STATISTICAL MODELING OF NETWORK PANEL DATA: GOODNESS-OF-FIT

Michael Schweinberger*

Abstract

A popular approach to model network panel data is to embed the discrete observations of the network in a latent, continuous-time Markov process. A score-type test statistic for goodness-of-fit tests is proposed, which is useful for studying the goodness-of-fit of a wide range of models. The finite-sample behavior of the test statistic is evaluated by a Monte Carlo simulation study, and its usefulness is demonstrated by an application to empirical data.

Keywords: continuous-time Markov process, regular estimating functions, goodness-of-fit, Lagrange multiplier / Rao score test, Neyman C(α) test.

1 Introduction

Social network analysis (Wasserman and Faust, 1994) is concerned with links among entities. The network data considered here correspond to the directed ties among the members of a set of actors. It is common that the ties are binary, but ties may take on arbitrary values.

When modeling networks, the ties between the actors are treated as random variables. The tie variables are, however, not independent. Some well-known examples of dependencies are reciprocity (second-order dependence) and transitivity (Holland and Leinhardt, 1970, 1976), which represents third-order dependence among ties and implies clustering ("group structure") in social networks.

Since the tie variables are dependent, statistical inference proved to be hard. (Curved) exponential random graph models (ERGMs) (Snijders, Pattison, Robins, and Handcock, 2006, Hunter and Handcock, 2006) have been used to model networks observed at one time point, but—in spite of recent advances in the specification and estimation of ERGMs—some theoretical and practical issues remain.

Here, the focus is on longitudinal network data. Longitudinal network data come frequently in the form of panel data. There is a large agreement in the literature (see Frank, 1991) that the most promising models for network panel data are continuous-time Markov models which assume that the network was observed at discrete time points, and that between these time points latent, continuous-time Markov processes shape the network. Holland and Leinhardt (1977) and Wasserman (1979, 1980) proposed methods for statistical inference for such Markov models, but these methods are limited to models with second-order dependence, and thus neglect fundamental third- and higher-order dependencies.

Snijders (2001) considered a family of continuous-time Markov models which allows to model third- and higher-order dependencies, and proposed the method of moments to estimate the parameter $\boldsymbol{\theta}$. The probabilistic framework may be described as actor-

driven, that is, the nodes are assumed to represent actors who make choices concerning the ties to other actors by either adding or deleting ties, and the choices are assumed to be based on mathematical functions, containing choice "determinants" (tendencies) such as reciprocity, transitivity, and covariate-related effects.

In the tradition of K. Popper's conception of science, deeply rooted in contemporary social science and statistics (Healy, 1978), a natural question to ask is "whether the observed data support a given specification" (Rao, 2002, p. 9). The study of goodness-of-fit, dating back to Pearson (1900), is considered so natural that model-based inference without goodness-of-fit evaluation is almost inconceivable. However, as is argued below in some more detail, no goodness-of-fit measure has been proposed up to now which is applicable in a wide range of applications.

The present paper proposes a new goodness-of-fit test statistic that (1) has many applications; (2) does not require to estimate the parameters to be tested, which is in practice a decisive advantage as it (a) saves valuable computation time, and (b) allows to test hard-to-estimate parameters (which are not uncommon); (3) admits multi-parameter tests; and (4) has an appealing interpretation in terms of goodness-of-fit, in the sense that it compares the expected value of some function of the data—evaluated under some assumed model—to the observed value of the function; that is, it uses the observed data as an external benchmark to which model predictions are compared.

The paper is structured as follows. Model specification and estimation are sketched in section 2. A goodness-of-fit test statistic is proposed in section 3. Section 4 reviews interesting goodness-of-fit tests. In section 5, the finite-sample behavior of the test statistic is studied by Monte Carlo simulation. The usefulness of the test statistic for real-world problems is demonstrated in section 6 by an application to the cross-ownership network among more than 400 business firms in Slovenia (EU).

3

2 Model Specification and Estimation

Due to space restrictions, the discussion is restricted to some basic model specifications. Extensions to model the co-evolution of networks and other outcome variables (Snijders, Steglich, and Schweinberger, 2006) are possible.

A binary, directed relation $\longrightarrow$ (or digraph) on a finite set of nodes $\mathcal{N} = \{1, 2, \ldots, n\}$ is considered. The digraph is observed at discrete, ordered time points $t_1 < t_2 < \cdots < t_M$, and the observations are represented as binary matrices $\mathbf{x}(t_1), \mathbf{x}(t_2), \ldots, \mathbf{x}(t_M)$, where element $x_{ij}(t_m)$ of $n \times n$ matrix $\mathbf{x}(t_m)$ is defined by

$$x_{ij}(t_m) = \begin{cases} 1 & \text{if } i \longrightarrow j \text{ at time point } t_m, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where $i \longrightarrow j$ means that node i is related to node j; the diagonal elements $x_{ii}(t_m)$ are regarded as structural zeros.

Although the model described below assumes that $x_{ij}(t_m)$ is binary, it is possible to extend the model to the case where $x_{ij}(t_m)$ takes on discrete, ordered values.

2.1 Model Specification

The observed digraph $\mathbf{x}(t_1)$ is taken for granted, that is, the model conditions on $\mathbf{x}(t_1)$. It is postulated that the observed digraphs $\mathbf{x}(t_2), \ldots, \mathbf{x}(t_M)$ are generated by an unobserved Markov process operating in time interval $[t_1, t_M]$. Consider the case $M = 2$; the extension to the case $M > 2$ is straightforward due to the Markov property.

The model is specified by the generator of the Markov process, which corresponds to a $W \times W$ matrix $\mathbf{Q}_{\boldsymbol{\theta}}$ indexed by a parameter $\boldsymbol{\theta}$, where $W = 2^{n(n-1)}$ is the number of digraphs on $\mathcal{N}$. The elements $q_{\boldsymbol{\theta}}(\mathbf{x}^\star, \mathbf{x})$ of generator $\mathbf{Q}_{\boldsymbol{\theta}}$ are the rates of moving from digraph $\mathbf{x}^\star$ to digraph $\mathbf{x}$. If $\mathbf{x}$ deviates from $\mathbf{x}^\star$ in more than one arc variable $X_{ij}^\star$, then $q_{\boldsymbol{\theta}}(\mathbf{x}^\star, \mathbf{x}) = 0$ by assumption (see Snijders, 2001); in other words, it is assumed that the process moves forward by changing not more than one arc variable $X_{ij}^\star$ at the time.

Let $\mathbf{x}^{\star}$ be an arbitrary digraph on $\mathcal{N}$, and let $\mathbf{x}$ be the digraph that is obtained from $\mathbf{x}^{\star}$ by changing one and only one specified arc variable, say $X_{ij}^{\star}$. Since the transition from $\mathbf{x}^{\star}$ to $\mathbf{x}$ involves only the ordered pair of nodes (i, j), one can rewrite $q_{\boldsymbol{\theta}}(\mathbf{x}^{\star}, \mathbf{x})$ as $q_{\boldsymbol{\theta}}(\mathbf{x}^{\star}, i, j)$ and decompose $q_{\boldsymbol{\theta}}(\mathbf{x}^{\star}, i, j)$ as follows:

$$q_{\boldsymbol{\theta}}(\mathbf{x}^{\star}, i, j) \;=\; \lambda_i(\boldsymbol{\theta}, \mathbf{x}^{\star}) \, \mathrm{r}_i(\boldsymbol{\theta}, \mathbf{x}, j), \tag{2}$$

where

$$\lambda_i(\boldsymbol{\theta}, \mathbf{x}^{\star}) \;=\; \sum_{h \neq i}^{n} q_{\boldsymbol{\theta}}(\mathbf{x}^{\star}, i, h) \tag{3}$$

is called the rate function of node ("actor") i, while

$$\mathrm{r}_i(\boldsymbol{\theta}, \mathbf{x}, j) \;=\; \frac{q_{\boldsymbol{\theta}}(\mathbf{x}^{\star}, i, j)}{\lambda_i(\boldsymbol{\theta}, \mathbf{x}^{\star})} \tag{4}$$

is the conditional probability that i changes $X_{ij}^{\star}$, given that i changes some arc variable $X_{ih}^{\star}$, $h \neq i$.

