

## Assessing and accounting for time heterogeneity in stochastic actor oriented models

Joshua A. Lospinoso · Michael Schweinberger ·  
Tom A. B. Snijders · Ruth M. Ripley

Received: 7 March 2010 / Revised: 21 September 2010 / Accepted: 8 October 2010 /  
Published online: 5 November 2010  
© Springer-Verlag 2010

**Abstract** This paper explores time heterogeneity in stochastic actor oriented models (SAOM) proposed by Snijders (Sociological methodology. Blackwell, Boston, pp 361–395, 2001) which are meant to study the evolution of networks. SAOMs model social networks as directed graphs with nodes representing people, organizations, etc., and dichotomous relations representing underlying relationships of friendship, advice, etc. We illustrate several reasons why heterogeneity should be statistically

---

This research was funded in part by U.S. Army Project Number 611102B74F and MIPR Number 9FDATXR048 (JAL); by U.S. N.I.H. (National Institutes of Health) Grant Number 1R01HD052887-01A2, for the project *Adolescent Peer Social Network Dynamics and Problem Behavior* (TABS and RMR); and by U.S. N.I.H. (National Institutes of Health) Grant Number 1R01GM083603-01 (MS).

The authors are grateful to Professor Laurence Moore of the Cardiff Institute for Society, Health and Ethics (CISHE) for the permission to use the ASSIST data and to Christian Steglich for his help with using these data.

---

J. A. Lospinoso (✉) · T. A. B. Snijders · R. M. Ripley  
Department of Statistics, University of Oxford, Oxford, UK  
e-mail: lospinos@stats.ox.ac.uk

T. A. B. Snijders  
e-mail: snijders@stats.ox.ac.uk

R. M. Ripley  
e-mail: ruth@stats.ox.ac.uk

J. A. Lospinoso  
Network Science Center, United States Military Academy, New York, USA

M. Schweinberger  
Department of Statistics, Pennsylvania State University, University Park, USA  
e-mail: michael.schweinberger@stat.psu.edu

T. A. B. Snijders  
Department of Sociology, University of Groningen, Groningen, The Netherlands

tested and provide a fast, convenient method for assessment and model correction. SAOMs provide a flexible framework for network dynamics which allow a researcher to test selection, influence, behavioral, and structural properties in network data over time. We show how the forward-selecting, score type test proposed by Schweinberger (Chapter 4: Statistical modeling of network panel data: goodness of fit. PhD thesis, University of Groningen 2007) can be employed to quickly assess heterogeneity at almost no additional computational cost. One step estimates are used to assess the magnitude of the heterogeneity. Simulation studies are conducted to support the validity of this approach. The ASSIST dataset (Campbell et al. In Lancet 371(9624):1595–1602, 2008) is reanalyzed with the score type test, one step estimators, and a full estimation for illustration. These tools are implemented in the RSiena package, and a brief walkthrough is provided.

**Keywords** Stochastic actor oriented models · Longitudinal analysis of network data · Time heterogeneity · Score-type test

**Mathematics Subject Classification (2000)** 62M02

## 1 Introduction

Social networks are relational structures between social actors. They evolve according to potentially complex and dynamic rules. Behavioral variables and static characteristics, the changing network structure itself, and environmental factors may all weigh in on how social networks and their component relationships and actors evolve over time. Understanding this process has become an area of increasing importance among researchers (see, e.g., the three special issues in the *Journal of Mathematical Sociology* edited by Doreian and Stokman 1996, 2001, 2003, literature on the evolution of friendship networks in Pearson and Michell 2000; Burk et al. 2007, on organizational networks in Borgatti and Foster 2003; Brass et al. 2004).

Increasingly, network data is being collected over time, which gives information about rules for network dynamics that are responsible for changes in network structure. Stochastic actor oriented models (SAOM) proposed by Snijders (2001) and elaborated in Snijders et al. (2010a) provide a flexible framework for statistical modeling of longitudinal network data. Social networks are modeled as digraphs in a Markov chain evolving in continuous time. This continuous time Markov chain (CTMC) changes from one state to the next by modifying a single link at a time. One possible interpretation of this setup is that actors are selected to update an outgoing link according to probabilities resulting from myopic stochastic optimization (see the econometric literature on discrete choice modeling, e.g., Maddala 1983; McFadden 1973). The model is explained in detail in Sect. 2.

Analysis of longitudinal data observed over multiple time periods permits researchers to make inference about the rules governing network evolution; however, the temporal component of panel data may make time heterogeneity an important issue. There are at least two plausible reasons that researchers might neglect time heterogeneity in SAOMs. First, the prospect of including numerous parameters to capture time heterogeneity when such heterogeneity is not a component of the research question is

an onerous and time-consuming process. To address this concern, we implement the score-type test of [Schweinberger \(2007\)](#). This allows us to create statistical tests for time heterogeneity without additional computationally intensive estimation runs. Second, it is not well studied in the literature under what circumstances omission of such heterogeneity leads to erroneous conclusions. This paper addresses the potential consequences of omitting time heterogeneity and provides an approach for assessing and accounting for it.

### 1.1 Motivation for assessing time heterogeneity

The most important consideration in formulating a SAOM is the set of research questions we wish to answer. Let us consider a partitioning of the statistical parameters included in the SAOM,  $\theta = (\psi, \nu)$ , where  $\psi$  are the parameters of interest and  $\nu$  are so-called nuisance parameters. Accordingly,  $\psi$  is formulated in such a way that it is relevant to the research questions at hand. Keeping in mind that the correctness of inferences about  $\psi$  is the principal motivation for considering time heterogeneity, there are at least three important reasons why it should be assessed:

1.  *$\psi$  is time heterogeneous*: Simulation results given later in the article indicate that incorrectly homogeneous specifications will result in estimates that average over the heterogeneity. Since parameters of interest are intended to answer research questions, detecting heterogeneity in  $\psi$  will at least be intrinsically interesting, and sometimes may suggest that some important explanatory covariate has been erroneously omitted. Disruptions in actor behavior may be crucial in certain settings. For example, drawing inference on time heterogeneity will help with understanding periods of disrupted behavior in illicit networks, cooperation networks for natural disasters, and planned restructuring of organizations.
2. *Undetected heterogeneity may lead to bias in other parameters*: It is currently unknown in the literature whether time heterogeneity in  $\nu$  can have serious consequences for inferences about  $\psi$ . Accordingly, it is prudent to formulate models which take into consideration time heterogeneity in all of the parameters in the model.
3. *Asymptotic degeneracy of the model*: SAOMs are based on a right-continuous process on the set of all possible networks, and the unique, limiting distribution of a continuous-time Markov process as  $t \rightarrow \infty$  under time-homogeneous parameterizations may be near-degenerate in the sense of placing much—sometimes almost all—probability mass on a small number of graphs which do not resemble real-world networks. [Strauss \(1986\)](#), [Snijders \(2002\)](#), [Handcock \(2003\)](#) show that exponential random graph models can be near-degenerate, and the same may hold for SAOMs if time runs on infinitely long (although in practice, time usually is limited). While the limiting distribution in itself is rarely of interest, this suggests that statistical inference can be affected when the amount of change between consecutive observations is large.

Since it is not possible to determine a priori that these cases do not apply, the study of time heterogeneity is motivated anywhere it could exist—namely whenever a dataset contains more than two time periods.

We proceed with a review on SAOM in Sect. 2. We then review the score type test proposed by [Schweinberger \(2007\)](#) and develop a specific test for time heterogeneity in Sect. 3. A simulation study follows to explore the validity of the score type approach for detecting time heterogeneity in Sect. 4. The paper culminates with a case study on a dataset collected by [Audrey et al. \(2004\)](#) with suspected time heterogeneity in Sect. 5. This study also provides an opportunity to demonstrate functionality implementing the aforementioned score type test now available in the RSiena package (see [Ripley and Snijders 2009](#)) in Sect. 5.1.

## 2 Stochastic actor oriented models (SAOM)

A social network composed of  $n$  actors is modeled as a directed graph (digraph), represented by an adjacency matrix  $(x_{ij})_{n \times n}$ , where  $x_{ij} = 1$  if actor  $i$  is tied to actor  $j$ ,  $x_{ij} = 0$  if  $i$  is not tied to  $j$ , and  $x_{ii} = 0$  for all  $i$  (self ties are not permitted). It is assumed here that the social network evolves in continuous time over an interval  $\mathcal{T} \subset \mathbb{R}^1$  according to a Markov process. Accordingly, the digraph  $x(t)$  models the state of social relationships at time  $t \in \mathcal{T}$ . Changes to the network called *updates*, occur at discrete time points defining the set  $\mathcal{L} \subset \mathcal{T}$ . Elements of the set are denoted  $L_a$  with consecutive natural number indices  $a$  so that  $L_1 < L_2 < \dots < L_{|\mathcal{L}|}$ , where the notation  $|\cdot|$  is used to denote the number of elements in a set. The network is observed at discrete time points called *observations* defining the set  $\mathcal{M}$  with elements  $M_a$  indexed similarly with consecutive natural number indices  $a$  so that  $M_1 < M_2 < \dots < M_{|\mathcal{M}|}$ . Define a set of *periods* with elements  $W_a \in \mathcal{W}$  each representing the continuous time interval between two consecutive observations  $M_a$  and  $M_{a+1}$ :

$$W_a = [M_a, M_{a+1}] = \{t \in \mathcal{T} : M_a \leq t \leq M_{a+1}\}.$$

By definition,  $|\mathcal{W}| = |\mathcal{M}| - 1$ . When  $\mathcal{L} = \mathcal{M}$ , we have *full information* on the network updates over the interval  $\mathcal{T}$ . We use upper case to denote random variables (e.g.  $X, M, L$ ).

There is a variety of models proposed in the literature for longitudinally observed social networks. In this paper, we consider the approach proposed by [Snijders \(2001\)](#), the stochastic actor oriented model (SAOM). Here, the stochastic process  $\{X(t) : t \in \mathcal{T}\}$  with digraphs as outcomes is modeled as a Markov process so that for any time  $t_a \in \mathcal{T}$ , the conditional distribution for the future  $\{X(t) : t > t_a\}$  given the past  $\{X(t) : t \leq t_a\}$  depends only on  $X(t_a)$ .

