value of r_{YX} is called a spurious correlation between Y and X , because it reflects an apparent relationship between Y and X , which is in fact induced by the presence of a third variable Z . When we control for Z in a multiple regression setting the bivariate effect of X on Y disappears. In the following, "spuriousness" hence indicates the correlation between Y and X due to the presence of Z .

More specifically, we generate raw autocorrelated data with the following base model:

$$Y_{ij} = K_R U_{RY_i} + K_C U_{CY_j} + K_B U_{Y_{ij}} \quad (20)$$

$$X_{ij} = K_R U_{RX_i} + K_C U_{CX_j} + K_B U_{X_{ij}} \quad (21)$$

$$Z_{ij} = K_R U_{RZ_i} + K_C U_{CZ_j} + K_B U_{Z_{ij}} \quad (22)$$

where K_R and K_C are parameters determining the levels of structural autocorrelation in respectively the rows and columns of the matrices, all U variables are independent standard normal distributed random square matrices, with identical values in rows or columns as suggested by the notation U_{RY_i} (R for row) or U_{CY_j} (C for column)(etc.), respectively.

In our simulations the parameters take the values $K_R = K_C = 1 - K_B = K$. The autocorrelation within rows and columns is given for these parameters by

$$\rho(Y_{ij}, Y_{i'j}) = \rho(Y_{ij}, Y_{ij'}) = \frac{2K^2}{1 - 2K + 3K^2} \quad (i \neq i', j \neq j') \quad (23)$$

Values were chosen for K such that autocorrelation ranges from .000 to .900.

To represent non-normally distributed data we consider count data which are quite usual in network analysis, for example, when frequencies of interactions are measured. In a similar vein as above we generate random count data matrices

$$Y_{ij} = W_{RY_i}(K_R) + W_{CY_j}(K_C) + W_{Y_{ij}}(K_B) \quad (24)$$

$$X_{ij} = W_{RX_i}(K_R) + W_{CX_j}(K_C) + W_{X_{ij}}(K_B) \quad (25)$$

$$Z_{ij} = W_{RZ_i}(K_R) + W_{CZ_j}(K_C) + W_{Z_{ij}}(K_B) \quad (26)$$

where K_R and K_C are parameters determining the levels of structural autocorrelation in respectively the rows and columns of the matrices. The W variables have Poisson distributions with parameters K_R , etc. These variables again have identical values in rows or columns as suggested by the notation. In our simulations the autocorrelations take the values $K_R = K_C = 8 - K_B = K$ with K increasing in 3 steps from .000 to 72 (not equal steps). Here the autocorrelation within rows and columns is given by

$$\rho(Y_{ij}, Y_{i'j}) = \rho(Y_{ij}, Y_{ij'}) = \frac{K}{8 + K} \quad (i \neq i', j \neq j') \quad (27)$$

which ranges from .000 to 0.900.

In addition, we impose collinearity conditions on Y_{ij} , X_{ij} and Z_{ij} (similar to Legendre, 2000). To impose different collinearity conditions we use a (symmetric) correlation matrix T of dimension 3×3 that represents the correlations between Y , X , and Z . Furthermore, we vectorize and concatenate our square data matrices into the data matrix D of dimension $n(n-1) \times 3$. A Cholesky decomposition of T provides the weights for the construction of the correlated data set DT . From this correlated data set DT of dimension $(n(n-1) \times 3)$ we reconstruct the 3 square $(n \times n)$ variable matrices Y , X , and Z . In our simulations the collinearity conditions expressed by ρ_{XZ} increase in 3 steps by .3 from 0 to .9. To keep integer values in the count data sets, we round the resulting values to the nearest integer.

In the “type I error”-simulations we record the percentage of rejections (based on 1000 runs per step) of the (true) null hypotheses, where there is no relation between Y and X . We assess robustness of MRQAP tests by evaluating results at a significance level of $\alpha = .10$ (similar to Krackhardt, 1988). To assess the consistency of the different tests based on the Monte Carlo simulations we use the lower and upper boundaries of the 99% expected outcome interval (Edgington, 1969).

For the power study we generated data for which the null hypothesis was false, i.e. $\beta \neq 0$. Data were generated in a similar fashion as in the “type I error”- study described above. The difference was that we set the correlation as follows:

$$r_{YX} = r_{YZ}r_{ZX} + .1 \quad (28)$$

where again $r_{YZ} = r_{ZX}$.

In the power study we thus vary the level of spuriousness given a extra correlation of .1 between Y and X , and we vary the levels of autocorrelation in the data.

Results

The consequences for type I error of using the non-pivotal statistic $\hat{\beta}$ in the "raw data" methods are shown in Figures 1 to 4. They show how increasing spuriousness due to collinearity of Y and X with a third variable Z harms the accuracy of these MRQAP tests. The MRQAP test using $\hat{\beta}$ with the Y -permutation will almost never reject the null-hypothesis when spuriousness increases (see Figure 1). This should translate in reduced power when there is an effect. In contrast, the MRQAP test that uses $\hat{\beta}$ with the X -permutation almost always rejects the null hypothesis (see Figure 3). Both tests are therefore uninformative. Figures 2 and 4 show that under these conditions pivotal statistics do give results close to the expected α level of .10 for the same set of permutations on the same data. These results corroborate the findings obtained by Legendre (2000) for matrix and, by Legendre and Anderson (1999) and Anderson and Robinson (2001) for i.i.d. (vector) data.

Figures 1 to 4 about here

Note that in Figures 1 to 4 the results seem to be independent of the level of autocorrelation. A close inspection of our results for the other methods we have studied (partialing, *FLSP*, *DSP*) shows similar results (see also Tables 1 and 2). This confirms and extends the results of Krackhardt (1988) as it shows that accuracy of MRQAP tests is robust against autocorrelation. However, the methods do differ in their ability to approximate the exact test. The 99% expected outcome interval for 1000 runs and $\alpha = .10$ is (.075 - .125). Values outside this interval significantly differ from the nominal value $\alpha = .10$ and indicate a biased reference distribution.