A simple specification of λ_i is

$$\lambda_i(\boldsymbol{\theta}, \mathbf{x}^{\star}) = \rho, \tag{5}$$

where ρ is a parameter, while non-constant rate functions are given by

$$\lambda_i(\boldsymbol{\theta}, \mathbf{x}^{\star}) = \rho \, \exp\left[\boldsymbol{\alpha}' \mathrm{a}_i(\mathbf{x}^{\star}, c_i) \right], \tag{6}$$

where $\boldsymbol{\alpha} = (\alpha_k)$ is a vector valued parameter and $\mathrm{a}_i = (\mathrm{a}_{ik})$ is a vector valued function of node-bound covariates c_i and graph-dependent statistics involving the arcs of node i. If there is more than one time interval ($M > 2$), then parameter ρ can be made dependent on time interval $[t_{m-1}, t_m]$.

A convenient, multinomial logit parametrization of r_i is given by

$$\mathrm{r}_i(\boldsymbol{\theta}, \mathbf{x}, j) \;=\; \frac{\exp\left[\mathrm{f}_i(\boldsymbol{\beta}, \mathbf{x}, j) \right]}{\displaystyle\sum_{h \neq i}^{n} \exp\left[\mathrm{f}_i(\boldsymbol{\beta}, \mathbf{x}, h) \right]}, \tag{7}$$

where the real valued function

$$f_i(\boldsymbol{\beta}, \mathbf{x}, j) = \boldsymbol{\beta}' s_i(\mathbf{x}, j) \tag{8}$$

is called the objective function, while $\boldsymbol{\beta} = (\beta_k)$ is a vector valued parameter and $s_i = (s_{ik})$ is a vector valued statistic. Examples of statistics s_{ik} are the number of arcs $\sum_{h=1}^{n} x_{ih}$, the number of transitive triplets $\sum_{h,l=1}^{n} x_{ih} x_{hl} x_{il}$, and interactions of covariates with these and other statistics; see section 4. Such statistics can be used to define third- and higher-order dependencies.

2.2 Model Estimation

Model estimation is concerned with estimating parameter $\boldsymbol{\theta}$— corresponding to rate parameters $\rho_1, \ldots, \rho_{M-1}, \boldsymbol{\alpha}$ and objective function parameter $\boldsymbol{\beta}$— from the observed data $\mathbf{z}_n$, consisting of the observed digraphs $\mathbf{x}(t_1), \mathbf{x}(t_2), \ldots, \mathbf{x}(t_M)$ and covariates.

Since the Markov process is not observed in continuous time, the likelihood function is intractable. The parameter $\boldsymbol{\theta}$ is therefore estimated by the method of moments (Pearson, 1902a,b). The moment estimate $\hat{\boldsymbol{\theta}}$ is defined here as the solution of the moment equation

$$\begin{aligned} g_n(\mathbf{z}_n, \boldsymbol{\theta}) &= \sum_{m=1}^{M-1} \Big[E_{\boldsymbol{\theta}} \left[u(\mathbf{X}(t_{m+1})) \mid \mathbf{X}(t_m) = \mathbf{x}(t_m) \right] - u(\mathbf{x}(t_{m+1})) \Big] \\ &= \mathbf{0}, \end{aligned} \tag{9}$$

where $E_{\boldsymbol{\theta}}$ denotes the expectation under $\boldsymbol{\theta}$. The coordinates of the vector valued function $u = (u_k)$ in (9) correspond to the coordinates of parameter $\boldsymbol{\theta} = (\theta_k)$. In the absence of formal methods to derive statistics with certain optimum properties (neglecting some close-to-trivial models), statistics u_k are chosen heuristically so as to be sensitive to changes of the value of θ_k. To estimate objective function parameter

β_k, the statistic $\sum_{i,j=1}^{n} s_{ik}(\mathbf{x}, j)$ is a natural choice; with regard to the rate parameters, see Snijders (2001).

In terms of numerical implementation, finding the moment estimate involves finding the root(s) of $g_n = g_n(\mathbf{z}_n, \boldsymbol{\theta})$ as a function of $\boldsymbol{\theta}$. A suitable root-finding algorithm, based on iterative, stochastic approximation methods (Robbins and Monro, 1951) and Monte Carlo simulation, is described in Snijders (2001). The variance-covariance matrix of the moment estimator $\hat{\boldsymbol{\theta}}$ can be derived by the delta method (see Snijders, 2001).

3   Test Statistic

Two papers address significance testing and / or goodness-of-fit in the considered family of models, though both of them have limitations, as is argued in section 3.1. A new test statistic is proposed in section 3.2, which has important advantages compared to the existing approaches.

3.1   Existing literature

Snijders (1996) proposed to base significance tests on the pseudo-t-test statistic

$$\frac{\hat{\theta}_k}{\text{s.e.}(\hat{\theta}_k)}, \tag{10}$$

where $\hat{\theta}_k$ is the moment estimator of coordinate θ_k of $\boldsymbol{\theta}$ and s.e.$(\hat{\theta}_k)$ is its standard error. The distribution of the test statistic is unknown. Snijders (1996) assumed that its sampling distribution under the null hypothesis $H_0 : \theta_k = 0$ is approximately the standard Gaussian distribution.

Snijders (2003) stresses the modeling of (out)degree distributions. A goodness-of-fit plot with confidence intervals is provided as a means to evaluate the fit of the model with respect to the observed (out)degree distribution. However, the goodness-of-fit study is limited to the (out)degree distribution, and while the (out)degree distribution

represents a fundamental feature of the data which should be taken into account, the (out)degree distribution in itself is in most applications a nuisance. In the social sciences, the focus is on second- and third-order dependencies among arcs and covariate-related parameters. The goodness-of-fit of corresponding models cannot, however, be addressed within the Snijders (2003) framework.

3.2   A goodness-of-fit test

As to the choice of suitable test statistics for goodness-of-fit tests, observe that the choice is constrained, in the first place, by the intractable likelihood function, the absence of saturated models, and the computational burden imposed by estimating non-trivial models. Thus the "holy trinity", corresponding to the Wald, likelihood ratio (LR), and Lagrange multiplier / Rao score (RS) test (Rao and Poti, 1946, Rao, 1948), is not readily available.

Although the "holy trinity" is not available, it can be instructive to consider some of its features. An important practical concern is computation time. As the LR test requires the estimation of both the restricted and the unrestricted model, it is inferior to the Wald and RS test in terms of computation time. The RS test is most appealing, since only the restricted model must be estimated, while the Wald test requires the more computation-intensive estimation of the unrestricted model. Besides, it is not uncommon to encounter convergence problems in high-dimensional parameter spaces. In fact, it may be argued that, in the considered family of models (given the absence of saturated models), forward model selection, combined with the RS test, is preferable to backward model selection, because one can start with simple models and is not required to estimate additional parameters in the first place.

In the method of moments framework, the RS test is not available, but a RS-type test can be obtained by generalizing the $C(\alpha)$ test (Neyman, 1959) and the RS test by replacing the Fisher score function by regular estimating functions along the lines of

Basawa (1985, 1991). The classical $C(\alpha)$ test is designed to test parametric hypotheses in the presence of nuisance parameters, where the nuisance parameters are replaced by consistent estimates under the null hypothesis. If maximum likelihood estimates (under the null hypothesis) are used to estimate the nuisance parameter, the classical $C(\alpha)$ test reduces to the RS test. The RS test encompasses the classical goodness-of-fit test of Pearson (1900) as a special case (see Rao, 2002). Interrelations among the "holy trinity" test statistics and the $C(\alpha)$ test statistic were studied by, among others, Godfrey (1988, pp. 27—28), Hall and Mathiason (1990), Basawa (1991), and Bera and Bilias (2001a,b).

Partition $\boldsymbol{\theta}' = (\boldsymbol{\theta}'_1, \boldsymbol{\theta}'_2)$, where $\boldsymbol{\theta}_1$ is the (vector valued) nuisance parameter and $\boldsymbol{\theta}_2$ the (vector valued) parameter of primary interest. In the classical Neyman and Pearson tradition, goodness-of-fit can be studied by specifying hypotheses regarding the postulated family of probability distributions $\{P_{\boldsymbol{\theta}}, \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$, for instance, the null hypothesis

$$H_0 : \boldsymbol{\theta}_2 = \boldsymbol{\theta}_{20}, \tag{11}$$

tested against

$$H_1 : \boldsymbol{\theta}_2 \neq \boldsymbol{\theta}_{20}, \tag{12}$$

where $\boldsymbol{\theta}_{20}$ is some specified value (commonly, $\boldsymbol{\theta}_{20} = \mathbf{0}$), and $\boldsymbol{\theta}_1$ is unspecified; that is, both H_0 and H_1 are composite.