From the general theory of continuous-time Markov chains ([Norris 1997](#)) follows the existence of the *intensity matrix* that describes the rate at which  $X(t) = x$  tends to transition into  $\tilde{X}(t + dt) = \tilde{x}$  as  $dt \rightarrow 0$ :

$$q(x, \tilde{x}) = \lim_{dt \downarrow 0} \frac{P\{X(t + dt) = \tilde{x} \mid X(t) = x\}}{dt} \quad (\tilde{x} \neq x), \quad (1)$$

where  $\tilde{x} \in \mathcal{X}$ . The SAOM supposes that a digraph update consists of exactly one tie variable change. Such a change is referred to as a *ministep*. This property can be

expressed as

$$q(x, \tilde{x}) > 0 \Rightarrow \sum_{i,j} \text{abs}(x_{ij} - \tilde{x}_{ij}) = 1 \quad (2)$$

where  $\text{abs}(\cdot)$  denotes the absolute value. Therefore we can use the notation

$$q_{ij}(x) = q(x, \tilde{x}) \quad \text{where } x_{ij} \neq \tilde{x}_{ij} \quad (3)$$

SAOMs consider two principal concepts in constructing the intensity matrix: how often actors update their tie variables and what motivates their choice of which tie variable to update. This is expressed by the formulation

$$q_{ij}(x) = \lambda_i(x) p_{ij}(x). \quad (4)$$

The interpretation is that actor  $i$  gets *opportunities* to make an update in her/his outgoing tie at a rate of  $\lambda_i(x)$  (which might, but does not need to, depend on the current network); if such an opportunity occurs, the probability that  $i$  selects  $x_{ij}$  as the tie variable to change is given by  $p_{ij}(x)$ . The actors are not required to make a change when an opportunity occurs, which is reflected by the requirement  $\sum_j p_{ij}(x) \leq 1$ , without the need for this to be equal to 1. The probabilities  $p_{ij}(x)$  are dependent on the so-called *evaluation function*, as described below.

## 2.1 Rate function

The rate function describes the rate at which an actor  $i$  updates tie variables. Waiting times between opportunities for actor  $i$  to make an update to the digraph are exponentially distributed with rate parameter  $\lambda_i(x)$ , and it follows that waiting times between any two opportunities for updates across all actors are exponentially distributed with rate parameter

$$\lambda_+(x) = \sum_i \lambda_i(x). \quad (5)$$

It is possible to specify any number of functional forms for  $\lambda_i(x)$ , to include combinations of actor-level covariates and structural properties of the current state of the network  $x$ ; however, in many applications, rate functions are modeled as constant terms.

## 2.2 Evaluation function

Once an actor  $i$  is selected for an update, the actor must select a tie variable  $x_{ij}$  to change. Define  $x(i \rightsquigarrow j) \in \mathcal{X}$  as the digraph resulting from actor  $i$  modifying his tie variable with  $j$  during a given time period  $t$  so that  $x_{ij}(i \rightsquigarrow j) = 1 - x_{ij}$ , and formally define  $x(i \rightsquigarrow i) = x$ .

The SAOM assumes that the probabilities  $p_{ij}(x)$  depend on the *evaluation function* that gives an evaluation of the attraction toward each possible next state of the network, denoted here by  $f_{ij}(x)$ . This attraction is conveniently modeled as a linear combination of the relevant features of each potential change  $i \rightsquigarrow j$ :

$$f_{ij}(x) = \beta^T s_i(x(i \rightsquigarrow j)) \quad (6)$$

where  $s_i$  is a vector-valued function containing structural features of the digraph as seen from the point of view of actor  $i$ , and  $\beta$  is a statistical parameter. Snijders (2001), following the econometric literature on discrete choice (see e.g., Maddala 1983; McFadden 1973), models the choice of  $i \rightsquigarrow j$  as a myopic, stochastic optimization of a conditional logit. This amounts to choosing the greatest value  $f_{ij}(x) + \epsilon_{ij}$ , where  $\epsilon_{ij}$  is a Gumbel distributed error term. This leads to the conditional choice probabilities  $p_{ij}(x)$  that actor  $i$  chooses to change tie variable  $i \rightsquigarrow j$  given the current digraph  $x$ :

$$p_{ij}(x) = \frac{\exp f_{ij}(x)}{\sum_{k=1}^n \exp f_{ik}(x)}. \quad (7)$$

In accordance with the formal definition  $x(i \rightsquigarrow i) = x$ , the choice  $j = i$  is interpreted as keeping the current digraph as it is, without making a change. For a thorough menu of what kinds of statistics  $s_i$  are appropriate for actor oriented models (see Ripley and Snijders 2009; Snijders et al. 2010a). We will present here three fundamental structural statistics which are used throughout this paper: outdegree (density), reciprocity, and transitive triplet effects:

1. *Outdegree (density) effect*, defined by the number of outgoing ties that an actor has:

$$s_{i1}(x) = \sum_j x_{ij}. \quad (8)$$

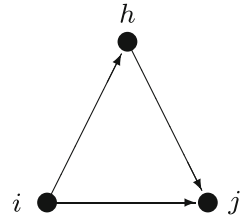
2. *Reciprocity effect*, defined by the number of outgoing ties that are matched (or *reciprocated*) by a corresponding incoming tie:

$$s_{i2}(x) = \sum_j x_{ij} x_{ji}. \quad (9)$$

3. *Transitive Triplet*, defined by the number of patterns matching the following Fig. 1:

$$s_{i3}(x) = \sum_h \sum_j x_{ih} x_{hj} x_{ij}. \quad (10)$$

These three features are considered fundamental to social network dynamics, but more features may also be formulated. One possibility is to include exogenous actor-level covariates  $c \in \mathbb{R}^n$ . In Sect. 5, a simple covariate statistic is used:

**Fig. 1** The transitive triplet

4. *Same covariate* defined by the number of actors  $j$  tied to actor  $i$  sharing the same value of  $c$ :

$$s_{i4}(x) = \sum_j x_{ij} \mathcal{E}(c_i, c_j), \quad (11)$$

where we define the covariate equivalence function

$$\mathcal{E}(c_i, c_j) = \begin{cases} 1 & \text{if } c_i = c_j \\ 0 & \text{if } c_i \neq c_j \end{cases}. \quad (12)$$

5. *Covariate similarity* defined by the number of actors  $j$  tied to actor  $i$  having a similar value of  $c$ :

$$s_{i5}(x) = \sum_j x_{ij} \mathcal{S}(c_i, c_j), \quad (13)$$

where we define the covariate similarity function

$$\mathcal{S}(c_i, c_j) = 1 - \frac{|c_i, c_j|}{\max_{k,l} (c_k, c_l)}. \quad (14)$$

See also [Snijders \(2001\)](#), [Snijders et al. \(2007\)](#), [Snijders \(2009\)](#), [Snijders et al. \(2010a\)](#) for thorough developments of possible statistics and extensions to models which consider the coevolution of network and endogenous behavioral characteristics.

### 2.3 Estimation

A key feature of (7) is the convenient form of its *log odds ratios* between any two potential next networks  $x(i \rightsquigarrow j)$  and  $x(i \rightsquigarrow k)$ :

$$\log \left[ \frac{p_{ij}(x)}{p_{ik}(x)} \right] = \log \left[ \frac{\frac{\exp f_{ij}(x)}{\sum_h \exp f_{ih}(x)}}{\frac{\exp f_{ik}(x)}{\sum_h \exp f_{ih}(x)}} \right] = \log \left[ \frac{\exp f_{ij}(x)}{\exp f_{ik}(x)} \right] = f_{ij}(x) - f_{ik}(x) \quad (15)$$

This is the characteristically simple property that makes estimation in classical discrete choices quite straightforward (see e.g., [Greene 2007](#)), and is also the basis for the SAOM. If the network updates  $\mathcal{L}$  are all observed so that we have full information (i.e.  $\mathcal{L} = \mathcal{M}$ ), and if the rate parameters are independent of the parameters of the objective function (as is usually the case), a convenient partial likelihood is available for the statistical parameters of the objective function. We use the notation  $x^{(a)} = x(M_a) = x(L_a)$  for the  $a$ th network observed in the dataset.

Define a vector of binary variables  $d$  such that

$$d_{ij}^{(a)} = \begin{cases} 1 & \text{if } x^{(a+1)} = x^{(a)}(i \rightsquigarrow j) \\ 0 & \text{if } x^{(a+1)} \neq x^{(a)}(i \rightsquigarrow j) \end{cases}, \quad (16)$$

which denotes whether actor  $i$  selected variable  $x_{ij}^{(a)}$  to update. Note that  $\sum_{j=1}^n d_{ij}^{(a)} = 1$  for all  $a$ . The partial log likelihood for the objective function parameters is

$$\begin{aligned} l(\beta) &= l(\beta|d) = \sum_{ija} d_{ij}^{(a)} \log p_{ij}(x^{(a)}) = \sum_{ija} d_{ij}^{(a)} \log \left( \frac{\exp f_{ij}(x^{(a)})}{\sum_k \exp f_{ik}(x^{(a)})} \right) \\ &= \sum_{ija} d_{ij}^{(a)} \left[ f_{ij}(x^{(a)}) - \log \left( \sum_k \exp f_{ik}(x^{(a)}) \right) \right]. \end{aligned} \quad (17)$$

Under regularity conditions, a solution  $\hat{\beta}_{\text{ML}}$  to

$$\nabla l(\beta) = 0$$

where

$$\nabla = \left( \frac{\partial}{\partial \beta_1}, \frac{\partial}{\partial \beta_2}, \dots \right)$$

is a maximum likelihood estimate for  $\beta$ .

Unfortunately, complete information on each network update is extremely rare, and a likelihood function as in (17) is not readily available if some transitions are unobserved. Even when tie creation is observed in continuous time, it is not often that observations concerning termination of ties are also available. It is far more often the case that the data observed will be in the form of panel data, where typically  $|\mathcal{M}| \ll |\mathcal{L}|$ . Accordingly, the method of moments as proposed by [Snijders \(2001\)](#) can be used as an alternative method for obtaining reasonable estimates. Consider the estimating function

$$g_n(\theta; z_n) = \sum_{m_a \in \mathcal{M}} \left( E_\theta \left\{ u \left( X^{(a+1)} \right) \mid X^{(a)} = x^{(a)} \right\} - u \left( x^{(a+1)} \right) \right), \quad (18)$$

which is simply the sum of deviations between the expected value of the statistics for the random (simulated) networks and the observed networks;  $z_n$  simply means all of



the available data, and  $\theta$  is a vector of parameters for the objective and rate functions described earlier.  $u(x)$  is a function that corresponds to appropriately chosen statistics calculated from the digraph for the parameters  $\theta$  (based on the statistics in 2.2). It is helpful to note that  $E_{\theta}\{g_n(\theta; z_n)\}$  does not depend on  $z_n$ , so it is not random. The method of moments involves finding the *moment estimate*  $\hat{\theta}$  solving the moment equation

$$g_n(\theta; z_n) = 0 \quad (19)$$

The specific details of fitting moment estimates are rather involved, and entail simulating networks  $X^{(a)}$  many times to achieve a reliable result for the expectation in (19).<sup>1</sup> This simulation is very straightforward. Take an initial network  $x(l_1)$  and proceed as follows for each update  $l_a \in \mathcal{L}$ :

1. Set  $l_a = l_{a-1} + \text{Expon}(\lambda_+)$
2. Select actor  $i$  with probability  $\frac{\lambda_i(l_{a-1})}{\lambda_+(l_{a-1})}$ .
3. Select actor  $j$  with probability  $p_{ij}(l_{a-1})$ .
4. If  $i \neq j$ , set  $x_{ij}(l_a) = 1 - x_{ij}(l_{a-1})$ .
5. Repeat until some specified conditions (e.g. number of updates  $|\mathcal{L}|$  or some holding time  $(l_{|\mathcal{L}|} - l_a)$  is exceeded) are satisfied.