Tables 1 and 2 about here

Table 1 shows the results for Gaussian data. The values remain within the expected outcome interval. Consistent with the findings of Anderson and Robinson (2001) the results for type I error obtained for partialing are always larger than or equal to the results of *FLSP*. Interestingly, the much refuted "raw data"-methods in the permutation test literature perform very well in these simulations. In particular, Y -permutation on average performs best. Note that the statistic used in these simulations is a pivotal statistic under the specified data conditions, and not the multiple regression coefficient used in UCINET.

For skewed data all methods have outcomes closer to the boundaries of the expected outcome interval as can be seen from the consistently higher absolute deviation values (see Table 2). In line with the analytic results of Anderson and Robinson (2001) the partialing method, which utilizes residuals of both Y on Z and X on Z , shows a markedly higher average absolute deviation from $\alpha = .10$ than the other four methods. Under these conditions Y -permutation seems to outperform the other methods, although the differences are not large. It seems that count data evoke some (minor) problems for the methods presented here in a linear model.

The results of the power study are summarized in Figures 5 and 6, and in Tables 3 and 4. The results show that the power of the different methods does rely on both the levels of spuriousness and of autocorrelation in the data. The figures show that autocorrelation decreases power, while spuriousness enhances power. The fact that autocorrelation decreases power results from the fact that autocorrelated data contain less statistical information that can be utilized in this analysis. The results show that for both types of data, when autocorrelation becomes excessive, power of the test heavily deteriorates.

Figures 5 and 6 about here

A more detailed look at the results in Tables 3 and 4 better distinguishes the power of the different MRQAP tests. As would be expected partialing has the highest apparent power, which is due to its inflated rate of type I error. Because the rate of type I error is higher than the significance level, the resulting tests are invalid (see Edgington, 1995, p.13); thus they should not be considered further as a viable option here. The other methods give similar results although the residual methods perform slightly better than the "raw data permutation" methods as we expected. The power of the tests is lower for the count data than for the normally distributed data.

Tables 3 and 4 about here

Discussion and Conclusion

Many studies in many different fields employ MRQAP for analysis of network data in the linear model framework. It is a popular approach mainly because it utilizes the benefits of multiple regression analysis, and offers a solution to problems of autocorrelation that prohibit use of standard statistical tests. However, different approaches to MRQAP described in the literature and/or

implemented in standard software packages show different performances. Raw data permutation methods that use non-pivotal statistics perform so poorly that they are useless for practical purposes. In the permutation test literature these difficulties have been recognized (e.g., Anderson and Robinson, 2001). Our analysis expands on this literature and adds to it more extensive insights into the problems of the different approaches in the case of structurally autocorrelated data such as they are encountered in network analysis.

Our results corroborate earlier findings that show that using Y -permutation and X -permutation tests that employ multiple regression coefficients has severe consequences for statistical inference. Y -permutation, which has been applied in many studies, gives oppressively conservative results under conditions of spurious correlation. X -permutation tends to overly reject the null hypothesis under conditions of spurious correlation. This leads to inferring relationships when in fact there is no empirical evidence to support these relationships. This clearly is very damaging for theory development. This has consequences for all work done that uses these methods.

Our simulations show that using a pivotal statistic for most methods is most appropriate, also for residual methods. Comparing residual methods and raw data permutation methods for pivotal statistics shows that residual methods are not necessarily better as is usually suggested in the permutation literature on i.i.d. (vector) data. The Y -permutation (pivotal statistic), X -permutation (pivotal statistic), DSP (pivotal and non-pivotal), and $FLSP$ (pivotal and non-pivotal) methods all have type I errors close to α for skewed (count) data. For Gaussian and Poisson type data all these methods (raw data and residual) stay within the 99% expected outcome interval.

Despite the findings of Legendre (2000) that outliers inflate type I error for raw data methods (specifically Y -permutation), some might insist that our results suggest that Y -permutation on a pivotal statistic is a good choice, especially for Gaussian data. This method would be the most efficient implementation option for software developers, because it immediately generates results for all coefficients. However, there are questions here concerning the concept of a pivotal statistic. This is defined as a statistic that is not dependent on unknown population parameters. As such a t -statistic and a partial correlation coefficient are pivotal statistics for multivariate normal distributions. However, when data stem from other distributions where higher moments are informative, these statistics are no longer pivotal. Using these statistics in "raw data" methods may again become harmful, because of their effect through ancillary variables. The residual methods, especially the semi-partialing methods, may suffer less from this problem because they

a priori minimize ancillary effects.

The simulation results show that under various conditions the different methods perform reasonably similarly and satisfactorily for pivotal statistics with a relatively small number of objects ($n = 10$ rows and columns). However, on the basis of the preceding paragraph, we advise against using "raw data" methods. On somewhat different grounds this was done before in permutation tests literature. Also, earlier suggestions to use partialing methods must be refuted, because Anderson and Robinson (2001) show that this method always performs less well than the method proposed by Freedman and Lane (1983) (in our analysis *FLSP*). The final conclusion therefore is in favor of *DSP* and *FLSP*.

These methods allow for some straightforward extensions. Earlier we defined X as a $n \times n$ matrix. It follows that the results for the semi-partialing approaches should hold when we define X as a $n \times n \times s$ matrix, where $s + q = k$, q is the number of Z variables and k is the total number of explanatory variables. For *DSP*, we can then employ the residuals of the regressions of multiple X -variables on Z in the construction of the reference distribution. This would allow us to test the joint significance of multiple X -variables. For example, as was proposed to assess models for longitudinal network data in Dekker et al. (2003).