A test statistic for testing such composite hypotheses is proposed below, based on the estimating function g_n along the lines of Basawa (1985, 1991). The estimating function and some additional notation is introduced first.

The estimating function

For convenience, it is assumed that $M = 2$; the extension to the case $M > 2$ is immediate.

The parameter $\boldsymbol{\theta}$ is estimated by the solution of the estimating equation $\mathrm{g}_n =$

$g_n(\mathbf{z}_n, \boldsymbol{\theta}) = \mathbf{0}$, where g_n is the estimating function satisfying some mild smoothness conditions (see Godambe, 1960). Examples of estimating functions include the Fisher score function, moment functions, or more generally, some function involving parameter $\boldsymbol{\theta}$ and the data. Here, the estimating function g_n is defined in (9).

Observe that g_n as defined in (9) is an unbiased estimating function in the sense that

$$E_{\boldsymbol{\theta}}[g_n(\mathbf{Z}_n, \boldsymbol{\theta}) \mid \mathbf{X}(t_1) = \mathbf{x}(t_1)] = \mathbf{0} \tag{13}$$

holds for all n and all $\boldsymbol{\theta}$.

Let $\boldsymbol{\Sigma}_n$ be the $L \times L$ variance-covariance matrix of g_n, and denote the limit of $w_n \boldsymbol{\Sigma}_n$ as $n \longrightarrow \infty$ by $\boldsymbol{\Sigma}$, where w_n are appropriate norming constants.

The function u defined in section 2.2 does not depend on $\boldsymbol{\theta}$, so that the definition of g_n implies that the derivative of g_n with respect to $\boldsymbol{\theta}$ is deterministic and given by

$$\boldsymbol{\Delta}_n(\boldsymbol{\theta}) = \frac{\partial g_n(\mathbf{z}_n, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} = \frac{\partial}{\partial \boldsymbol{\theta}'} E_{\boldsymbol{\theta}}\left[u(\mathbf{X}(t_2)) \mid \mathbf{X}(t_1) = \mathbf{x}(t_1)\right]. \tag{14}$$

Denote the limit of $w_n \boldsymbol{\Delta}_n$ as $n \longrightarrow \infty$ by $\boldsymbol{\Delta}$, where g_n is of order $L \times 1$ and $\boldsymbol{\theta}'$ is of order $1 \times L$, so that $\boldsymbol{\Delta}$ is of order $L \times L$.

Last, partition g_n, $\boldsymbol{\Sigma}$, and $\boldsymbol{\Delta}$ in accordance with $\boldsymbol{\theta}' = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')$:

$$g_n(\mathbf{z}_n, \boldsymbol{\theta}) = \begin{pmatrix} g_{1n}(\mathbf{z}_n, \boldsymbol{\theta}) \\ g_{2n}(\mathbf{z}_n, \boldsymbol{\theta}) \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}' & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

$$\boldsymbol{\Delta}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{\Delta}_{11}(\boldsymbol{\theta}) & \boldsymbol{\Delta}_{12}(\boldsymbol{\theta}) \\ \boldsymbol{\Delta}_{21}(\boldsymbol{\theta}) & \boldsymbol{\Delta}_{22}(\boldsymbol{\theta}) \end{pmatrix}. \tag{15}$$

The test statistic

As was mentioned above, the test statistic is based on estimating function g_n. Since the statistics contained in g_n are commonly based on sums of weakly dependent random variables, it is plausible that

$$w_n^{1/2} g_n(\mathbf{Z}_n, \boldsymbol{\theta}) \xrightarrow{\text{d}} N_L(\mathbf{0}, \boldsymbol{\Sigma}) \text{ as } n \longrightarrow \infty \tag{16}$$

for the norming constants w_n introduced above, where $\xrightarrow{\mathrm{d}}$ denotes convergence in distribution, N_L refers to the L-variate Gaussian distribution, and $\mathbf{\Sigma}$ is non-singular. In the case of the Independent Arcs model (Snijders and van Duijn, 1997), where the arc variables $X_{ij}(t)$ follow independent Markov processes, the asymptotic normality can be proved directly from the Central Limit Theorem. However, in the general case where the arc variables $X_{ij}(t)$ are dependent and the Markov process is non-stationary, a rigorous proof of (16) is beyond the scope of the present paper (see section 7). The approximate normality is supported by some limited simulation studies (not further reported here); in the remainder of the section, it is assumed that (16) holds.

Let $\boldsymbol{\theta}_0' = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_{20}')$ be the parameter under $H_0 : \boldsymbol{\theta}_2 = \boldsymbol{\theta}_{20}$. To eliminate the impact of the estimated nuisance parameter $\boldsymbol{\theta}_1$ on the test, Neyman's (1959) orthogonalization method can be exploited, as suggested by Basawa (1985, 1991). Let

$$\mathrm{b}_n(\mathbf{z}_n, \boldsymbol{\theta}_0) = \mathrm{g}_{2n}(\mathbf{z}_n, \boldsymbol{\theta}_0) - \mathbf{\Gamma}(\boldsymbol{\theta}_0)\, \mathrm{g}_{1n}(\mathbf{z}_n, \boldsymbol{\theta}_0), \tag{17}$$

where $\mathbf{\Gamma} = \mathbf{\Delta}_{21}\mathbf{\Delta}_{11}^{-1}$ and $\mathbf{\Delta}_{11}$ is non-singular. Since both g_{2n} and $\mathbf{\Gamma}\,\mathrm{g}_{1n}$ have zero expectation by (13) and are asymptotically Gaussian distributed by (16), one obtains

$$\mathrm{w}_n^{1/2}\mathrm{b}_n(\mathbf{Z}_n, \boldsymbol{\theta}_0) \xrightarrow{\mathrm{d}} N_R(\mathbf{0}, \mathbf{\Xi}) \text{ as } n \longrightarrow \infty, \tag{18}$$

where the variance-covariance matrix $\mathbf{\Xi}$ is given by

$$\mathbf{\Xi} = \mathbf{\Sigma}_{22} - (\mathbf{\Sigma}_{12}'\,\mathbf{\Gamma}(\boldsymbol{\theta}_0)' + \mathbf{\Gamma}(\boldsymbol{\theta}_0)\,\mathbf{\Sigma}_{12}) + \mathbf{\Gamma}(\boldsymbol{\theta}_0)\,\mathbf{\Sigma}_{11}\,\mathbf{\Gamma}(\boldsymbol{\theta}_0)', \tag{19}$$

and R is the number of coordinates of $\boldsymbol{\theta}_2$. Thus, under H_0,

$$\mathrm{w}_n\,\mathrm{b}_n(\mathbf{Z}_n, \boldsymbol{\theta}_0)'\,\mathbf{\Xi}^{-1}\,\mathrm{b}_n(\mathbf{Z}_n, \boldsymbol{\theta}_0) \xrightarrow{\mathrm{d}} \chi_R^2 \text{ as } n \longrightarrow \infty \tag{20}$$

is asymptotically central chi-square distributed with R degrees of freedom.

The entities $\mathbf{\Delta}$ and $\mathbf{\Sigma}$ can be replaced by the constant matrices $w_n\mathbf{\Delta}_n$ and $w_n\mathbf{\Sigma}_n$, respectively; it is evident that these plug-ins do not change the asymptotic distribution

of (20). The unknown nuisance parameter $\boldsymbol{\theta}_1$ can be replaced by a consistent estimator under H_0 without changing the asymptotic distribution of (20). The parameter $\boldsymbol{\theta}_1$ can be estimated by the $\hat{\boldsymbol{\theta}}_1$ that solves $\mathrm{g}_n = \mathrm{g}_n(\mathbf{z}_n, \boldsymbol{\theta}_0) = \mathbf{0}$ (see section 2.2), where $\boldsymbol{\theta}_0' = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_{20}')$ and $\boldsymbol{\theta}_{20}$ is the value of $\boldsymbol{\theta}_2$ under H_0. It is plausible that $\hat{\boldsymbol{\theta}}_1$ is a consistent estimator of $\boldsymbol{\theta}_1$, although a rigorous proof is beyond the scope of the present paper (see section 7). Denote by C_n the test statistic (20) obtained by plugging in $\hat{\boldsymbol{\theta}}_1$, $\boldsymbol{\Delta}_n$, and $\boldsymbol{\Sigma}_n$ for $\boldsymbol{\theta}_1$, $\boldsymbol{\Delta}$, and $\boldsymbol{\Sigma}$, respectively.