See [Snijders \(2001\)](#), [Snijders et al. \(2010a\)](#) for guidelines on the selection of appropriate statistics for  $u(X(t))$  and information on how to estimate the root of  $g_n(z_n, \theta)$ , and [Schweinberger and Snijders \(2007\)](#) for the estimation of the derivative matrix, covariance matrix, and standard errors.

## 2.4 Estimator ( $\hat{\beta}$ ) properties

The properties of parameter estimates  $\hat{\beta}$  in SAOM have yet to receive a great deal of attention in the literature. Understanding the conditions under which desirable estimator properties (i.e. approximate unbiasedness, consistency, and efficiency) exist is an important feature for non-experimental studies with social networks, since erroneous conclusions could be drawn from the results. In the classical discrete choice literature, early attempts at uncovering the relationship between omitted variables and biased estimates in multinomial/conditional logit models include [Amemiya and Nold \(1975\)](#), [Lee \(1982\)](#), [Ruud \(1983\)](#), [Yatchew and Griliches \(1985\)](#), [Wooldridge \(2002\)](#), but fail to uncover an analytical form for bias in general. It is desirable to investigate similar properties for the estimators of SAOM, especially the form of any omitted variable biases. Unfortunately, closed forms for estimates of  $\beta$  and of  $\text{var}(\hat{\beta})$  are unavailable for all but the most simple SAOMs (see [Snijders 2005](#); [Van De Bunt et al. 1999](#)). Compounded with the lack of analytical results for even classical discrete choice models, straightforward analysis of SAOM estimates is very difficult.

<sup>1</sup> That this can take a considerable amount of time per estimation motivates the use of the score-type test in the next section.

Accordingly, we utilize simulation study to uncover some basic results on bias and efficiency under model misspecifications in Sect. 4.

### 3 Assessing time heterogeneity with the score type test

Consider a SAOM formulated as in (6) with some set of effects  $\mathcal{K}$  included. We initially assume that  $\beta$  does not vary over time, yielding a *restricted model*. Our data contains  $|\mathcal{M}| \ll |\mathcal{L}|$  observations, so we estimate the restricted model by the method of moments mentioned in Sect. 2.3. For reasons introduced in Sect. 1.1, we wish to test whether the *restricted model* is misspecified with respect to time heterogeneity. An *unrestricted model* which allows for time heterogeneity in all of the effects is considered as a modification of (6):

$$f_{ij}^{(a)}(x) = \sum_{K_k \in \mathcal{K}} \left( \beta_k + \delta_k^{(a)} \right) s_{ik}(x(i \rightsquigarrow j)) \quad (20)$$

where  $\delta_k^{(a)}$  is called the time dummy interacted effect parameter for effect  $k$  and period  $a$ . Define also the vectors  $\delta_k = \left( \delta_k^{(2)}, \dots, \delta_k^{(|\mathcal{W}|)} \right)$  and  $\delta = (\delta_1, \dots, \delta_{|\mathcal{K}|})$ . Equation (20) applies for updates occurring during the period  $W_a$ . By convention,  $\delta_k^{(1)} = 0$  for all  $k \in \mathcal{K}$  so that the first period is called the *base period*; therefore, the vector of time dummy interacted effect parameters  $\delta$  has length  $(|\mathcal{W}| - 1)|\mathcal{K}|$ .<sup>2</sup> One way to formulate the testing problem of assessing time heterogeneity is the following omnibus test:

$$\begin{aligned} H_0 : \delta &= 0 \\ H_1 : \delta &\neq 0. \end{aligned} \quad (21)$$

If we consider the unrestricted model parameters for (20) as  $\theta_1 = (\beta, \delta)$ , and the restricted model parameters for (6) as  $\theta_0 = (\beta, 0)$ , there are three broad routes in likelihood-based inference for testing  $H_0 : \theta_0 = 0$ : the likelihood ratio test, the Wald test, and the score test. Each of these routes requires different estimates to be calculated.

A likelihood ratio test statistic would take the form

$$\eta_L = 2 \left[ l(\hat{\theta}_1) - l(\hat{\theta}_0) \right]. \quad (22)$$

To construct  $\eta_L$ , we estimate the unrestricted model parameters  $\hat{\theta}_1$ , the restricted model parameters  $\hat{\theta}_0$ , and evaluate the likelihoods for the restricted and unrestricted models. There is a major difficulty with using this approach for a SAOM: when  $\mathcal{M} \neq \mathcal{L}$ , the likelihood for SAOM parameters is not available in closed form, although Snijders et al. (2010b) consider a maximum likelihood estimation approach for SAOM using data augmentation methods. As we will see, pursuing the likelihood ratio testing route will be the most computationally expensive among those asymptotic approaches

<sup>2</sup> Because  $\delta_k^{(1)}$  is fixed, it is implicitly omitted from  $\delta_k$  and  $\delta$  throughout the notation.

considered here. Note three important features of evaluating  $\eta_L$ : (1) both the restricted and unrestricted model parameters must be estimated, (2) we must evaluate the likelihood functions at these estimates, and (3) finding maximum likelihood estimates is more computationally expensive than finding method of moments estimates.

We might also consider a Wald test statistic of the form

$$\eta_W = \hat{\delta}^T \left[ i^{-1}(\hat{\theta}_1) \right]_{\delta\delta} \hat{\delta}, \quad (23)$$

where  $i(\hat{\theta}_1)$  is the *expected (Fisher) information matrix* and  $\left[ i^{-1}(\hat{\theta}_1) \right]_{\delta\delta}$  refers to the block corresponding to  $\delta$ .<sup>3</sup> Unlike  $\eta_L$ , the Wald statistic  $\eta_W$  does not require estimation of  $\theta_0$ ; however, since we must know  $\hat{\theta}_1$ , we still must estimate the unrestricted model.

The score test statistic is defined by

$$\eta_U = U(\hat{\theta}_0)^T i^{-1}(\hat{\theta}_0) U(\hat{\theta}_0) \quad (24)$$

where we only require the expected information matrix  $i(\hat{\theta}_0)$  and the score function  $U(\hat{\theta}_0)$ .<sup>4</sup> An advantage here is that the score test statistic  $\eta_U$  only requires a score function and the expected information matrix, and these ingredients do not require estimates for the unrestricted model. For a review of the classic asymptotic approaches to approximate inference presented above (see e.g. [Cox and Hinkley 1974](#); [Cox 2006](#); [Lehmann and Romano 2005](#)).

These three tests are classical tests for inference based on maximum likelihood estimation. However, for the SAOM, the maximum likelihood estimators recently developed by [Snijders et al. \(2010b\)](#) are computationally so demanding that in practice the more easily calculated method of moment estimators of [Snijders \(2001\)](#) and [Schweinberger and Snijders \(2007\)](#) are used. This type of inference is also called inference based on estimating functions or M-estimation. For inference using the method of moments, generalizations of the Wald test and score test are available, as proposed for the score test by [Rao and Poti \(1946\)](#), [Rao \(1948\)](#), [Neyman \(1959\)](#), [Basawa \(1985, 1991\)](#), and reviewed by [Rippon and Rayner \(2010\)](#).

Because  $\delta$  has length  $(|\mathcal{W}| - 1)|\mathcal{K}|$ , estimating it with method of moments is infeasible for even modest numbers of periods and effects. For example, given a modestly sized model containing 8 effects and a dataset with 5 periods, 40 parameters must be estimated (8 base effects  $\beta$  and 32 time dummy interaction effect parameters  $\delta$ ). Estimations of this size can begin to become unstable and computationally costly (particularly with larger numbers of actors, e.g.  $n > 150$ ). We therefore take the *forward-selecting* approach to model selection, where we start by estimating only the base effects and add time dummy interaction effects only when there is empirical evidence

<sup>3</sup> The Fisher information matrix is the second derivative of minus the log likelihood function  $l(\theta)$  with respect to  $\theta$ . Formally,  $i(\theta) = E_{\theta} \left( -\nabla \nabla^T l(\theta) \right)$  where  $\nabla = \left( \frac{\partial}{\partial \theta_1}, \frac{\partial}{\partial \theta_2}, \dots \right)$ . Often, the observed Fisher information matrix  $j(\hat{\theta}_1)$  is used, which is simply the sample-based version of the Fisher information matrix.

<sup>4</sup> The score function is the first derivative of the likelihood function, so that  $U(\theta) = \nabla l(\theta)$ .

or when there is some theoretical reason. The score-type test manifests naturally in such an approach. For the SAOM, it was proposed by [Snijders \(2001\)](#) to test parameters by a test which can be regarded as the analogue of the Wald test for estimates obtained by the method of moments. The generalization of the score test based on moment estimators for the SAOM was elaborated by [Schweinberger \(2007\)](#) following methods proposed by [Basawa \(1985, 1991\)](#). This requires analogues of  $U(\theta)$  and  $i(\theta)$  as used in (24), which are derived in the following way.