We must emphasize that this study has limitations that restrict our earlier, BorgattiEA02 recommendations. First, we do not know the behavior of the different approaches for other than linear models. For example, many data in network studies are nominal or ordinal and hence require other models, e.g., those of Snijders et al. (2004) for dichotomous variables. It remains for further research to show what MRQAP approaches (possibly vastly different from those discussed here) have to offer for these types of models. Finally, the advice given by Krackhardt (1988) remains important. Researchers should check whether MRQAP tests are appropriate for the models and the types of data they use in their studies. Three specific issues that can be mentioned here are the linearity assumption; the absence of gross outliers, a requirement that follows from the conditions in Freedman and Lane (1983); and the assumption of permutation invariance of the n objects which might be invalid, e.g., in the case of strong transitivity effects in the data.

References

- Anderson, M. J. and P. Legendre. 1999. "An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model." *Journal of Statistical Computation and Simulation* 62:271–303.
- Anderson, M. J. and J. Robinson. 2001. "Permutation tests for linear models." *Australian & New Zealand Journal of Statistics* 43:75–88.
- Anderson, M. J. and C. J. F. Ter Braak. 2003. "Permutation tests for multi-factorial analysis of variance." *Journal of Statistical Computation and Simulation* 73:85–113.
- Baker, F.B. and L.J. Hubert. 1981. "The analysis of social interaction data." *Sociological Methods and Research* 9:339–361.
- Borgatti, S.P., M.G. Everett, and L.C. Freeman. 2002. *Ucinet for Windows: Software for Social Network Analysis*. Harvard Analytic Technologies.
- Borgatti, S. P. and R. Cross. 2003. "A relational view of information seeking and learning in social networks." *Management Science* 49:432–445.
- Dekker, D., P. H. Franses, and D. Krackhardt. 2003. "An equilibrium-correction model for dynamic network data." *Journal of Mathematical Sociology* 27:193–215.
- Edgington, E.S. 1969. *Randomization Tests*. New York: Marcel Dekker.
- Edgington, E.S. 1995. *Randomization Tests. 3rd edition*. New York: Marcel Dekker, Inc.
- Frank, O. and D. Strauss. 1986. "Markov Graphs." *Journal of the American Statistical Association* 81:832–842.
- Freedman, D. and D. Lane. 1983. "A nonstochastic interpretation of reported significance levels." *Journal of Business and Economic Statistics* 1:292–298.
- Gibbons, D. and P. M. Oik. 2003. "Individual and structural origins of friendship and social position among professionals." *Journal of Personality and Social Psychology* 84:340–351.
- Heagerty, P. J. and S. R. Lele. 1998. "A composite likelihood approach to binary spatial data." *Journal of the American Statistical Association* 93:1099–1123.

- Holland, P. W. and S. Leinhardt. 1981. "An Exponential Family of Probability-Distributions for Directed-Graphs." *Journal of the American Statistical Association* 76:33–50.
- Hubert, Lawrence J. 1987. *Assignment Methods in Combinatorial Data Analysis*, volume 73 of *Statistics: Textbooks and Monographs*. New York and Basel: Marcel Dekker, INC.
- Jackson, D. A. and K. M. Somers. 1989. "Are probability estimates from the permutation model of Mantel's test stable?" *Canadian Journal of Zoology* 67:766–769.
- Kennedy, P. E. 1995. "Randomization Tests in Econometrics." *Journal of Business & Economic Statistics* 13:85–94.
- Kennedy, P. E. and B. S. Cade. 1996. "Randomization tests for multiple regression." *Communications in Statistics-Simulation and Computation* 25:923–936.
- Krackhardt, D. 1987. "QAP Partialling as a Test of Spuriousness." *Social Networks* 9:171–186.
- Krackhardt, D. 1988. "Predicting with Networks - Nonparametric Multiple-Regression Analysis of Dyadic Data." *Social Networks* 10:359–381.
- Krackhardt, D. and M. Kilduff. 1999. "Whether Close or Far: Perceptions of Balance in Friendship Networks in Organizations." *Journal of Personality and Social Psychology* 76:770–782.
- Laumann, Edward O., Peter V. Marsden, and Joseph Galaskiewicz. 1977. "Community-Elite Influence Structures: Extension of a Network Approach." *American Journal of Sociology* 83:594–631.
- Laumann, Edward O. and Franz U. Pappi. 1976. *Networks of Collective Action: A Perspective on Community Influence Systems*. New York: Academic Press.
- Legendre, P. 2000. "Comparison of Permutation Methods for the Partial Correlation and Partial Mantel Tests." *Journal of Statistical Computation and Simulation* 67:37–73.
- Legendre, P. and M.J. Anderson. 1999. "Distance-based redundancy analysis: testing multi-species responses in multi-factorial ecological experiments." *Ecological Monographs* 69:1–24.

- Legendre, P., F.-J. Lapointe, and Casgrain. 1994. "Modeling brain evolution from behavior: a permutational regression approach." *Evolution* 48:1487–1499.
- Manly, B.J.F. 1997. *Randomization, Bootstrap, and Monte Carlo Methods in Biology*. London: Chapman & Hall, 2nd edition.
- Mantel, Nathan. 1967. "The Detection of Disease Clustering and a Generalized Regression Approach." *Cancer Research* 27:209–220.
- Mizruchi, M. S. 1990. "Capitalist Collective Action - Competition, Cooperation, and Conflict in the Coal Industry - Bowman, Jr." *American Journal of Sociology* 96:473–475.
- Nelson, R. E. 1989. "The Strength of Strong Ties: Social Networks and Intergroup Conflict in Organizations." *Academy of Management Journal* 32:377–401.
- Oden, N. L. and R. R. Sokal. 1992. "An Investigation of 3-Matrix Permutation Tests." *Journal of Classification* 9:275–290.
- Oja, H. 1987. "On Permutation Tests in Multiple Regression and Analysis of Covariance Problems." *Australian Journal of Statistics* 29:81–100.
- Pattison, Philippa E. 1988. "Network models: Some comments on papers in this special issue." *Social Networks* 10:383–411.
- Pitman, E.J.G. 1937. "Significance Tests Which May Be Applied to Samples from Any Population." *Royal Statistical Society, Supplement* 4:119–130.
- Proctor, Charles 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* pp. 457–465.
- Smouse, Peter E., Jeffery C. Long, and Robert R. Sokal. 1986. "Multiple Regression and Correlation Extensions of the Mantel Test of Matrix Correspondence." *Systematic Zoology* 35:627–632.
- Snijders, T.A.B., P.E. Pattison, G.L. Robins, and M.S. Handcock. 2004. "New specifications for exponential random graph models." *Manuscript under review*. .