The entities g_n, $\boldsymbol{\Delta}_n$, and $\boldsymbol{\Sigma}_n$ are not available in closed form, but can be estimated by Monte Carlo methods. Crude Monte Carlo estimation (Hammersley and Handscomb, 1964) of the expectations of which g_n is composed and of $\boldsymbol{\Sigma}_n$ is straightforward. Monte Carlo estimators of $\boldsymbol{\Delta}_n$ can be found in Schweinberger and Snijders (2005). These Monte Carlo estimators are simulation-consistent in the sense that the Monte Carlo estimators converge in probability to the desired quantities as the number of Monte Carlo simulations increases without bound. Therefore, the Monte Carlo estimator of test statistic C_n, which is obtained by plugging in the Monte Carlo estimators of g_n, $\boldsymbol{\Delta}_n$, and $\boldsymbol{\Sigma}_n$, is simulation-consistent.

Remarks and extensions

Observe that, to test restrictions on parameter vector $\boldsymbol{\theta}_2$, $\boldsymbol{\theta}_2$ needs not be estimated, as is generally the case for RS tests.

If $\boldsymbol{\theta}_2$ is a scalar, the test statistic (20) can be used both in its quadratic form, as presented above, and in its corresponding linear form

$$-\mathrm{w}_n^{1/2} \frac{\mathrm{b}_n(\mathbf{Z}_n, \boldsymbol{\theta}_0)}{\boldsymbol{\Xi}^{1/2}}, \tag{21}$$

where b_n and $\boldsymbol{\Xi}$ are scalars. The linear form is convenient when one-sided one-parameter tests are desired. The minus sign in (21) facilitates the interpretation in the sense that, if u_2 denotes the statistic corresponding to parameter $\boldsymbol{\theta}_2$ and its conditional expectations are non-decreasing functions of $\boldsymbol{\theta}_2$, then, by the definition of g_n

in (9), $\boldsymbol{\theta}_2 - \boldsymbol{\theta}_{20} > 0$ is associated with positive values of (21). By (18), the asymptotic distribution of (21) under $H_0 : \boldsymbol{\theta}_2 = \boldsymbol{\theta}_{20}$ is standard Gaussian.

Furthermore, tests with $R > 1$ degrees of freedom can be complemented with one degree of freedom tests, testing the restrictions one by one; two-sided one-parameter tests can be based on (20), while one-sided one-parameter tests can be based on (21).

Note that test statistic (20) has an appealing interpretation in terms of goodness-of-fit. Let $M = 2$ and observe that the test statistic is based on

$$g_{2n}(\mathbf{z}_n, \boldsymbol{\theta}) = E_{\boldsymbol{\theta}}\left[\mathrm{u}_2(\mathbf{X}(t_2)) \mid \mathbf{X}(t_1) = \mathbf{x}(t_1)\right] - \mathrm{u}_2(\mathbf{x}(t_2)), \qquad (22)$$

where u_2 is the part of the statistics vector u which corresponds to $\boldsymbol{\theta}_2$. In other words, the test statistic is based on the "distance" between the expected value of the function u of the data—evaluated under the model restricted by H_0—and the observed value of u. The argument extends to the case $M > 2$.

When the test indicates that there is empirical evidence against $H_0 : \boldsymbol{\theta}_2 = \boldsymbol{\theta}_{20}$, it may be desired to estimate $\boldsymbol{\theta}_2$. If g_n is differentiable at $\hat{\boldsymbol{\theta}}_0$, then, by definition (Magnus and Neudecker, 1988, p. 82),

$$\mathrm{g}_n(\mathbf{z}_n, \boldsymbol{\theta}) = \mathrm{g}_n(\mathbf{z}_n, \hat{\boldsymbol{\theta}}_0) + \boldsymbol{\Delta}_n(\hat{\boldsymbol{\theta}}_0)\,(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_0) + \mathrm{h}_{\hat{\boldsymbol{\theta}}_0}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_0), \qquad (23)$$

where the Jacobian matrix $\boldsymbol{\Delta}_n$ is given by (14), and

$$\lim_{(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_0) \longrightarrow \mathbf{0}} \frac{\mathrm{h}_{\hat{\boldsymbol{\theta}}_0}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_0)}{||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_0||} = \mathbf{0}. \qquad (24)$$

Thus, solving $\mathrm{g}_n = \mathrm{g}_n(\mathbf{z}_n, \boldsymbol{\theta}) = \mathbf{0}$ is asymptotically the same as solving

$$\mathrm{g}_n(\mathbf{z}_n, \hat{\boldsymbol{\theta}}_0) + \boldsymbol{\Delta}_n(\hat{\boldsymbol{\theta}}_0)\,(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_0) = \mathbf{0}, \qquad (25)$$

giving rise to the one-step estimator

$$\boldsymbol{\theta}^{\star} = \hat{\boldsymbol{\theta}}_0 - \boldsymbol{\Delta}_n(\hat{\boldsymbol{\theta}}_0)^{-1}\,\mathrm{g}_n(\mathbf{z}_n, \hat{\boldsymbol{\theta}}_0), \qquad (26)$$

13

where $\mathbf{\Delta}_n$ is non-singular. In general, one-step estimators can be useful as approximations of estimators which are hard to obtain (see Lehmann, 1999). Here, the one-step estimator $\boldsymbol{\theta}^\star$ is an approximation of the unrestricted moment estimator $\hat{\boldsymbol{\theta}}$, and is useful because all ingredients required to evaluate $\boldsymbol{\theta}^\star$ are available once the ingredients of (20) are available, while the estimation of $\hat{\boldsymbol{\theta}}$ requires an additional, time-consuming estimation run (see section 2.2). Note that $\boldsymbol{\theta}^{\star\prime} = (\boldsymbol{\theta}_1^{\star\prime}, \boldsymbol{\theta}_2^{\star\prime})$ is an estimator of both $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, and in practice both $\boldsymbol{\theta}_1^\star$ and $\boldsymbol{\theta}_2^\star$ are interesting, because in network models these estimators can be considerably correlated.

4 Model Misspecifications

Two classes of model misspecifications are briefly considered. First, the model is misspecified in the sense that the true parameter $\boldsymbol{\theta}$ contains other (additional) coordinates than the specified model. Sections 4.1 deals with such cases. Second, the model is misspecified in the sense that homogeneity assumptions regarding $\boldsymbol{\theta}$ do not hold. Such homogeneity assumptions are convenient for statistical and computational reasons. Two important cases are the assumption that

(1) $\boldsymbol{\beta}^{(1)} = \boldsymbol{\beta}^{(2)} = \cdots = \boldsymbol{\beta}^{(M-1)}$, where the superscript refers to the period, provided $M \geq 3$.

(2) $\boldsymbol{\beta}^{(1)} = \boldsymbol{\beta}^{(2)} = \cdots = \boldsymbol{\beta}^{(n)}$, where the superscript refers to the node.

Testing such homogeneity assumptions is considered in sections 4.2 and 4.3.

4.1 Misspecifications I

The simple case is considered where the specified objective function f_i by mistake excludes effects inherent to the true f_i.

Besides second- and third-order dependencies, covariates may have an impact on the digraph evolution. As an example, adolescents may favor friendship ties to adolescents of the same gender. Let c be some

(1) node-bound covariate c_i, containing information about node i: (a) demographic covariate; (b) behavioral or other covariate (such as smoking).

(2) dyadic covariate c_{ij}, containing information about the ordered pair of nodes (i, j): a frequently used form of dyadic covariates is "similarity" of nodes i and j with respect to node-bound covariates.

The objective function can be specified as

$$f_i(\boldsymbol{\beta}, \mathbf{x}) = \sum_{k=1}^{K} \beta_k \, s_{ik}(\mathbf{x}, j) + \beta_{K+1} \, s_{iK+1}(\mathbf{x}, c, j). \tag{27}$$

Some possible specifications of statistic s_{iK+1} are

(1) node-bound covariate: $\sum_{h=1}^{n} x_{ih} \, c_h$ or $c_i \sum_{h=1}^{n} x_{ih}$.

(2) dyadic covariate: $\sum_{h=1}^{n} x_{ih} \, c_{ih}$.

(3) interaction node-bound covariate $\times$ reciprocated arcs: $\sum_{h=1}^{n} x_{ih} \, x_{hi} \, c_h$.

Numerous other interactions between covariates and digraph statistics are conceivable. The covariate c may, or may not, depend on time. If c depends on time, it is—in the presence of panel data—convenient to assume that it is constant in time interval $[t_m, t_{m+1}]$, so that the value of $c(t_m)$ observed at time point t_m can be used to evaluate the corresponding statistic in time interval $[t_m, t_{m+1}]$.