By  $\Delta(\theta)$  we denote the Jacobian,

$$\Delta(\theta) = \left( \frac{\partial E_{\theta}\{g_n(\theta; Z_n)\}}{\partial \theta} \right).$$

We partition the estimating function  $g_n(\hat{\theta}_0; z_n)$ , the covariance matrix of the estimating functions  $\Sigma(\hat{\theta}_0)$ , and the Jacobian  $\Delta(\hat{\theta}_0)$ ,

$$g_n(\hat{\theta}_0; z_n) = \begin{pmatrix} g_1(\hat{\theta}_0; z_n) \\ g_2(\hat{\theta}_0; z_n) \end{pmatrix} \quad (25)$$

$$\Sigma(\hat{\theta}_0) = \begin{pmatrix} \Sigma_{11}(\hat{\theta}_0) & \Sigma_{12}(\hat{\theta}_0) \\ \Sigma_{12}^T(\hat{\theta}_0) & \Sigma_{22}(\hat{\theta}_0) \end{pmatrix} \quad (26)$$

$$\Delta(\hat{\theta}_0) = \begin{pmatrix} \Delta_{11}(\hat{\theta}_0) & \Delta_{12}(\hat{\theta}_0) \\ \Delta_{21}(\hat{\theta}_0) & \Delta_{22}(\hat{\theta}_0) \end{pmatrix}, \quad (27)$$

so that subscript 1 corresponds to the restricted model and subscript 2 corresponds to the additional terms of the unrestricted model (i.e. the time dummy interacted effect parameters). We orthogonalize the moment function  $g_2(\hat{\theta}_0; z_n)$  of the nuisance parameters in  $g_1(\hat{\theta}_0; z_n)$  to yield:

$$b_n(\hat{\theta}_0; z_n) = g_2(\hat{\theta}_0; z_n) - \Gamma(\hat{\theta}_0)g_1(\hat{\theta}_0; z_n) \quad (28)$$

where

$$\Gamma(\hat{\theta}_0) = \Delta_{21}(\hat{\theta}_0)\Delta_{11}^{-1}(\hat{\theta}_0).$$

We now calculate  $\Xi(\hat{\theta}_0)$ , the asymptotic covariance matrix for  $b_n(\hat{\theta}_0; z_n)$ :

$$\Xi(\hat{\theta}_0) = \Sigma_{22}(\hat{\theta}_0) - \left[ \Sigma_{12}^T(\hat{\theta}_0)\Gamma^T(\hat{\theta}_0) + \Gamma(\hat{\theta}_0)\Sigma_{12}(\hat{\theta}_0) \right] + \Gamma(\hat{\theta}_0)\Sigma_{11}(\hat{\theta}_0)\Gamma^T(\hat{\theta}_0), \quad (29)$$

giving rise to the test statistic

$$b_n^T(\hat{\theta}_0; Z_n) \Xi^{-1}(\hat{\theta}_0) b_n(\hat{\theta}_0; Z_n) \quad (30)$$

with an asymptotic  $\chi^2$  distribution with degrees of freedom equal to the number of terms in  $b_n(Z_n, \hat{\theta}_0)$ . We will make use of this test statistic in assessing null hypotheses for time dummy interacted effect parameters.

So long as  $g_n(\hat{\theta}_0; z_n)$  is differentiable at  $\hat{\theta}_0$ , the same Taylor-series approximation used to calculate the covariance matrix  $\Xi(\hat{\theta}_0)$  (the delta method) can be manipulated to solve for a one-step estimator, which is a “quick and dirty” estimator that does not require the use of likelihood or moment equations (see e.g., [Schweinberger \(2007\)](#)):

$$\hat{\theta}^* = \hat{\theta}_0 - \Delta^{-1}(\hat{\theta}_0)g_n(\hat{\theta}_0; z_n). \quad (31)$$

We will test the effectiveness of the one-step estimator thoroughly in Sect. 4.

After SAOM parameters are estimated, we typically simulate the network evolution according to the estimated parameters thousands of times so that we can estimate standard errors. Using these simulations, we estimate the expected values for the estimating functions in (18) using the simulated and observed statistics. Conveniently, the simulations are conducted independently for each period in  $\mathcal{W}$ , so that we also have the estimating functions in (18) for the time dummy interacted effects.

### 3.1 Guidelines on selecting a decision procedure and constructing test statistics

There is an array of hypothesis tests which are available from these ingredients, which permit a number of procedural approaches to assessing time heterogeneity and ultimately refining the model. We present an iterative, forward-selecting procedure below. This decision procedure is informal, but in the absence of a universally superior rule, it is a reasonable approach to refining model selection. We also note here that outdegree effects will be given a privileged position in the selection of time dummy terms. The reason is that the outdegree is highly correlated with most other effects.

The decision procedure proceeds as follows:

1. Estimate the parameters  $\theta_0^\dagger$  of some arbitrary *restricted* model using the method of moments. We refer to this model as restricted with respect to the set of time dummy interacted effect parameters, because many or all of these are assumed to be zero in the model. Denote by  $\delta^\dagger$  the vector of time dummy interacted parameters which are zero in the restricted model and which the researcher would like to test.
2. Test the composite hypothesis

$$\begin{aligned} H_0^J : \delta^\dagger &= 0 \\ H_1^J : \delta^\dagger &\neq 0 \end{aligned} \quad (32)$$

Evaluate  $b_n(z_n, \hat{\theta}_0)$  and  $\Xi$  by including all  $\delta^\dagger$  terms in those ingredients of (25), (26), and (27) with subscript 2. Follow the procedure outlined above for orthogonalizing the testing function (28) of the nuisance parameters  $\theta^\dagger$  and estimating  $\Xi$  in (29). Construct the test statistic

$$\eta_U = b_n(Z_n, \hat{\theta}_0^\dagger)^T \Xi^{-1} b_n(Z_n, \hat{\theta}_0^\dagger) \quad (33)$$

which is asymptotically distributed  $\chi^2$  with degrees of freedom equal to the number of elements in  $\delta^\dagger$ . If we fail to reject  $H_0^J$ , stop.

3. If  $H_0^J$  is rejected, select one  $\delta_k^{(w)} \in \delta^\dagger$  to include in the model by considering two quantities. First, evaluate

$$z_k^{(w)} = \frac{[b_n(Z_n, \hat{\theta}_0^\dagger)]_{w,k}}{[\text{diag } \Sigma^{1/2}]_{w,k}}, \quad (34)$$

for each  $a, k$  combination. Each  $z_k^{(a)}$  is a test statistic for

$$\begin{aligned} H_0^{ak} : \delta^\dagger &= 0 \\ H_1^{ak} : \delta_k^{(a)} &\neq 0 \quad \text{and} \quad \delta_c^{(b)} = 0 \quad \text{for all } (c, b) \neq (k, a) \end{aligned} \quad (35)$$

and is distributed standard normal. In interpreting this array of test results, one should take into account that each test is directed at the overall null hypothesis  $H_0^J$ , and it is possible that the test based on  $z_k^{(a)}$  yields a significant result not because of  $\delta_k^{(a)}$  itself, but because some other  $\delta_h^{(b)}$  is non-zero.

4. Evaluate the one step estimators  $\hat{\theta}^*$  to see which  $\hat{\delta}_k^{(a)}$  has the greatest magnitude.
5. Select one  $\delta_k^{(a)}$  using the hypothesis test results and the one step estimators for inclusion. Because of what was mentioned in Step 3, the results for all  $k, a$  should be considered simultaneously, and substantive background knowledge and the researchers judgment will be important in making this choice. We highly recommend that outdegree heterogeneity is given a privileged position in this selection.
6. Set  $\delta^\dagger \rightarrow \delta^\dagger \setminus \delta_k^{(a)}$ , i.e. remove  $\delta_k^{(a)}$  from the vector  $\delta^\dagger$ .
7. Set  $\theta_0^\dagger \rightarrow \theta_0^\dagger \cup \delta_k^{(a)}$ , i.e. add  $\delta_k^{(a)}$  to the restricted model.
8. Return to Step 1.

This model selection process is ended when we fail to reject  $H_0^J$ . An advantage to this approach is that all of the evidence for time heterogeneity is assessed near the current estimate  $\hat{\theta}_0^\dagger$ , and we iteratively update it in an attempt to keep the local approximations valid. A disadvantage is that we must make one successively more computationally costly estimation for each time dummy term that is included.

#### 4 Simulation study

The simulation study is conducted to achieve three purposes: (1) to better understand the properties of  $\hat{\beta}$  and its standard errors when a model is improperly specified with respect to time heterogeneity, (2) to investigate the validity and relative efficiency of one step estimates as a tool for assessing time heterogeneity, and (3) to analyze the type I and II errors of the various available hypothesis tests from the score-type approach. The same simulation setup is used to address each of these research questions. A  $|\mathcal{W}| = 2$ -period dynamic network of  $n = 50$  actors is generated with parameters for outdegree  $\beta_{11} = -1.25$ , reciprocity  $\beta_{12} = 2$ , and transitivity  $\beta_{13} = 0.25$  at period

$w_1$ , and rate parameters  $\lambda = 1.5$  for both periods, values which are in line with many observed datasets (see e.g., [Snijders 2001](#); [Snijders et al. 2010a](#)).

We create a basis for comparison by simulating a *base model* SAOM with no time heterogeneity using the aforementioned setup. The parameters for the *restricted specification*, which correspond to the outdegree, reciprocity, and transitive triplets effects (i.e. no time dummy interactions), are estimated in a time homogeneous *restricted estimation*. The series of score-type tests and one step estimators described in Sect. 3 are then carried out using the results from the restricted estimation.

Next, a series of nine *perturbed models* are considered. At period  $w_2$ , some level of perturbation is introduced to one of the effects during generation (there are three levels). Since the restricted specification is improper for a perturbed model, we can compare the results from the base model by conducting a restricted estimation to see how the improperly specified parameter estimates behave. The score-type tests and one step estimators are obtained from the ingredients in the restricted estimation results. Finally, a properly specified set of parameters (i.e. the *unrestricted specification*) is estimated in an *unrestricted estimation* so that we can compare the relative efficiency of the one step estimate with the traditional method of moments estimate.<sup>5</sup> Using the results from these simulations, we treat each research purpose in turn:

#### 4.1 Validity of the method of moments estimators

We investigate the performance of these estimators here by the Monte Carlo simulations detailed above. Let  $[\hat{\beta}_k]_i$  be the  $i$ th from a set of  $N$  estimates from independently simulated data with identical effect parameter vectors. We use a one sample  $t$  test to investigate bias. Assuming that  $\hat{\mu}_k = \frac{1}{N} \sum_{i=1}^N [\hat{\beta}_k]_i$  is asymptotically distributed  $N(\beta_k + \text{bias}, \frac{\sigma_k}{\sqrt{N}})$ , we test  $H_0 : \text{bias} = 0$  by evaluating the  $t$  statistic

$$t_{N-1}^{(k)} = \frac{\sqrt{N}(\hat{\mu}_k - \beta_k)}{\hat{\sigma}_k}$$

for each parameter. Table 1 contains the results of this hypothesis test: outdegree, transitive triplet, and rate parameters indicate a statistically significant bias. However, the magnitude of this bias is very small and of no concern; the method of moments estimators for properly specified SAOMs are approximately unbiased and have appropriate standard errors.

Restricted estimations of the perturbed models indicate that restricted estimators for *unperturbed effects* (e.g. the estimated parameter for reciprocity effect in a restricted specification when outdegree is perturbed) are also approximately unbiased with appropriate standard errors. Further, the distributions of the method of moments

<sup>5</sup> Each model was generated and estimated using the R package RSiena v1.10. For estimation, a conditional approach was used, which simulates periods until the number of changes in the observed network matches the number of changes in the simulated network. For details on this approach, see [Snijders \(2001\)](#). Five phase 2 subphases and 1,500 phase 3 iterations were used for each restricted/unrestricted estimation. For more information on what these quantities mean, see the RSiena Manual ([Ripley and Snijders 2009](#)). Each perturbed model and the base model are generated 1,000 times for a total of 10,000 iterations.