- Sorenson, O. and T. E. Stuart. 2001. "Syndication networks and the spatial distribution of venture capital investments." *American Journal of Sociology* 106:1546–1588.
- ter Braak, C.J.F. 1992. "Permutation versus Bootstrap Significance Tests in Multiple Regression and ANOVA." In *Bootstrapping and Related Techniques*, edited by K-H. Jckel, G. Rothe, and W. Sendler. Berlin: Springer-Verlag.
- Wasserman, S. and P. Pattison. 1996. "Logit models and logistic regressions for social networks .1. An introduction to Markov graphs and p." *Psychometrika* 61:401–425.
- Welch, W.J. 1990. "Construction of Permutation Tests." *Journal of the American Statistical Association* 85:693–698.

Figures

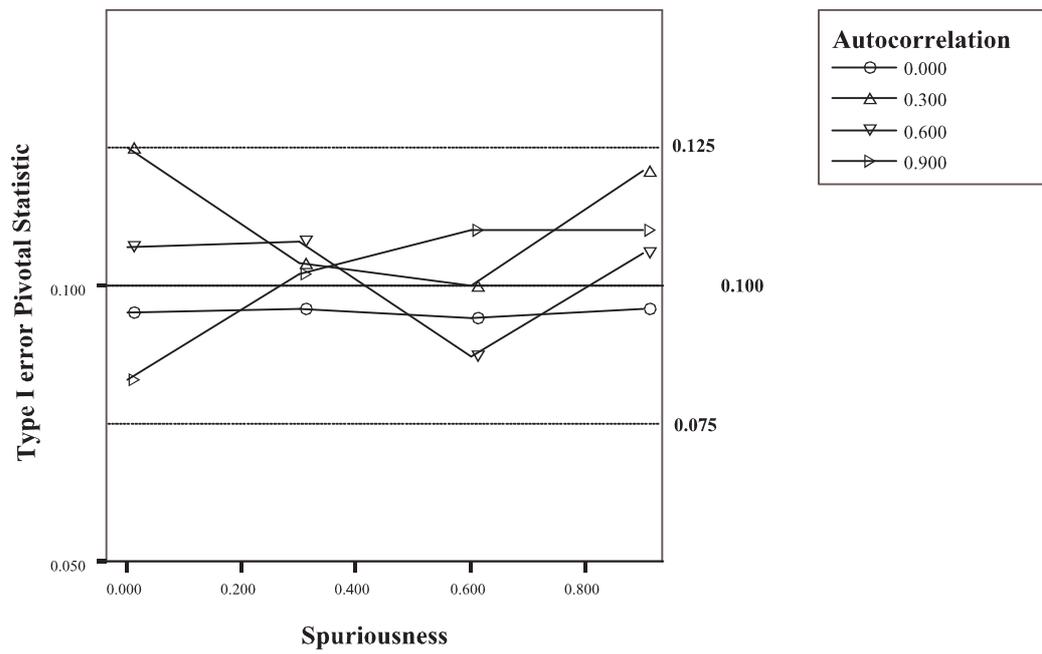


Figure 1: Type I Error of a MRQAP Test based on Y -permutation method with a non-pivotal statistic ($\hat{\beta}$) under increasing Spuriousness and Autocorrelation ($n=10, \alpha=.100$)

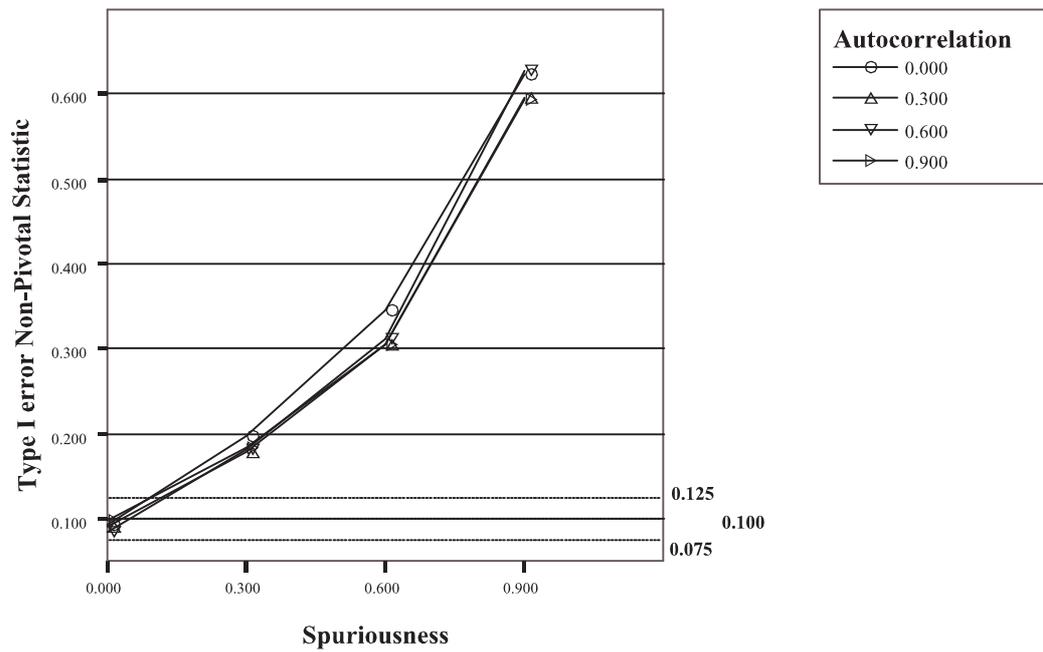


Figure 2: Type I Error of a MRQAP Test based on Y -permutation method with a pivotal statistic ($r_{XY|Z}$) under increasing Spuriousness and Autocorrelation ($n=10$, $\alpha=.100$)