A natural approach is to test the model restricted by

$$H_0 : \beta_{K+1} = 0 \tag{28}$$

against the unrestricted model

$$H_1 : \beta_{K+1} \neq 0. \tag{29}$$

4.2 Misspecifications II: Homogeneous Periods

An important assumption underlying the Snijders (2001) family of models is that $\boldsymbol{\beta}$ is constant across time intervals $[t_m, t_{m+1}]$, $m = 1, \ldots, M - 1$. To test whether the data-generating mechanism is indeed constant across time intervals, the model can be specified as

$$\mathrm{f}_i(\boldsymbol{\beta}, \mathbf{x}) = \sum_{k=1}^{K} \beta_k^{(m)} \, \mathrm{s}_{ik}(\mathbf{x}, j), \tag{30}$$

where parameter $\beta_k^{(m)}$ depends on time interval $[t_m, t_{m+1}]$. Observe that an equivalent formulation is obtained by including, for the $M - 1$ periods, $M - 2$ period-dependent dummy variables interacting with statistic s_{ik} (assuming that the main effect of s_{ik} is included), where the dummies are equivalent to node- and time-dependent covariates $c_i^{(m)}$.

To test whether coordinate β_k of $\boldsymbol{\beta}$ is constant across time intervals, the model restricted by

$$H_0 : \beta_k^{(1)} = \beta_k^{(2)} = \cdots = \beta_k^{(M-1)} \tag{31}$$

can be tested against the unrestricted model

$$H_1 : \beta_k^{(a)} \neq \beta_k^{(b)} \text{ for some } a \neq b, \tag{32}$$

where a and b indicate periods.

4.3 Misspecifications II: Homogeneous Nodes

Another assumption of the Snijders (2001) family of models is that $\boldsymbol{\beta}$ is constant across nodes i. Models with node-dependent parameters can be specified, which is equivalent to using $n - 1$ node-dependent dummy variables c_i interacting with statistic s_{ik} (assuming that the main effect of s_{ik} is included). Tests can be carried out as in section 4.2.

5 Monte Carlo Study

The present section studies the finite-sample behavior of the proposed test statistic (20) and the pseudo-t-test (10). For convenience, the two tests will be referred to as "score test" and "t-test", respectively, although the tests are not equivalent to the tests which are known in the statistical literature under these names.

Data are generated by simulating the Markov process with $n = 30$ ("small data set") and $n = 60$ nodes ("moderate data set") in time interval $[0, 2]$, starting with a real-world digraph at time point $t_1 = 0$ and "observing" the random digraph at time points $t_2 = 1$ and $t_3 = 2$. The rate function λ_i is constant and equal to ρ_m for period $m = 1, 2$, and the objective function is given by

$$f_i(\boldsymbol{\beta}, \mathbf{x}) = \sum_{k=1}^{K} \beta_k \, s_{ik}(\mathbf{x}, j). \tag{33}$$

The choice of statistics s_{ik} is based on their importance in empirical social science research; the chosen statistics are:

$s_{i1}(\mathbf{x}, j) = \sum_{h=1}^{n} x_{ih}$: the number of arcs,

$s_{i2}(\mathbf{x}, j) = \sum_{h=1}^{n} x_{ih} x_{hi}$: the number of reciprocated arcs,

$s_{i3}(\mathbf{x}, j) = \sum_{h,l=1}^{n} x_{ih} x_{hl} x_{il}$: the number of transitive triplets,

$s_{i4}(\mathbf{x}, j) = \sum_{h=1}^{n} (1 - x_{ih}) \max_{l} x_{il} x_{lh}$: the number of indirect connections,

$s_{i5}(\mathbf{x}, j) = \sum_{h=1}^{n} x_{ih} c_{ih}$: interaction of arcs and dyadic covariate c_{ih},

$s_{i6}(\mathbf{x}, j) = \sum_{h=1}^{n} x_{ih} d_h$: interaction of arcs and node-bound covariate d_h.

The values of the dyadic covariate c_{ih} are generated by independent draws from the Poisson(1) distribution; the values of the node-bound covariate d_h are generated by independent draws from the Bernoulli(1/2) distribution.

Two main purposes of goodness-of-fit testing can be distinguished, and hence the simulation study consists of two main parts: (1) testing parameters which capture structural features of the data such as third-order dependence among arcs ("clustering"), and (2) testing the impact of covariates on the digraph evolution. The basic data-generating model corresponds to parameters $\rho_1 = \rho_2 = 4$, $\beta_1 = -1$ (corresponding to s_{i1}), and $\beta_2 = 1$ (corresponding to s_{i2}), which are common to all models used for generating data. The first part of the simulation study tests hypotheses involving the parameters β_3 and β_4, corresponding to s_{i3} and s_{i4}, respectively; both parameters capture third-order dependence among arcs. The second part of the simulation study tests hypotheses involving the parameters β_5 and β_6, corresponding to s_{i5} and s_{i6}, respectively. All tests are two-sided.

Part I: testing effects capturing third-order dependence   The basic model is $P_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\rho_1, \rho_2, \beta_1, \beta_2, \beta_3, \beta_4)' = (4, 4, -1, 1, \beta_3, \beta_4)'$; three values of β_3 (0, .2, and .4) and β_4 (0, $-.3$, and $-.6$) are considered including all combinations, giving 9 models. The values of β_3 and β_4 are primarily chosen to study the power of the test against local alternatives. For each model, 500 data sets are generated; the reason for limiting the number of data sets to 500 is the computing time required to estimate models. For each model and each data set, the score test is evaluated in one estimation run (where the 4 unrestricted parameters are estimated) and the t-test in another estimation run (where all 6 parameters are estimated); the parameters are estimated by the conditional method of moments of Snijders (2001).

Figure 1 (p. 19) shows histograms of the distributions of the score test statistic for testing $H_0 : \beta_3 = \beta_4 = 0$, $H_0 : \beta_3 = 0$, and $H_0 : \beta_4 = 0$, and histograms of the distributions of the t-test statistic for testing $H_0 : \beta_3 = 0$ and $H_0 : \beta_4 = 0$, where the true, data-generating model is $P_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\rho_1, \rho_2, \beta_1, \beta_2, \beta_3, \beta_4)' = (4, 4, -1, 1, 0, 0)'$. For the score tests, the distributions agree very well with the chi-square distributions

Figure 1: Monte Carlo results, model $P_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\rho_1, \rho_2, \beta_1, \beta_2, \beta_3, \beta_4)' = (4, 4, -1, 1, 0, 0)'$, $n = 30$ and $n = 60$: distribution of test statistics
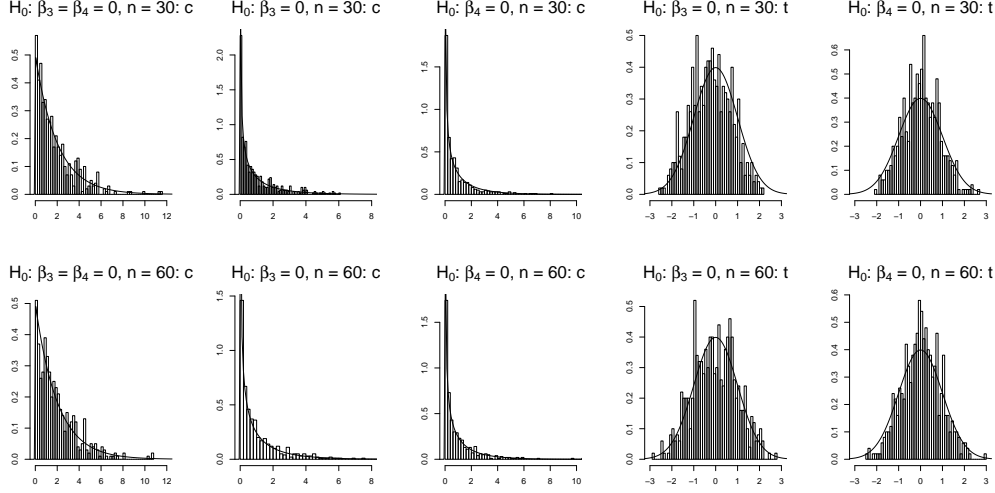
Note: the curves represent the expected distributions under H_0; c and t refer to score test (20) and t-test (10), respectively.

expected under H_0, but the t-tests appear to be slightly conservative.