**Table 1** This table provides a comparison of observed and population parameters for a properly specified SAOM from  $N = 1,000$  independent Monte Carlo simulations with an independent, one sample  $t$  test

Effect	Bias	$\hat{\mu}_k$	$\hat{\sigma}_k$	$t_{N-1}^{(k)}$
Outdegree (density)	0.012	-1.24	0.163	2.24
Reciprocity	0.005	2.00	0.250	0.57
Transitive triplets	-0.013	0.24	0.105	-3.86
Rate 1	0.010	1.51	0.213	1.48
Rate 2	0.019	1.52	0.210	2.83

estimates are almost identical to the restricted estimates for the unperturbed effects. Figures 2 and 3 illustrate these findings (refer to the black, dotted lines centered on the vertical reference line). The figures for outdegree effect are similar, but omitted for brevity.

This finding is surprising, as we might expect some difficulty in estimating parameters and standard errors in a misspecified model. This robustness might be due in part to the rather simple specification of the models used in the study, but it is an encouraging indication that method of moments estimation is resilient to time heterogeneity when a model is improperly specified. Along the same lines, we find that the expected value for the restricted estimates of perturbed effect parameters is  $w\beta_{1k} + (1 - w)\beta_{2k}$  where  $w \approx 0.5$ —which may be due to the constant rate parameters across periods. Figures 4 and 5 illustrate these findings; the black density plots in the left column correspond to the restricted estimates and lie centered between  $\beta_{1k}$  and  $\beta_{2k}$ . Again, the figures for outdegree effect are similar, but omitted for brevity.

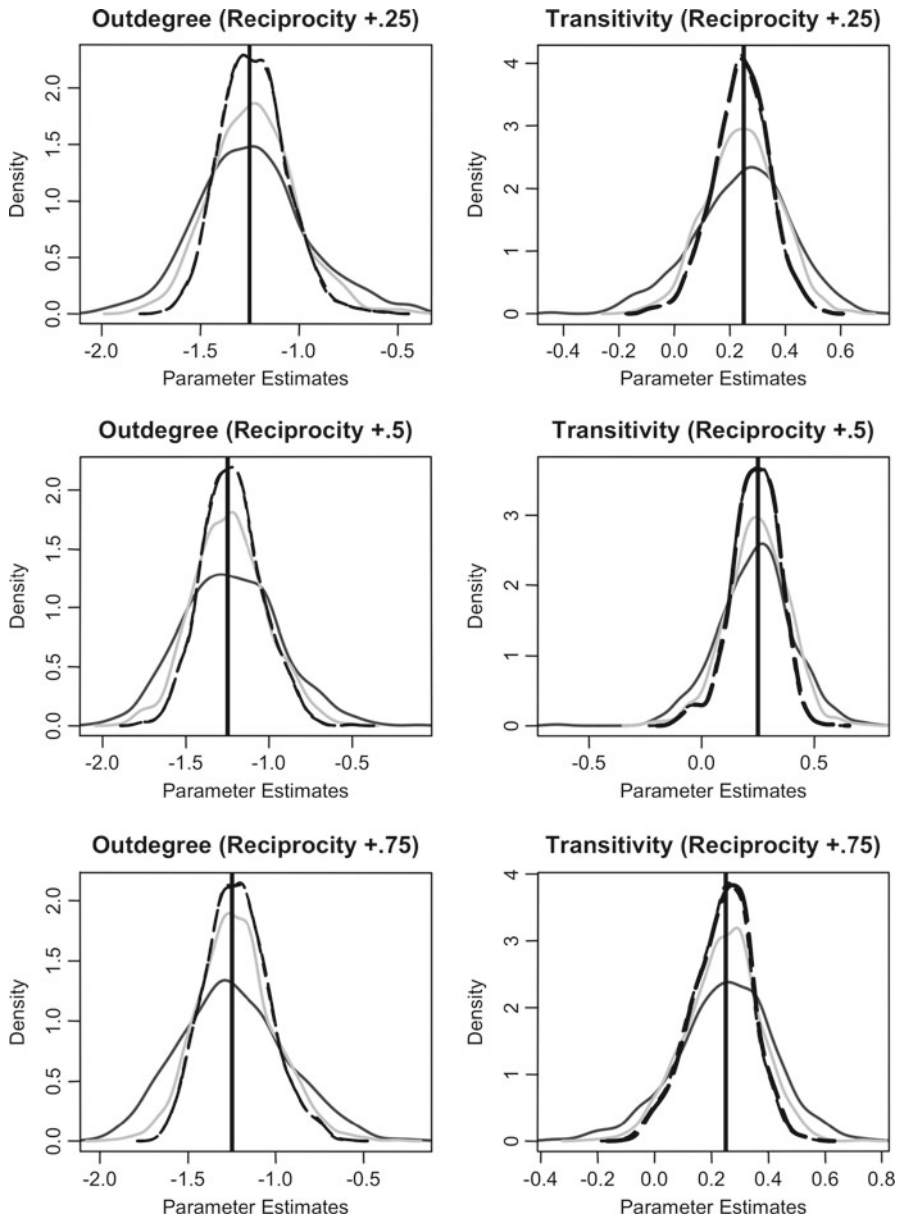
Also evident from these figures is the approximate unbiasedness of the method of moments estimates of effect parameters with simulated time heterogeneity as expected from the base model results (i.e. proper specifications correspond with approximately unbiased parameter estimates).

#### 4.2 Approximate validity and relative efficiency of one step estimates

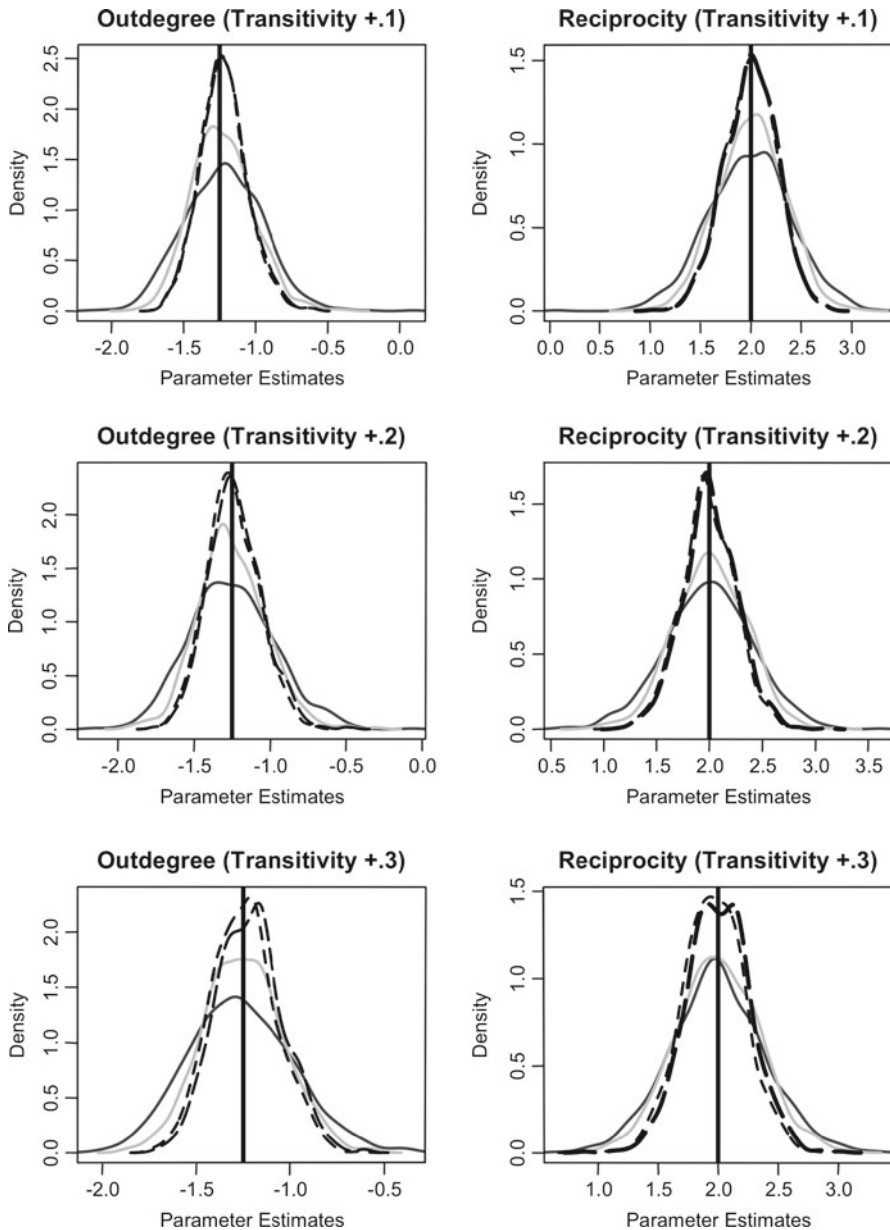
Restricted estimations of the perturbed models indicate that one step estimates using the ingredients from restricted estimators for unperturbed effects (e.g. the one step estimates for reciprocity effect and its time dummy interaction in a restricted specification when outdegree is perturbed) are approximately unbiased. Referring to the dotted density plots in Figs. 2 and 3 the reader may visually confirm this finding. The one step estimators for time dummy interacted effect parameters are, however, less effective at detecting time homogeneity than method of moments estimates as evidenced by their relatively flat density plots. In other words, when time heterogeneity does not exist, one step estimates will have wider dispersion than method of moments estimates, illustrating their “quick and dirty” nature.