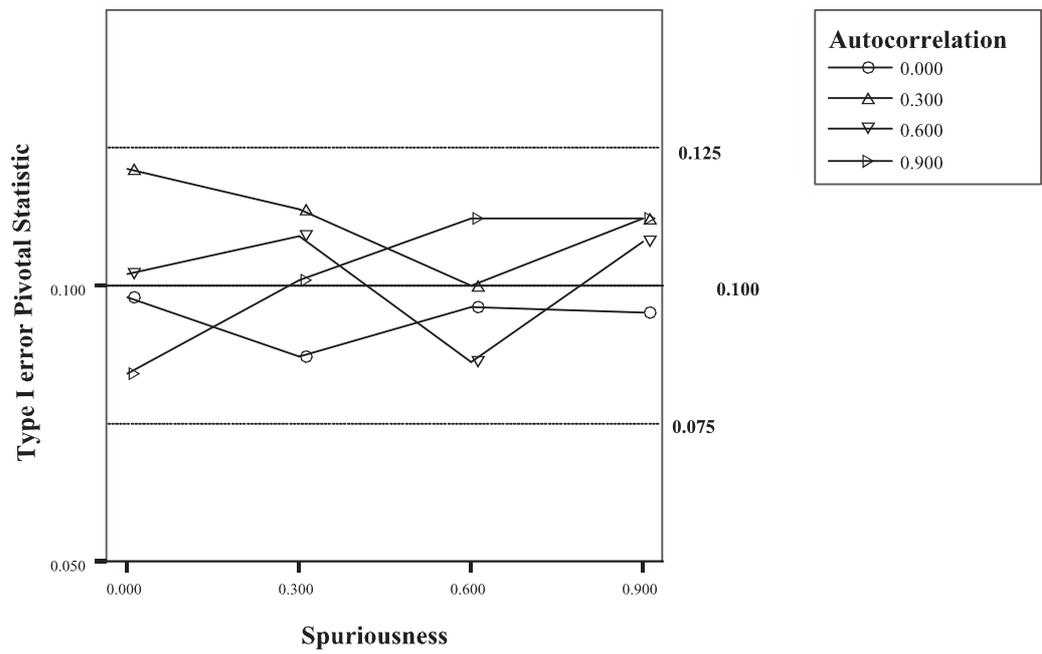


Figure 3: Type I Error of a MRQAP Test based on X -permutation method with a non-pivotal statistic ($\hat{\beta}$) under increasing Spuriousness and Autocorrelation ($n=10$, $\alpha=.100$)

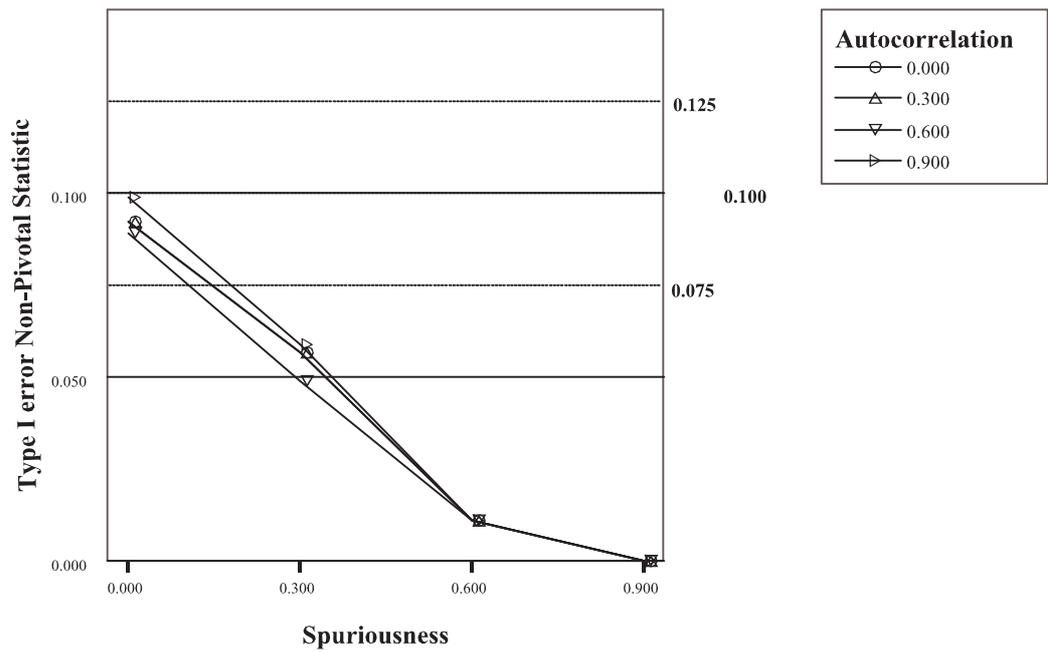


Figure 4: Type I Error of a MRQAP Test based on X -permutation method with a pivotal statistic ($r_{XY|Z}$) under increasing Spuriousness and Autocorrelation ($n=10, \alpha=.100$)