Table 1 (p. 20) shows the empirical rejection probabilities for the mentioned hypotheses and the 9 data-generating models using the nominal significance level .05; space restrictions do not allow to show the results for other nominal significance levels (which do not alter the conclusions). Note that if H_0 is true and the test is of size .05, then the binomial distribution of the number of rejections implies that the empirical rejection rate should be roughly between .03 and .07.

The two-parameter score test behaves reasonable under $H_0 : \beta_3 = \beta_4 = 0$ and seems to have (for all practical purposes) sufficient power to detect departures from H_0.

Table 1 indicates that the one-parameter score test has considerably more power than the t-test, in particular if $n = 30$ and the departure from H_0 is small.

Concerning the one-parameter score test, note that the two-parameter test and the

Table 1: **Monte Carlo results, model** $P_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\rho_1, \rho_2, \beta_1, \beta_2, \beta_3, \beta_4)' = (4, 4, -1, 1, \beta_3, \beta_4)'$, $n = 30$ **and** $n = 60$**: empirical rejection probabilities for tests with nominal significance level** $.05$

true model:		score test:	score test:	score test:	t-test:	t-test:
	$H_0:$	$\beta_3 = \beta_4 = 0$	$\beta_3 = 0$	$\beta_4 = 0$	$\beta_3 = 0$	$\beta_4 = 0$
$P_{\boldsymbol{\theta}}, \beta_3 = 0, \beta_4 = 0$:	$n = 30$	.024	.050	.034	.032	.014
	$n = 60$	.036	.054	.044	.048	.024
$P_{\boldsymbol{\theta}}, \beta_3 = .2, \beta_4 = 0$:	$n = 30$	.446	.540	.136	.422	.026
	$n = 60$	.744	.812	.128	.788	.030
$P_{\boldsymbol{\theta}}, \beta_3 = .4, \beta_4 = 0$:	$n = 30$	.984	.990	.664	.980	.026
	$n = 60$	1.000	1.000	.650	1.000	.038
$P_{\boldsymbol{\theta}}, \beta_3 = 0, \beta_4 = -.3$:	$n = 30$	.492	.078	.596	.018	.458
	$n = 60$	.972	.140	.988	.052	.972
$P_{\boldsymbol{\theta}}, \beta_3 = 0, \beta_4 = -.6$:	$n = 30$	.892	.180	.940	.014	.862
	$n = 60$	1.000	.472	1.000	.030	1.000
$P_{\boldsymbol{\theta}}, \beta_3 = .2, \beta_4 = -.3$:	$n = 30$	.816	.512	.800	.220	.612
	$n = 60$	1.000	.832	.998	.506	.990
$P_{\boldsymbol{\theta}}, \beta_3 = .4, \beta_4 = -.3$:	$n = 30$	.992	.978	.976	.870	.786
	$n = 60$	1.000	1.000	1.000	.986	.990
$P_{\boldsymbol{\theta}}, \beta_3 = .2, \beta_4 = -.6$:	$n = 30$	.982	.658	.990	.162	.944
	$n = 60$	1.000	.946	1.000	.316	1.000
$P_{\boldsymbol{\theta}}, \beta_3 = .4, \beta_4 = -.6$:	$n = 30$	1.000	.952	1.000	.574	.990
	$n = 60$	1.000	1.000	1.000	.900	1.000

Figure 2: Monte Carlo results, model $P_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\rho_1, \rho_2, \beta_1, \beta_2, \beta_3, \beta_5, \beta_6)' = (4, 4, -1, 1, .2, 0, 0)'$, $n = 30$ and $n = 60$: distribution of test statistics
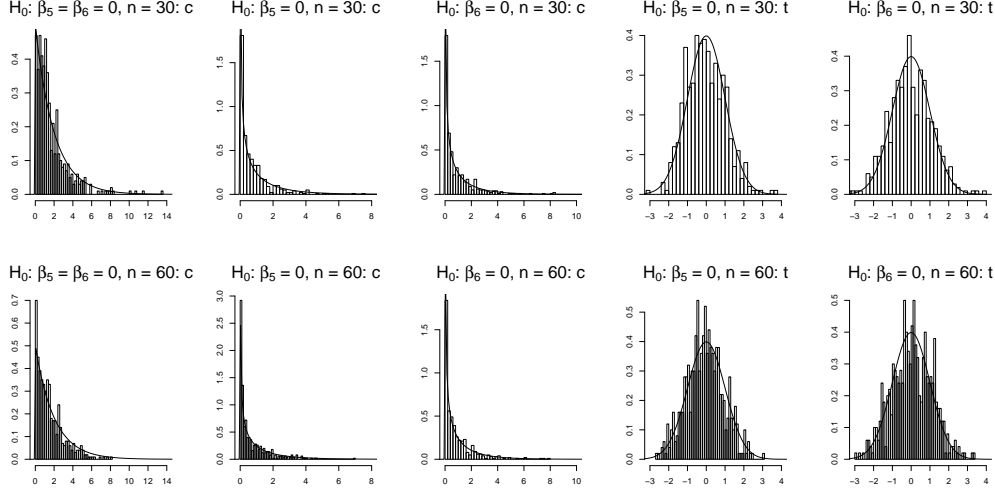
Note: the curves represent the expected distributions under H_0; c and t refer to score test (20) and t-test (10), respectively.

two corresponding one-parameter tests are computed from the same estimation run, where $\boldsymbol{\theta}$ was estimated under $H_0 : \beta_3 = \beta_4 = 0$. Therefore, the one-parameter score tests in table 1 do not control for the other parameter (β_3 or β_4). Note that, in principle, it is straightforward to carry out one-parameter score tests of β_3 and β_4 where the other parameter is controlled for, but that would require two additional estimation runs.

Concerning the t-test, it is notable that the power of the t-test for testing $H_0 : \beta_3 = 0$ is the smaller the larger the departure of β_4 from 0.

Part II: testing covariate-related effects The basic model is $P_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\rho_1, \rho_2, \beta_1, \beta_2, \beta_3, \beta_5, \beta_6)' = (4, 4, -1, 1, .2, \beta_5, \beta_6)'$, which includes transitivity parameter β_3; three values of β_5 (0, .1, and .2) and β_6 (0, .2, and .4) are considered including all combinations, giving 9 models.

Figure 2 (p. 21) shows histograms of the distributions of the score test statistic for testing $H_0 : \beta_5 = \beta_6 = 0$, $H_0 : \beta_5 = 0$, and $H_0 : \beta_6 = 0$, and histograms of the distributions of the t-test statistic for testing $H_0 : \beta_5 = 0$ and $H_0 : \beta_6 = 0$, where the true, data-generating model is $P_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\rho_1, \rho_2, \beta_1, \beta_2, \beta_3, \beta_5, \beta_6)' = (4, 4, -1, 1, .2, 0, 0)'$. For the score tests, the distributions agree quite well with the chi-square distributions expected under H_0, although the two-parameter score test appears to be slightly conservative.

Table 2 (p. 23) shows the empirical rejection probabilities for the mentioned hypotheses and the 9 data-generating models using the nominal significance level .05. Table 2 indicates that the one-parameter score test has less power than the t-test, in particular if $n = 30$ and the departure from H_0 is small.

6 Application

Pahor (2003) studied the directed cross-ownerships among 413 business firms in Slovenia (EU) observed at 5 time points, where directed cross-ownership $A \longrightarrow B$ means that firm A holds stock market shares of firm B. The present paper re-analyzes the data in the light of the new goodness-of-fit test statistic. Once again, it is convenient to refer to the two tests as "score test" and "t-test", respectively, and all tests are two-sided, unless stated otherwise.