The one step estimates for the perturbed effects (i.e. in the case with time heterogeneity in the simulated effect) are likewise approximately unbiased. Surprisingly, the one step estimates for perturbed reciprocity and transitive triplet effect parameters are

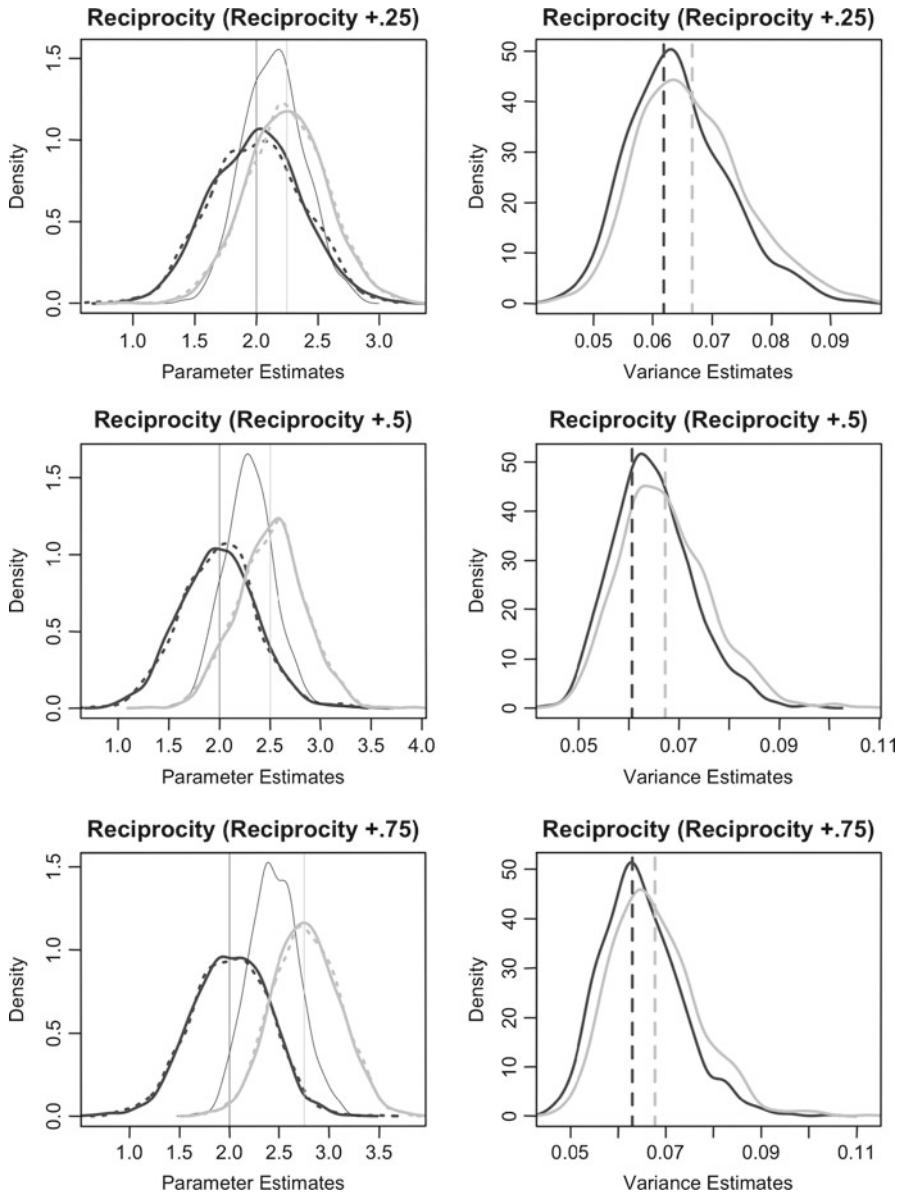




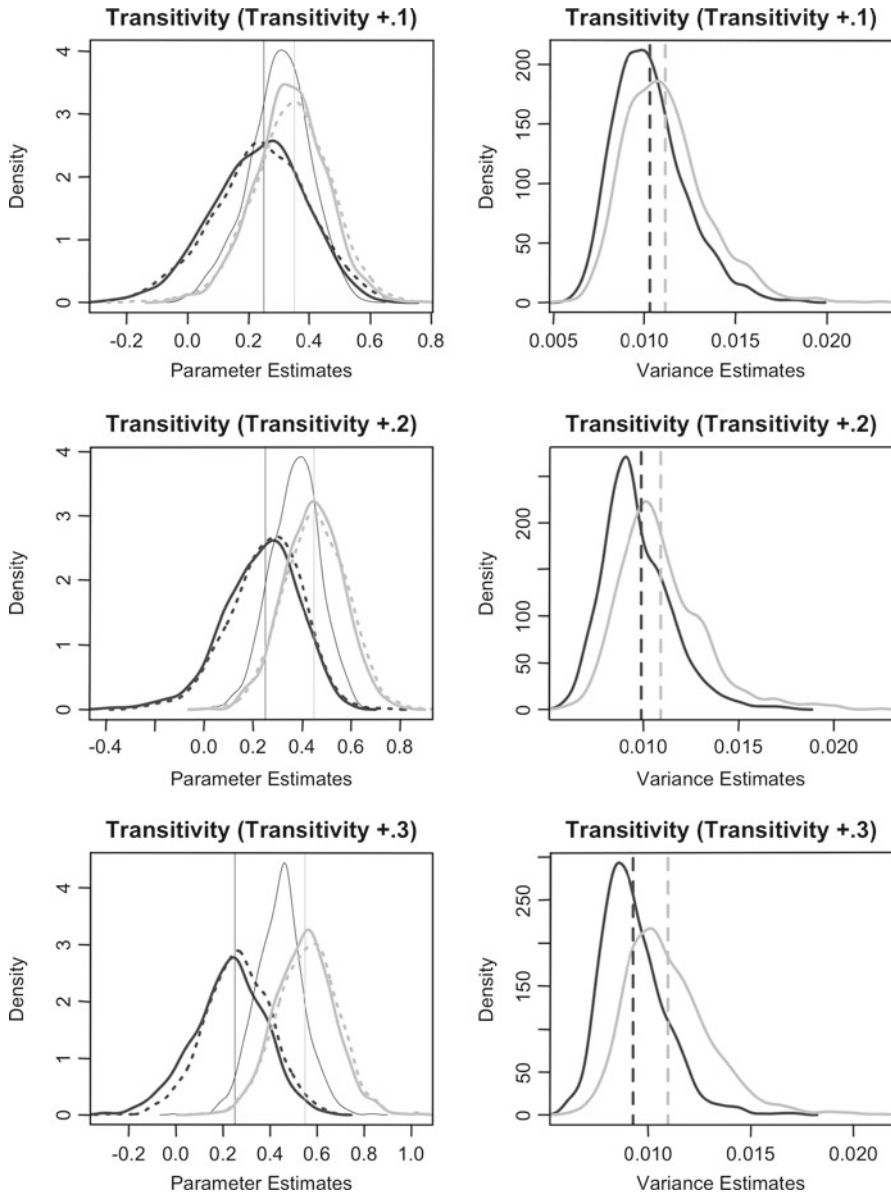
**Fig. 2** Omitting reciprocity time heterogeneity: density plots illustrating approximate unbiasedness of estimates for time homogeneous parameters in the presence of simulated reciprocity time heterogeneity. The two solid curves correspond to the distributions of the one step estimates  $\hat{\beta}^*$  and  $\hat{\beta}^* + \hat{\delta}^*$  (i.e. the base period parameter—in gray—and the base period plus time heterogeneity term—in black, slightly broader than the base period). The two (indistinguishable) dotted curves correspond to the method of moments estimates for  $\beta$  under the unrestricted and restricted models (i.e. estimated with and without the time heterogeneity parameter included for reciprocity). The vertical reference lines correspond with the population generating quantities for each effect parameter



**Fig. 3** Omitting transitive triplets time heterogeneity: Density plots illustrating approximate unbiasedness of estimates for time homogeneous parameters in the presence of simulated transitive triplets time heterogeneity. The *two solid curves* correspond to the distributions of the one step estimates  $\hat{\beta}^*$  and  $\hat{\beta}^* + \hat{\delta}^*$  (i.e. the base period parameter—in gray—and the base period plus time heterogeneity term—in black, slightly broader than the base period). The two (indistinguishable) *dotted curves* correspond to the method of moments estimates for  $\beta$  under the unrestricted and restricted models (i.e. estimated with and without the time heterogeneity parameter included for transitive triplets). The *vertical reference lines* correspond with the population generating quantities for each effect parameter



**Fig. 4** Performance of the reciprocity effect parameter one step estimators: this figure plots parameter estimates in the *left column*. For these figures, the *fine, black density plot* (the tallest among the plots) represents the restricted estimate. To either side of this restricted parameter density plot are the period-wise estimates for the unrestricted model. The *dotted density plot* represents the one step estimates, and the *bold density plots* represent the method of moments estimators. The true period-wise values are given by *vertical reference lines*. On the *right*, estimates of the parameter estimate variances are shown. The *black plots* represent the variance estimators from the restricted estimation and the *gray plot* represents the variance of the unrestricted estimates. *Vertical lines* represent the sample variances



**Fig. 5** Performance of the transitive triplet effect parameter one step estimators: this figure plots parameter estimates in the *left column*. For these figures, the fine, *black density plot* (the tallest among the plots) represents the restricted estimate. To either side of this restricted parameter density plot are the period-wise estimates for the unrestricted model. The *dotted density plot* represents the one step estimates, and the *bold density plots* represent the method of moments estimators. The true period-wise values are given by *vertical reference lines*. On the *right*, estimates of the parameter estimate variances are shown. The *black plots* represent the variance estimators from the restricted estimation and the *gray plot* represents the variance of the unrestricted estimates. *Vertical lines* represent the sample variances

**Table 2** Correlations among the one step estimates and the method of moments estimates: this table presents the Pearson correlation coefficient between the one step estimates and the method of moments estimates for the unrestricted model parameters for each treatment group indicated by row title

Effect <sub>k</sub>	$\beta_k$	$\delta_k^{(2)}$
Outdegree (+0.3)	0.79	0.91
Outdegree (+0.6)	0.87	0.90
Outdegree (+0.9)	0.85	0.82
Reciprocity (+0.25)	0.92	0.96
Reciprocity (+0.5)	0.93	0.96
Reciprocity (+0.75)	0.94	0.97
Transitive triplets (+0.1)	0.85	0.93
Transitive triplets (+0.2)	0.85	0.93
Transitive triplets (+0.3)	0.67	0.94

The level of simulated time heterogeneity is provided in parenthesis

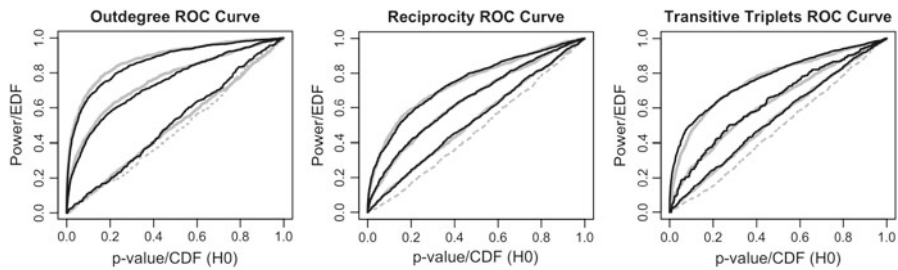
Correlations close to one indicate that the one step estimates and the method of moments estimates tend to agree on the direction of the heterogeneity

at least as efficient as their method of moments counterparts. The dotted density plots of Figs. 4 and 5 illustrate these results. The outdegree results are mixed; nonetheless, the estimators perform well on balance. Figures of outdegree results are omitted for brevity.

In short, the one step estimators are approximately valid in that they are unbiased (i.e. their expected values correspond to the true model under both time heterogeneous and homogeneous parameterizations). This simulation study indicates that the one step estimators are less efficient under homogeneous conditions, since method of moments estimates have less variance. When there is time heterogeneity in the effects, however, one step estimates perform as well as their method of moments counterparts. To further synthesize the relationship between the one step and method of moments estimates, we report Pearson correlation coefficients in Table 2. There is a clear, positive correlation between the two estimates, which further supports the approximate validity of the one step estimates as a tool to detect time heterogeneity.