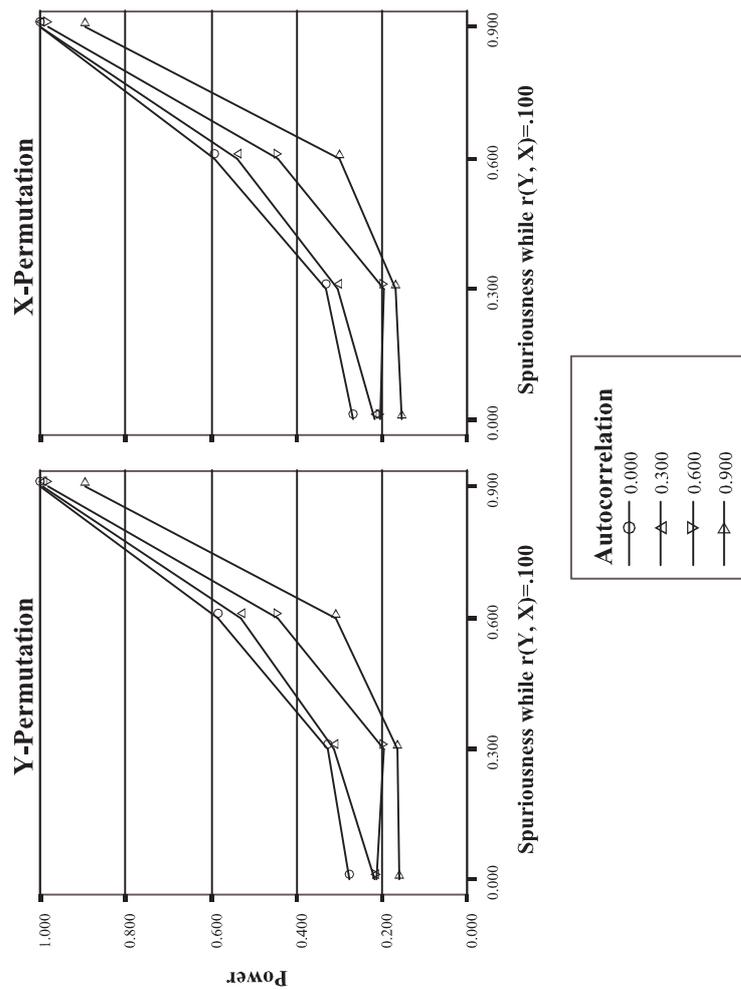


Figure 5: Power of Different MRQAP Tests under increasing Spuriousness and Autocorrelation ($n=10, Y, X, Z \sim N(0, 1)$).

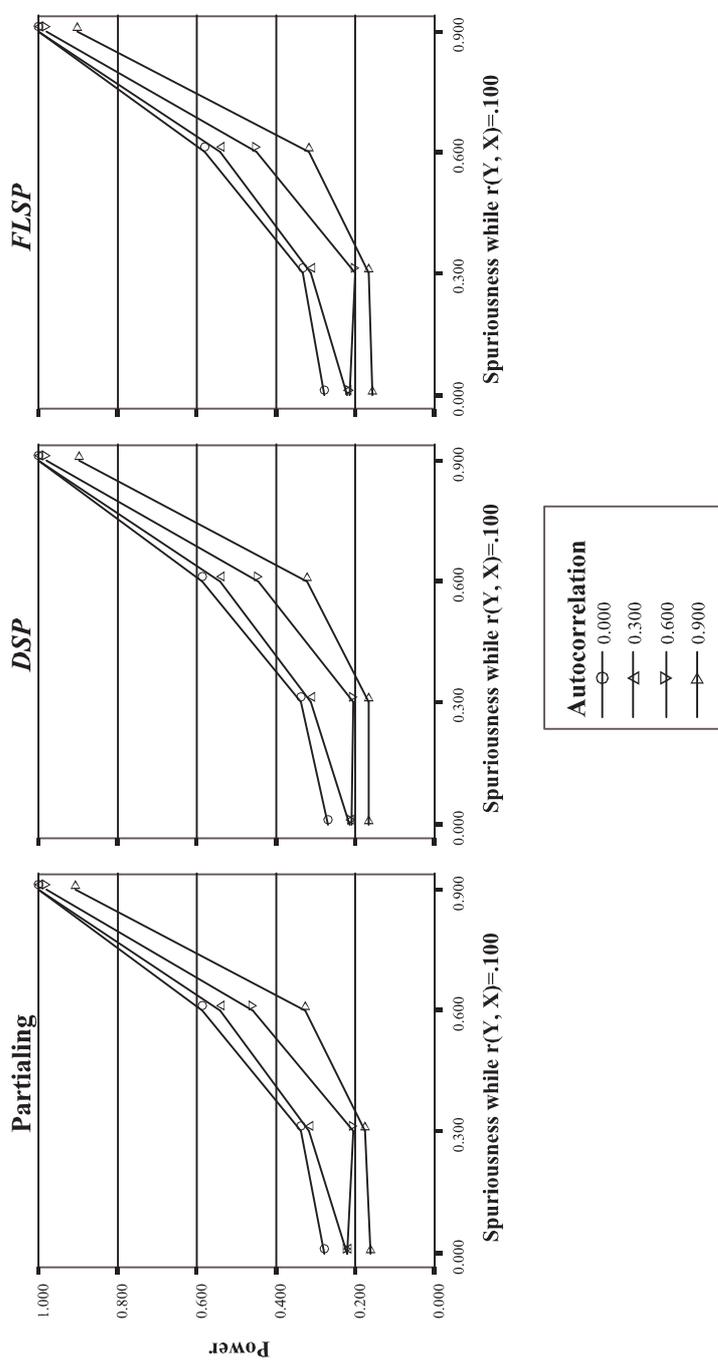


Figure 6: Power of Different MRQAP Tests under increasing Spuriousness and Autocorrelation ($n=10, Y, X, Z \sim N(0, 1)$).