The baseline model considered here corresponds to the rate function

$$\lambda_i(\boldsymbol{\theta}, \mathbf{x}^\star) = \rho_m \exp\left[\alpha \left(1 + \sum_{h=1}^{n} x_{ih}^\star\right)^{-1}\right] \tag{34}$$

for time interval $[t_m, t_{m+1}]$. The objective function is given by

$$\mathrm{f}_i(\boldsymbol{\beta}, \mathbf{x}) = \sum_{k=1}^{K} \beta_k \, \mathrm{s}_{ik}(\mathbf{x}, j), \tag{35}$$

where

Table 2: **Monte Carlo results, model** $P_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\rho_1, \rho_2, \beta_1, \beta_2, \beta_3, \beta_5, \beta_6)' = (4, 4, -1, 1, .2, \beta_5, \beta_6)'$, $n = 30$ **and** $n = 60$: **empirical rejection probabilities for tests with nominal significance level** .05

true model:		score test:	score test:	score test:	t-test:	t-test:
	$H_0:$	$\beta_5 = \beta_6 = 0$	$\beta_5 = 0$	$\beta_6 = 0$	$\beta_5 = 0$	$\beta_6 = 0$
$P_{\boldsymbol{\theta}}, \beta_5 = 0, \beta_6 = 0:$	$n = 30$	.024	.024	.030	.044	.068
	$n = 60$	.012	.014	.032	.058	.062
$P_{\boldsymbol{\theta}}, \beta_5 = .1, \beta_6 = 0:$	$n = 30$	.182	.250	.044	.326	.078
	$n = 60$	.394	.526	.038	.640	.056
$P_{\boldsymbol{\theta}}, \beta_5 = .2, \beta_6 = 0:$	$n = 30$	.690	.780	.034	.846	.058
	$n = 60$	.968	.988	.046	.994	.084
$P_{\boldsymbol{\theta}}, \beta_5 = 0, \beta_6 = .2:$	$n = 30$	.152	.034	.220	.068	.326
	$n = 60$	.336	.030	.512	.066	.626
$P_{\boldsymbol{\theta}}, \beta_5 = 0, \beta_6 = .4:$	$n = 30$	.588	.020	.696	.070	.808
	$n = 60$	.942	.022	.974	.056	.982
$P_{\boldsymbol{\theta}}, \beta_5 = .1, \beta_6 = .2:$	$n = 30$	.344	.276	.190	.374	.292
	$n = 60$	.660	.530	.462	.632	.596
$P_{\boldsymbol{\theta}}, \beta_5 = .1, \beta_6 = .4:$	$n = 30$	.734	.262	.728	.378	.822
	$n = 60$	.980	.500	.982	.616	.992
$P_{\boldsymbol{\theta}}, \beta_5 = .2, \beta_6 = .2:$	$n = 30$	.790	.792	.232	.858	.318
	$n = 60$	.996	.992	.482	.998	.600
$P_{\boldsymbol{\theta}}, \beta_5 = .2, \beta_6 = .4:$	$n = 30$	.932	.812	.712	.886	.818
	$n = 60$	1.000	.984	.984	.994	.990

Table 3: Pahor data: testing restricted models

nuisance parameters	test	score test	d.f.	p-value
$\boldsymbol{\eta}$	$H_0 : \beta_5 = \beta_6 = 0$	61.95	2	$< .0001$
$\boldsymbol{\eta}$	$H_0 : \beta_5 = 0$	61.82	1	$< .0001$
$\boldsymbol{\eta}$	$H_0 : \beta_6 = 0$	$< .01$	1	.9905
$\boldsymbol{\eta}, \beta_5$	$H_0 : \beta_7 = \beta_8 = \beta_9 = 0$	127.59	3	$< .0001$
$\boldsymbol{\eta}, \beta_5$	$H_0 : \beta_7 = 0$	43.38	1	$< .0001$
$\boldsymbol{\eta}, \beta_5$	$H_0 : \beta_8 = 0$	70.34	1	$< .0001$
$\boldsymbol{\eta}, \beta_5$	$H_0 : \beta_9 = 0$	14.04	1	.0002
$\boldsymbol{\eta}, \beta_5, \beta_7, \beta_8, \beta_9$	$H_0 : \beta_{10} = \beta_{11} = \beta_{12} = 0$	20.19	3	.0002
$\boldsymbol{\eta}, \beta_5, \beta_7, \beta_8, \beta_9$	$H_0 : \beta_{10} = 0$	10.37	1	.0013
$\boldsymbol{\eta}, \beta_5, \beta_7, \beta_8, \beta_9$	$H_0 : \beta_{11} = 0$	4.60	1	.0319
$\boldsymbol{\eta}, \beta_5, \beta_7, \beta_8, \beta_9$	$H_0 : \beta_{12} = 0$	13.83	1	.0002

The parameter vector $\boldsymbol{\eta}$ is given by $\boldsymbol{\eta} = (\boldsymbol{\rho}', \alpha, \beta_1, \beta_2, \beta_3, \beta_4)'$.

$s_{i1}(\mathbf{x}, j) = \sum_{h=1}^{n} x_{ih}$: the number of arcs,

$s_{i2}(\mathbf{x}, j) = \sum_{h=1}^{n} x_{ih} x_{hi}$: the number of reciprocated arcs,

$s_{i3}(\mathbf{x}, j) = \sum_{h=1}^{n} x_{ih} d_{h1}$: interaction of arcs and covariate d_{h1},

$s_{i4}(\mathbf{x}, j) = \sum_{h=1}^{n} x_{ih} d_{h2}$: interaction of arcs and covariate d_{h2},

where d_{h1} indicates whether or not firm h was quoted on the stock exchange, while d_{h2} refers to the size of firm h. The parameter $\boldsymbol{\theta}$ of the baseline model $P_{\boldsymbol{\theta}}$ corresponds to $\boldsymbol{\theta} = (\boldsymbol{\rho}', \alpha, \beta_1, \beta_2, \beta_3, \beta_4)'$, where $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3, \rho_4)'$.

Pahor (2003) suspected that the data may exhibit third-order dependence, leading to the transitivity parameter β_5 and the indirect connections parameter β_6 as candidates to be tested, which correspond to statistics $\sum_{h,l=1}^{n} x_{ih} x_{hl} x_{il}$ and $\sum_{h=1}^{n} (1 - x_{ih}) \max_{l} x_{il} x_{lh}$, respectively. According to table 3 (p. 24), the two-parameter score test

of $H_0 : \beta_5 = \beta_6 = 0$ clearly indicates that parameters capturing third-order dependence are required to improve the goodness-of-fit of the model, and the one-parameter score tests suggest to include β_5 but not β_6. A one-sided test of $H_0 : \beta_5 = 0$ can be carried out by using the linear form of the score test (see (21)), giving 7.86: using the asymptotic standard Gaussian distribution of the linear form under $H_0 : \beta_5 = 0$, it seems that β_5 is positive, which is supported by the one-step estimate of β_5 (see (26)) given by .996. As a result, β_5 is henceforth included in the model, while β_6 is not.

Pahor (2003) argued that firms tend to hold shares of other firms close to them with respect to region (c_{ih1}) and industry branch (c_{ih2}), and to other firms having the same owner (c_{ih3}). Three parameters are added, β_7, β_8, and β_9, corresponding to statistics $\sum_{h=1}^{n} x_{ih} c_{ihl}$, $l = 1, 2, 3$, respectively. The three-parameter score test of $H_0 : \beta_7 = \beta_8 = \beta_9 = 0$ and the three corresponding one-parameter tests, shown in table 3, suggest to add all three covariates to the model, which is done.

It is interesting to test whether the values of the parameters are constant across time intervals. A basic parameter for which such homogeneity tests frequently make sense is the "outdegree" parameter β_1, corresponding to statistic s_{i1}. A homogeneity test for β_1 is conducted by testing $H_0 : \beta_{10} = \beta_{11} = \beta_{12} = 0$, where β_{10}, β_{11}, and β_{12} correspond to statistics $e_i^{(m)} s_{i1}$, $m = 2, 3, 4$, respectively, where $e_i^{(m)}$ is a period-dependent dummy variable with value 1 in time interval $[t_m, t_{m+1}]$ for all i and 0 otherwise. According to table 3, the three-parameter score test of $H_0 : \beta_{10} = \beta_{11} = \beta_{12} = 0$ indicates that there is empirical evidence against H_0, and the three corresponding one-parameter tests suggest that it is sensible to add β_{10}, β_{11}, and β_{12} to the model. The one-sided one-parameter score test statistics (see (21)) are -3.22, -2.15, and 3.72, respectively, suggesting that the values of β_{10} and β_{11} are negative while β_{12} is positive. Table 4 (p. 26) gives the one-step estimate (see (26)) of the parameter $\boldsymbol{\theta} = (\boldsymbol{\rho}', \alpha, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_7, \beta_8, \beta_9, \beta_{10}, \beta_{11}, \beta_{12})'$ based on the moment estimate of $\boldsymbol{\theta}$ under the model restricted by $H_0 : \beta_{10} = \beta_{11} = \beta_{12} = 0$, and shows in addition the