#### 4.3 Effectiveness of the score-type test

There are two important properties of the score-type test which we consider here: type I error and power. The approach for evaluating the type I error is to first validate that the distributions used to evaluate the hypotheses of Sect. 3 are valid. We conduct the score-type tests for both joint significance and individual significance using  $\chi^2$  distributions with the appropriate d.f., yielding  $p$  values. These  $p$  values are assembled into receiver operating characteristic (ROC) curves, which represent the probability of rejection of the test as a function of its nominal significance level (i.e.  $\alpha$  or type I error level). These should resemble the cumulative density function for a uniform distribution  $\text{Unif}(0, 1)$  when the null hypothesis is true (i.e. there is time



**Fig. 6** Charts for statistical tests: receiver operating characteristic (ROC) curves. Dotted gray curves correspond to individual tests for the effect indicated by the plot's title when  $H_0^J$  is true. For three levels of time heterogeneity, the solid gray and solid black curves plot the results of the individual and joint tests, respectively. Uniformly more powerful curves correspond with greater levels of time heterogeneity

homogeneity in the effect parameters).<sup>6</sup> Figure 6 illustrates that the score-type test meets this criterion.

When  $H_0$  is false, the ROC curves indicate the power on the vertical axis for varying type I error rates on the horizontal axis. It is desired that power is as high as possible for all type I error rates.

While it is straightforward to validate the type I error indicated by the score-type test, there is no objective basis over which to evaluate the power of these tests. Nonetheless, there are some key features available from the results. As expected, greater heterogeneity in the effect parameter leads to greater power in detecting the heterogeneity. Additionally, these results enrich the previous findings by illustrating that perturbations of  $\{+0.6, +0.9\}$  on the outdegree effect parameter correspond to the most powerful of the score-type tests (while  $\{+0.3\}$  is the least powerful).

The score tests of the different parameters will be correlated to an a priori unknown extent. When computing test statistics for individual parameters  $\delta_k^{(a)}$ , a very rough approximation is possible based on the covariance matrix  $\mathcal{E}$  in (29): the scores of all other  $\delta_k^{(a)} \in \delta^\dagger$  are included in  $g_1(z_n, \hat{\theta}_0)$  of (25) in addition to the parameters  $\theta^\dagger$  which have been estimated by the method of moments. This approximation is valid only in the close vicinity of the overall null hypothesis  $H_0^J$ , and it may not hold for parameter values away from this null hypothesis. Table 3 shows that for the case of various levels of simulated outdegree time heterogeneity, the individual test statistics are strongly correlated.

The practical significance of this result is the following: when  $H_0^J$  is false, all of the individual tests have a tendency to show evidence for time heterogeneity—even those effect parameters which are time homogeneous. This complicates procedures such as the one given in Sect. 3.1, since the selection of how to iteratively update  $\theta^\dagger$  is based in part on these individual test statistics.<sup>7</sup> Using the approximate orthogonalization above, dependence may be substantially reduced in some limited situations.

<sup>6</sup> For more information on ROC curves (see Fawcett 2006; Zweig and Campbell 1993).

<sup>7</sup> We advised giving outdegree a privileged position in part for this reason.

**Table 3** Pearson correlation coefficients between  $t$  statistics for individual time heterogeneity parameter significance tests in the presence of simulated outdegree time heterogeneity

$t_a$	$t_b$	$\text{cor}\left(t_a^\perp, t_b^\perp\right)$	$\text{cor}(t_a, t_b)$
Outdegree (+0.0)	Reciprocity	0.06	0.46
Outdegree (+0.3)	Reciprocity	0.05	0.43
Outdegree (+0.6)	Reciprocity	0.05	0.44
Outdegree (+0.9)	Reciprocity	0.06	0.49
Outdegree (+0.0)	Transitive triplets	-0.06	0.64
Outdegree (+0.3)	Transitive triplets	-0.10	0.62
Outdegree (+0.6)	Transitive triplets	-0.09	0.69
Outdegree (+0.9)	Transitive triplets	-0.13	0.67

The magnitude of the heterogeneity is given in parenthesis.

Test statistics treated with approximate orthogonalization based on  $\mathcal{E}$  in (29) are denoted with a superscript  $\perp$

Table 3 shows that both reciprocity and transitive triplets time heterogeneity parameters are only weakly correlated after performing the approximate orthogonalization. The rough approximation may be valid in the present circumstances due to the sparsity of effect parameters and to the small magnitude of overall time heterogeneity in the parameters; however, the validity of this procedure needs further investigation before it can be applied generally.

## 5 Application: Bristol and Cardiff's ASSIST data

Campbell et al. (2008) conducted a study funded by the UK Medical Research Council which involved peer-nominated students aged 12–14. These individuals underwent intensive training on how to discourage their peers from smoking. The complete dataset contains data on over 10,000 students.<sup>8</sup> Smoking behavior, friendship nomination, sex, age, parental smoking habits, family affluence, and nominal information on classroom membership are all features of the dataset. There are a number of interesting research questions supported by this study. In the following analysis, we wish to find whether an ego's smoking behavior affects nomination of alters on the basis of the alter's smoking behavior (i.e. selection).<sup>9</sup>

Because the dataset was collected over a three year period where adolescents are likely to modify social behavior, any assumptions of time homogeneity are suspect. Additionally, Steglich et al. (2010) notes that the ASSIST dataset contains substantial

<sup>8</sup> For the purposes of this study, we select the 236 students in one of the 59 schools observed every year for 3 years. This particular school is in a relatively affluent area, geographically located in south Wales.

<sup>9</sup> In the full study, smoking behavior is treated as an endogenous variable, so that we can also investigate whether an ego's smoking behavior is affected by alter smoking behavior (i.e. influence). The addition of smoking behavior to the model is beyond the scope of this simple example, which seeks to simply identify heterogeneity in the network effects.



**Table 4** Estimates for a restricted SAOM model of the Cardiff ASSIST data (Campbell et al. 2008)

Effect name	$\hat{\theta}$	$se(\hat{\theta})$	t. conv.	$\hat{\theta}^*$	$p$ value
Rate (1)	16.762	1.290	–	–	–
Rate (2)	10.834	0.681	–	–	–
Outdegree	–2.892	0.029	–0.034	–2.910	–
(2)	–	–	–	0.042	0.467
Reciprocity	1.958	0.065	0	1.935	–
(2)	–	–	–	0.056	0.673
Transitive triplets	0.415	0.015	–0.025	0.457	–
(2)	–	–	–	–0.078	0.012
Age similarity	0.318	0.149	0.034	0.037	–
(2)	–	–	–	0.112	0.902
Smoking sim.	0.634	0.117	0.002	0.457	–
(2)	–	–	–	0.101	0.384
Same sex	0.876	0.133	–0.001	0.923	–
(2)	–	–	–	–0.044	0.492
Same form	0.566	0.042	–0.040	0.767	–
(2)	–	–	–	–0.450	0

$\hat{\theta}$  corresponds to the estimate obtained by conditional method of moments, and  $\hat{\theta}^*$  corresponds to the one step estimates

Rows marked (2) indicate the corresponding values for a dummy term interacted with the preceding effect  
Estimates for the time dummy terms are given below their respective base effects

The  $p$  value is for the score test of the hypothesis that the dummy term is equal to zero

time heterogeneity in some important parameters of interest. Ultimately, they chose to mitigate the heterogeneity by estimating period 1  $\rightarrow$  2 separately from 2  $\rightarrow$  3, effectively dummifying all of the effect parameters.

Due to potential convergence issues in estimation with models containing large numbers of parameters (and for illustrative purposes), we first specify a restricted model with period-wise independence for the rate parameters only. We specify as network effects the following: basic first and second order structural dependencies (outdegree, reciprocity, and transitive triplets), smoking similarity, same form (a control for the classroom to which the student is assigned), age similarity, and same sex effects (refer to Sect. 2.2 for an explanation of these effects).

We then estimated the restricted model using RSiena (Ripley and Snijders 2009). All of the effects achieved good convergence, as indicated by near-zero convergence  $t$  statistics. Results from this estimation are contained in Table 4. The rest of the terms support our intuition: age, smoking, sex, and classroom similarity increase the probability that an actor chooses an alter. Basic first and second order dependencies in reciprocity and transitive triplets also make links more attractive.

Using the tools developed in Sect. 3, we conduct score type test for time heterogeneity. The analysis of power from Fig. 6 illustrates that  $\alpha$  levels that are too low have very little power to detect heterogeneity. Over the ranges studied in Sect. 4,  $\alpha \approx 0.05$  gives reasonable power (near 0.5, e.g., for outdegree) for modest levels of heterogeneity.



**Table 5** Estimates obtained by conditional method of moments for a refined, unrestricted SIENA model to the Cardiff ASSIST data (Campbell et al. 2008)

Effect name	$\hat{\theta}$	se( $\hat{\theta}$ )
Rate (1)	15.278	1.010
Rate (2)	11.103	0.766
Outdegree	-2.895	0.029
Reciprocity	1.947	0.062
Transitive triplets	0.421	0.015
(2)	-0.042	0.024
Age similarity	0.303	0.152
Smoking sim.	0.547	0.115
Same sex	0.778	0.254
Same form	0.507	0.045
(2)	-0.503	0.077

Following the decision procedure from the previous section, we might elect to include dummy terms for those effects with score type tests yielding  $p$  values of 0.05 or less: transitive triplets and same form. We note that these results coincide with the findings of Steglich et al. (2010); even without introducing smoking behavior as a dependent variable or use of the full dataset, we find heterogeneity in transitive ties and in same classroom through the use of one-step estimates. We then further confirm the heterogeneity in same classroom and in transitive ties through an updated model containing time dummy terms for transitive ties and same classroom.

Using the iterative approach of the decision procedure given in the last section, we estimate an updated model with a time dummy interacted same form effect. Evidence for transitive triplets time heterogeneity is still present, so we update the model again and re-estimate. At this point, the joint test statistic of  $H_0^J$  does not indicate further time heterogeneity. Estimation of this updated model supports the indications of heterogeneity detected by the one step estimates: transitive ties has a very mild heterogeneity, but same form is highly heterogeneous and becomes almost insignificant in the second period. These results are presented in Table 5. It is interesting that same form is a prominent feature of actor behavior during period one, and the dummy completely negates the effect during period two. It appears that membership within the same classroom only encourages creation of ties during the lower age range, and that this effect diminishes when pupils get older. The transitive triplets time dummy interaction is statistically insignificant in the method of moments estimates, potentially a result of controlling for the strong effect of same form. All of the estimates from the updated model are stable in comparison to the restricted model, indicating good convergence of the estimates.