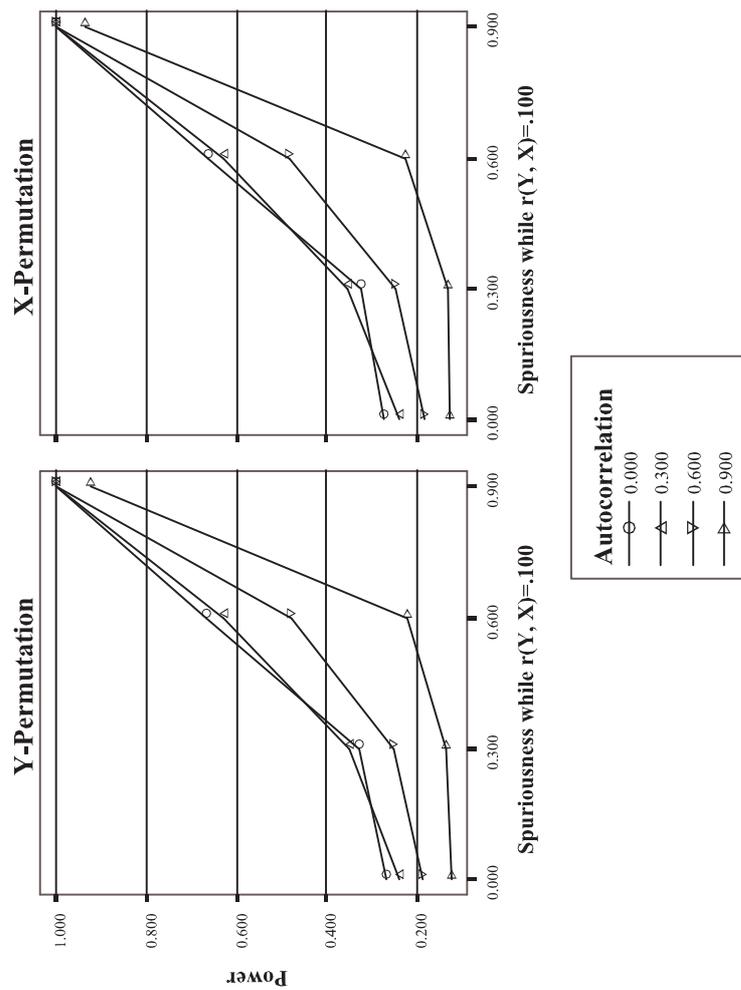


Figure 7: Power of Different MRQAP Tests under increasing Spuriousness and Autocorrelation ($n=10$, based on Poisson data).

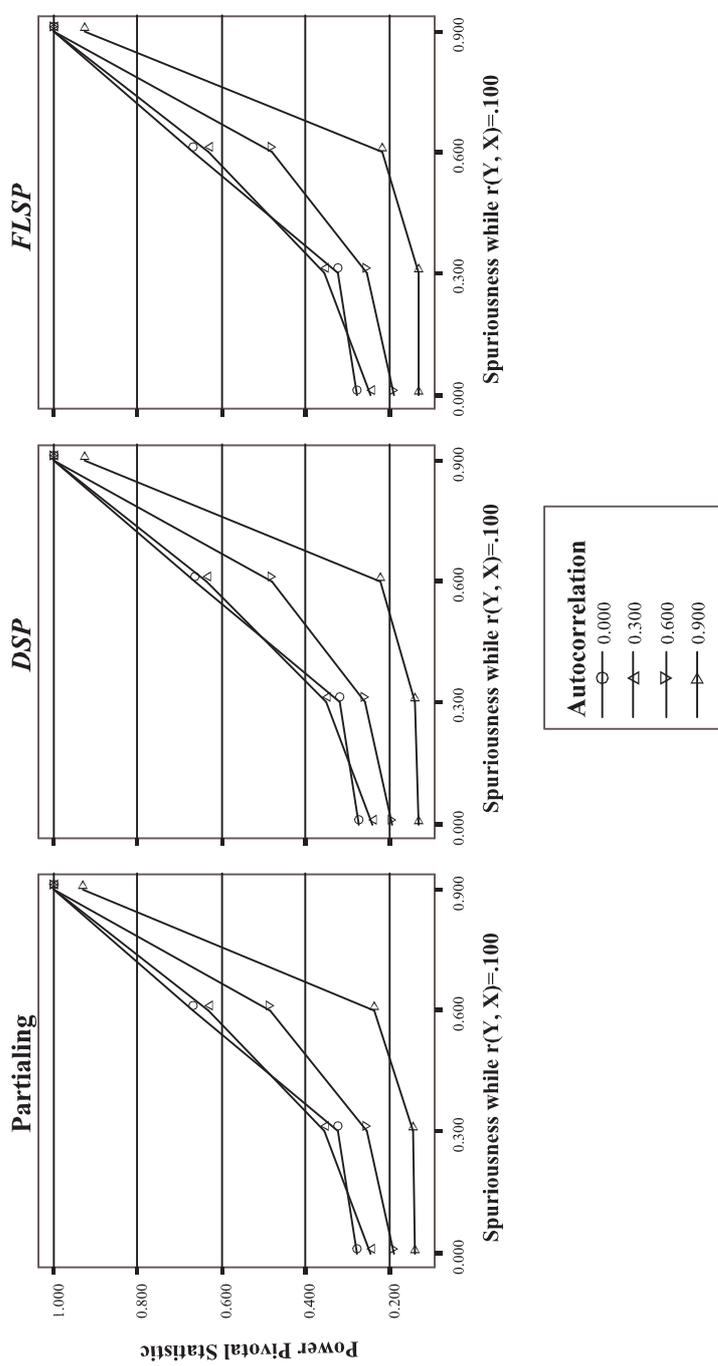


Figure 8: Power of Different MRQAP Tests under increasing Spuriousness and Autocorrelation ($n=10$, based on Poisson data).