Table 4: Pahor data, model $P_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\boldsymbol{\rho}', \alpha, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_7, \beta_8, \beta_9, \beta_{10}, \beta_{11}, \beta_{12})'$: estimates, standard errors, and t-tests

	one-step estimate $\boldsymbol{\theta}^{\star}$	moment estimate $\hat{\boldsymbol{\theta}}$	s.e.$(\hat{\boldsymbol{\theta}})$	t-test	p-value
β_1	-2.65	-2.67	.097	-27.54	$< .0001$
β_5	.81	.81	.102	7.91	$< .0001$
β_7	.57	.58	.076	7.60	$< .0001$
β_8	1.01	1.01	.106	9.57	$< .0001$
β_9	.89	.89	.258	3.45	.0006
β_{10}	$-.31$	$-.29$	.134	-2.12	.0337
β_{11}	$-.31$	$-.28$	.158	-1.79	.0735
β_{12}	.12	.13	.117	1.14	.2544

To save space, the nuisance parameters $\alpha, \boldsymbol{\rho}$, β_2, β_3, and β_4 are omitted. The one-step estimate $\boldsymbol{\theta}^{\star}$ is based on the moment estimate of $\boldsymbol{\theta}$ under the model restricted by $H_0 : \beta_{10} = \beta_{11} = \beta_{12} = 0$; the moment estimate $\hat{\boldsymbol{\theta}}$ is the unrestricted moment estimate of $\boldsymbol{\theta}$.

unrestricted moment estimate of $\boldsymbol{\theta}$, including standard errors and t-tests. The one-step estimates roughly agree with the unrestricted moment estimates. The t-tests by and large agree with the score tests (regarding β_5, β_7, β_8, β_9, and β_{10}), but slightly disagree about β_{11} and clearly disagree about β_{12}.

7   Discussion

A new goodness-of-fit test statistic was proposed, which has many applications, and, in contrast to the pseudo-t-test, does not require the estimation of the parameters to be tested (saving computation time and avoiding convergence problems), admits multi-parameter tests, and has an appealing interpretation in terms of goodness-of-fit.

The Monte Carlo study indicated that the finite-sample distribution of the goodness-of-fit test statistic under the considered null hypotheses matches the expected distri-

butions fairly well.

As to the unproven assertion (16), observe that the statistical modeling of social networks is in its infancy, and that asymptotic properties of estimators and test statistics are largely unknown; see, for instance, Hunter and Handcock (2006). In addition, asymptotics—despite of being an indispensable tool in mathematical statistics and useful guides in applied statistics—strictly speaking make no statement about samples of given finite size (De Bruijn, 1981, p. 2, Hampel, 1997). In combination with the Monte Carlo study, the approach in the present paper can be recommended for practical use, though care must be taken.

The proposed test statistic is implemented in the Windows-based computer program `Siena` embedded in the program collection `StOCNET`, which can be downloaded free of charge from `http://stat.gamma.rug.nl/stocnet`.

References

Basawa, I. V. (1985). Neyman-Le Cam tests based on estimation functions. In L. M. Le Cam and R. A. Ohlson (Eds.), *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer*, Volume II, pp. 811–825. Wadsworth.

Basawa, I. V. (1991). Generalized score tests for composite hypotheses. In V. P. Godambe (Ed.), *Estimating Functions*, Chapter 8, pp. 121–131. Oxford: Oxford Science Publications.

Bera, A. K. and Y. Bilias (2001a). On some optimality properties of Fisher-Rao score function in testing and estimation. *Communications in Statistics—Theory and Methods 30*, 1533–1559.

Bera, A. K. and Y. Bilias (2001b). Rao's score, Neyman's $C(\alpha)$ and Silvey's LM tests: an essay on historical developments and some new results. *Journal of Statistical Planning and Inference 97*, 9–44.

De Bruijn, N. G. (1981). *Asymptotic Methods in Analysis.* New York: Dover.

Frank, O. (1991). Statistical analysis of change in networks. *Statistica Neerlandica 45*, 283–293.

Godambe, V. P. (1960). An optimum property of regular maximum likelihood estimation. *The Annals of Mathematical Statistics 31*, 1208–1211.

Godfrey, L. G. (1988). *Misspecification Tests in Econometrics. The Lagrange Multiplier Principle and Other Approaches.* Cambridge: Cambridge University Press.

Hall, W. J. and D. J. Mathiason (1990). On large-sample estimation and testing in parametric models. *International Statistical Review 58*, 77–97.

Hammersley, J. M. and D. C. Handscomb (1964). *Monte Carlo methods.* London: Methuen.

Hampel, F. (1997). Is statistics too difficult? Research Report 81, Eidgenössische Technische Hochschule, Zürich. To download from http://stat.ethz.ch/Research-Reports/81.pdf.

Healy, M. J. R. (1978). Is statistics a science? *Journal of the Royal Statistical Society, Series A 141*, 385–393.

Holland, P. W. and S. Leinhardt (1970). A method for detecting structure in sociometric data. *American Journal of Sociology 76*, 492–513.

Holland, P. W. and S. Leinhardt (1976). Local structure in social networks. *Sociological Methodology*, 1–45.

Holland, P. W. and S. Leinhardt (1977). A dynamic model for social networks. *Journal of Mathematical Sociology 5*, 5–20.

Hunter, D. R. and M. S. Handcock (2006). Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics*. In press.

Lehmann, E. L. (1999). *Elements of Large Sample Theory.* New York: Springer.

Magnus, J. R. and H. Neudecker (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* New York: Wiley.

Neyman, J. (1959). Optimal asymptotic tests of composite statistical hypotheses. In U. Grenander (Ed.), *Probability and Statistics. The Harald Cramér Volume*, pp. 213–234. Stockholm/New York: Almquist & Wiksell/Wiley.

Pahor, M. (2003). *Causes and Consequences of Companies' Activity in Ownership Network.* Ph. D. thesis, Faculty of Economics, University of Ljubljana, Slovenia.

Pearson, K. (1900). On the criterion that a given system of observations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine 50*, 157–175.

Pearson, K. (1902a). On the systematic fitting of curves to observations and measurements. *Biometrika 1*(3), 256–303.

Pearson, K. (1902b). On the systematic fitting of curves to observations and measurements II. *Biometrika 2*(1), 1–23.

Rao, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with

applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society 44*, 50–57.

Rao, C. R. (2002). Karl Pearson chi-square test. The dawn of statistical inference. In C. Huber-Carol, N. Balakrishnan, M. S. Nikulin, and M. Mesbah (Eds.), *Goodness-of-fit Tests and Model Validity*, Chapter 2, pp. 9–24. Boston: Birkhäuser.

Rao, C. R. and S. J. Poti (1946). On locally most powerful tests when alternatives are one sided. *Sankhyā 7*, 439.

Robbins, H. and S. Monro (1951). A stochastic approximation method. *Annals of Mathematical Statistics 22*, 400–407.

Schweinberger, M. and T. A. B. Snijders (2005). Markov models for digraph panel data: Monte Carlo-based gradient estimation. Submitted.

Snijders, T. A. B. (1996). Stochastic actor-oriented dynamic network analysis. *Journal of Mathematical Sociology 21*, 149–172.

Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. In M. Sobel and M. Becker (Eds.), *Sociological Methodology*, pp. 361–395. Boston and London: Basil Blackwell.

Snijders, T. A. B. (2003). Accounting for degree distributions in empirical analysis of network dynamics. In R. Breiger, K. Carley, and P. Pattison (Eds.), *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, pp. 146–161. Washington, DC: National Academies Press.

Snijders, T. A. B., P. E. Pattison, G. L. Robins, and M. S. Handcock (2006). New specifications for exponential random graph models. *Sociological Methodology*. In press.

Snijders, T. A. B., C. E. G. Steglich, and M. Schweinberger (2006). Modeling the co-evolution of networks and behavior. In K. Montfort, H. Oud, and A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences*. Lawrence Erlbaum. In press.

Snijders, T. A. B. and M. A. J. van Duijn (1997). Simulation for statistical inference in dynamic network models. In R. Conte, R. Hegselmann, and P. Terna (Eds.), *Simulating social phenomena*, pp. 493–512. Berlin: Springer.

Wasserman, S. (1979). A stochastic model for directed graphs with transition rates determined by reciprocity. In K. F. Schuessler (Ed.), *Sociological Methodology*, pp. 392 – 412. San Francisco, CA: Jossey-Bass.

Wasserman, S. (1980). Analyzing social networks as stochastic processes. *Journal of the American Statistical Association 75*, 280 – 294.

Wasserman, S. and K. Faust (1994). *Social Network Analysis: Methods and Applications*. Cambridge:

Cambridge University Press.