Tables

		Spurious Correlation ρ_{YX}				
		Autocorr. par. (K)	.000	.300	.600	.900
Y-Permutation	.000	.092	.104	.108	.099	
Avg. Abs. Dev. = .007625	.300	.093	.104	.101	.113	
	.600	.0880	.113	.099	.101	
	.900	.0960	.117	.093	.121	
X-Permutation	.000	.0940	.110	.111	.101	
Avg. Abs. Dev. = .0091875	.300	.091	.108	.098	.117	
	.600	.0840	.113	.096	.108	
	.900	.0960	.111	.090	.117	
Partialing	.000	.0940	.111	.114	.100	
Avg. Abs. Dev. = .0094375	.300	.094	.110	.104	.120	
	.600	.0920	.111	.099	.105	
	.900	.107	.117	.110	.121	
DSP	.000	.0950	.117	.108	.0970	
Avg. Abs. Dev. = .0071875	.300	.0900	.103	.100	.112	
	.600	.088	.110	.091	.102	
	.900	.099	.108	.097	.112	
FLSP	.000	.094	.110	.112	.097	
Avg. Abs. Dev. = .0069375	.300	.094	.107	.097	.117	
	.600	.089	.107	.093	.101	
	.900	.101	.110	.100	.110	

Table 1: Type I errors for different MRQAP-methods applied to the pivotal t -statistic. Identical results were obtained for the partial correlation coefficient. Each cell is based on 1000 runs. Each run based on 1000 random draws with replacement from permutation distribution. The size of the data matrices was 10×10 . Standard normal data ($Y, X, Z \sim N(0, 1)$).

		Spurious Correlation ρ_{YX}				
		Autocorr. par. (K)	.000	.300	.600	.900
Y-Permutation		.000	.092	.104	.108	.099
Avg. Abs. Dev. = .008875		.300	.093	.104	.101	.113
		.600	.088	.113	.099	.101
		.900	.096	.117	.093	.121
X-Permutation		.000	.094	.110	.111	.101
Avg. Abs. Dev. = .009063		.300	.091	.108	.098	.117
		.600	.084	.113	.096	.108
		.900	.096	.111	.090	.117
Partialing		.000	.094	.111	.114	.100
Avg. Abs. Dev. = .011375		.300	.094	.110	.104	.120
		.600	.092	.111	.099	.105
		.900	.107	.117	.110	.121
DSP		.000	.095	.117	.108	.097
Avg. Abs. Dev. = .009188		.300	.0900	.103	.100	.112
		.600	.0880	.110	.0910	.102
		.900	.0990	.108	.0970	.112
FLSP		.000	.094	.110	.112	.097
Avg. Abs. Dev. = .010375		.300	.094	.107	.097	.117
		.600	.089	.107	.093	.101
		.900	.101	.110	.100	.110

Table 2: Type I errors for different MRQAP-methods applied to the pivotal t -statistic. Identical results were obtained for the partial correlation coefficient. Each cell is based on 1000 runs. Each run based on 1000 random draws with replacement from permutation distribution. The size of the data matrices was 10×10 . Based on Poisson data.

		Spurious Correlation ρ_{YX}			
	Autocorr. par. (K)	.000	.300	.600	.900
Y-Permutation	.000	.277	.326	.587	1.000
	.300	.219	.313	.532	.999
	.600	.213	.197	.445	.984
	.900	.158	.164	.307	.896
X-Permutation	.000	.268	.332	.592	1.000
	.300	.217	.305	.538	1.000
	.600	.203	.195	.443	.983
	.900	.156	.167	.301	.896
Partialing	.000	.276	.337	.587	1.000
	.300	.222	.316	.544	1.000
	.600	.222	.206	.457	.982
	.900	.161	.178	.328	.908
<i>DSP</i>	.000	.269	.339	.587	1.000
	.300	.214	.313	.542	1.000
	.600	.209	.204	.444	.982
	.900	.164	.166	.323	.900
<i>FLSP</i>	.000	.276	.332	.583	1.000
	.300	.218	.312	.542	1.000
	.600	.213	.201	.449	.981
	.900	.158	.167	.318	.902

Table 3: Power for different MRQAP-test methods applied to the pivotal t -statistic. Identical results were obtained for the partial correlation coefficient. Each cell is based on 1000 runs. Each run based on 1000 random draws with replacement from permutation distribution. The size of the data matrices was 10×10 . Standard normal data ($Y, X, Z \sim N(0, 1)$). An additional correlation was enforced of .100 between Y and X .

		Spurious Correlation ρ_{YX}			
	Autocorr. par. (K)	.000	.300	.600	.900
Y-Permutation	.000	.268	.327	.667	1.00
	.300	.239	.349	.631	1.00
	.600	.186	.253	.478	1.00
	.900	.123	.134	.221	.926
X-Permutation	.000	.271	.325	.663	1.00
	.300	.238	.353	.631	1.00
	.600	.183	.247	.481	1.00
	.900	.126	.132	.226	.936
Partialing	.000	.280	.328	.673	1.00
	.300	.247	.356	.636	1.00
	.600	.197	.259	.487	1.00
	.900	.142	.145	.240	.933
DSP	.000	.270	.317	.664	1.00
	.300	.241	.352	.637	1.00
	.600	.192	.260	.479	1.00
	.900	.127	.137	.221	.926
FLSP	.000	.277	.323	.668	1.00
	.300	.244	.354	.633	1.00
	.600	.189	.251	.479	1.00
	.900	.128	.131	.217	.927

Table 4: Power for different MRQAP-test methods applied to the pivotal t -statistic. Identical results were obtained for the partial correlation coefficient. Each cell is based on 1000 runs. Each run based on 1000 random draws with replacement from permutation distribution. The size of the data matrices was 10×10 . Based on Poisson data. An additional correlation was enforced of .100 between Y and X .