In summary, we were able to reproduce heterogeneity in both the one step estimates and the estimates from the unrestricted models. Even though Steglich et al. (2010) may have found greater heterogeneity in the transitive ties effect, we have only used a small subset of the data and we have not accounted for smoking behavior as an endogenous variable. This application has generally supported the use of the score type test to assess time heterogeneity.

### 5.1 A brief sketch of RSiena

RSiena has an extensive manual ([Ripley and Snijders 2009](#)). This section gives a concise walkthrough of how to operate RSiena v1.10 to produce the results of the foregoing example.

The `assist63` dataset is compiled from [Campbell et al. \(2008\)](#) and is loaded in the usual way.<sup>10</sup> To load the RSiena package, type the following:

```
library(RSiena)
```

To access information on the dataset, issue the command `?assist631`. Setting up the data can be done a number of ways. Perhaps the simplest method is to create the objects in batch mode:

```
nets <-sienaNet(array(c(assist631, assist632, assist633), dim = c(236, 236, 3)))
fas <-varCovar(assist63fas)
form <-varCovar(assist63form)
ps <-varCovar(assist63ps)
sex <-coCovar(assist63sa[, 1])
age <-coCovar(assist63sa[, 2])
dat <-sienaDataCreate(nets, fas, form, ps, sex, age)
eff <-getEffects(dat)
```

The `sienaNet` function sets up the dependent variables, while the `varCovar` and `coCovar` functions set up the time varying and constant covariates. `sienaDataCreate` joins all of the dependent variables and covariates into a single data object, and `getEffects` generates an effects object containing all of the potential interactions and effects we might want to include. The effects object contains the model specification by a column which, for each available effect, indicates whether this effect is included in the model. For those conversant in R, `eff` is a data frame which may be accessed and modified in the normal way. Otherwise, the effects to be included may be specified by typing

```
fix(eff)
```

and using the graphical user interface and thereby manipulating the effects object. Effects to be included are turned on by setting `include = TRUE`. For more information on the many fields available and the details of this operation, see [Ripley and Snijders \(2009\)](#). To run the estimation, the following may be used:

```
estimate <-sienaModelCreate(fn = simstats0c)
results.1 <-siena07(estimate, data = dat, effects = eff, batch = FALSE)
```

The `results.1` object now contains a wealth of information about the estimation. It is a list of various data objects which may be accessed individually, or issuing the

<sup>10</sup> A simulated version of this dataset is distributed with the RSiena package as of v1.10.

command `summary(results.1)` will display much of the important convergence and parameter information. Further, an output file, containing a large body of important diagnostic information, is generated in the current working directory (which is given by `getwd()`).

Two features have been added to assess and fix time heterogeneity. To run the score type test and display the results, the following code may be used:

```
timetest.1 <-sienaTimeTest(results.1)
summary(timetest.1)
plot(timetest.1, effects = c(1, 2, 4))
```

Plots of the one step estimates and approximate standard errors are presented alongside the diagnostic test results. Type `?sienaTimeTest` for more information on how to use these tools.

If it is determined that time dummied interaction terms should be included, the column `timeDummy` of the effects object `eff` may be employed to automatically generate the time dummy interaction term. For example, for the reciprocity effect this can be done by the command

```
eff$timeDummy[eff$shortName == 'recip' & eff$type == 'eval'] = "2"
```

or the convenience function

```
eff <-includeTimeDummy(eff, recip, timeDummy = "all")
```

and the updated model may be estimated in the usual way, with the modified effects object:

```
results.2 <-siena07(estimate, data = dat, effects = eff, batch = FALSE)
```

Using the iterative approach suggested in Sect. 3.1, we may again test the remaining unestimated time dummy terms, e.g.,

```
timetest.2 <-sienaTimeTest(results.2)
summary(timetest.2)
plot(timetest.2, effects = c(1, 4))
```

until we are satisfied with the model fit.

It is possible to produce individual test statistics with the approximate orthogonalization discussed in Sect. 4.3 by specifying

```
timetest.1.orthog <-sienaTimeTest(results.1, condition = TRUE)
```

so that if `results.1$theta` is close to the true parameter  $\theta$ , the individual test statistics may be less dependent.

## 6 Conclusion

Social networks evolve in potentially complicated ways, and relational dependencies cause difficulties in statistical modeling. Longitudinal social network data be used to

model the rules that drive the changes in network structure. Stochastic actor oriented models (SAOM) are a flexible family of statistical models that are designed to draw inference about these network dynamics through the use of longitudinal data.

Because SAOMs model network dynamics over time, there is a potential for effects to be time heterogeneous. The ramifications of omitting variables in the conditional logit models are not currently well understood, but we submit three motivations for testing time heterogeneity in datasets with more than two time periods: (1) if a parameter of interest has time heterogeneity, it is intrinsically interesting, (2) if a nuisance parameter has time heterogeneity and is not properly specified, it is possible that the estimators for the parameters of interest could take on undesirable properties like inconsistency and poor efficiency, and (3) SAOMs may be asymptotically degenerate, and introducing heterogeneity in the model can help to alleviate this problem.

We have presented the stochastic actor based model and two estimation procedures (maximum likelihood for complete continuous-time data and method of moments for panel data) which may be employed under different circumstances. After laying out the testing problem for time heterogeneity through the use of time dummy terms, we illustrated why here the score test approach is preferable to the likelihood ratio and Wald test approaches. We then reviewed the score type test of [Schweinberger \(2007\)](#) and applied it to tests for time heterogeneity and resulting one step estimators. This is based on deviations between observed and expected statistics, and on approximating the former by simulated values. Because the algorithm used for calculating the method of moments estimate generates the deviations between simulated and observed values period by period, calculating these deviations comes for free, making the score type approach computationally cheap compared with estimating an unrestricted model.

A simulation study indicates approximate unbiasedness of the one step estimators, the validity of the statistical tests, and acceptable levels of power for the perturbations studied. With the simple model used for the study, improper specification of a time heterogeneous effect did not cause important amounts of bias or inefficiency in the other effect parameters. This unexpected result deserves exploration in future work. The method of moments estimates under a misspecified model behaved nicely; estimates for the time homogeneous specification had expected values roughly equal to an affine combination over the time heterogeneous effect parameter values for the two time periods.

An example application to Bristol and Cardiff's ASSIST Data was supplied for illustrative purposes. The test, as implemented in the R package RSiena, was demonstrated briefly.

Applying the work of [Schweinberger \(2007\)](#) to time heterogeneity, we have shown that time heterogeneity of a SAOM can be tested with properly formulated score-type tests. A natural extension to this paper's goodness of fit for time heterogeneity is to develop a score-type test of actor homogeneity assumptions. The same machinery in Sects. 2 and 3 can be applied to create new tools to allow researchers to assess their model specification.

Assessing time heterogeneity is an important aspect of studying longitudinal data. Until now, this has taken the form of a time consuming process involving estimation of the unrestricted model and performing a Wald type test on the estimated time dummy interacted effect parameters. With the score type test now applied to time heterogeneity,

researchers can rapidly assess and respecify proper models. Quick tests for SAOM misspecification can be helpful for researchers as easily applied methods guarding against an important type of model misspecification. This test for time heterogeneity is one of many potential future applications that can help us better untangle the complex dynamics of social networks.

## References

- Amemiya T, Nold F (1975) A modified logit model. *Rev Econ Stat* 57:255–257
- Audrey S, Cordall K, Moore L, Cohen D, Campbell R (2004) The development and implementation of an intensive, peer-led training programme aimed at changing the smoking behaviour of secondary school pupils using their established social networks. *Health Educ J* 63(3):266–284
- Basawa I (1985) Neyman-Le Cam tests based on estimation functions. In: *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer*. Wadsworth, pp 811–825
- Basawa I (1991) Generalized score tests for composite hypotheses. In: *Estimating functions*, chap 8. Oxford Science Publications, Oxford, pp 131–131
- Borgatti S, Foster P (2003) The network paradigm in organizational research: a review and typology. *J Manage* 29:991–1013
- Brass D, Galaskiewicz J, Greve H, Tsai W (2004) Taking stock of networks and organizations: a multilevel perspective. *Acad Manage J* 47:795–817
- Burk W, Steglich C, Snijders T (2007) Beyond dyadic interdependence: actor-oriented models for co-evolving social networks and individual behaviors. *Int J Behav Dev* 31:397–404
- Campbell R, Starkey F, Holliday J, Audrey S, Bloor M, Parry-Langdon N, Hughes R, Moore L (2008) An informal school-based peer-led intervention for smoking prevention in adolescence (ASSIST): A cluster randomised trial. *Lancet* 371(9624):1595–1602
- Cox D (2006) *Principles of statistical inference*. Cambridge University Press, Cambridge
- Cox D, Hinkley D (1974) *Theoretical statistics*. Chapman and Hall, London
- Doreian P, Stokman F (eds) (1996) *Evolution of social networks*. Gordon and Breach Publishers, Newark
- Doreian P, Stokman F (eds) (2001) *Evolution of social networks*. Gordon and Breach Publishers, Newark
- Doreian P, Stokman F (eds) (2003) *Evolution of social networks*. Gordon and Breach Publishers, Newark
- Fawcett T (2006) An introduction to roc analysis. *Pattern Recognit Lett* 27:861–874
- Greene W (2007) *Econometric analysis*, 6th edn. Prentice Hall, New Jersey
- Handcock M (2003) *Assessing degeneracy in statistical models of social networks*. Center for Statistics and the Social Sciences, University of Washington. Available from: <http://www.csss.washington.edu/Papers>
- Lee L (1982) Specification error in multinomial logit models. *J Econom* 20:197–209
- Lehmann EL, Romano JP (2005) *Testing statistical hypotheses*, 3rd edn. Springer, New York
- Maddala G (1983) *Limited-dependent and qualitative variables in econometrics*, 3rd edn. Cambridge University Press, Cambridge
- McFadden D (1973) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed) *Frontiers in econometrics*. Academic Press, New York, pp 105–142
- Neyman J (1959) Optimal asymptotic tests of composite statistical hypotheses. In: Grenander U (ed) *Probability and statistics. The Harald Cramér Volume*. Wiley, New York, pp 213–234
- Norris J (1997) *Markov chains*. Cambridge University Press, Cambridge
- Pearson M, Michell L (2000) Smoke rings: social network analysis of friendship groups, smoking, and drug-taking. *Drugs Educ Prev Policy* 7:21–37
- Rao C (1948) Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proc Cambridge Philos Soc* 44:50–57
- Rao C, Poti S (1946) On locally most powerful tests when alternatives are one sided. *Indian J Stat* 7(4):439
- Ripley R, Snijders T (2009) *Manual for RSiena version 4.0*
- Rippon P, Rayner JCW (2010) Generalised score and Wald tests. *Adv Decis Sci* 2010:8
- Ruud P (1983) Sufficient conditions for the consistency of maximum likelihood estimation despite misspecification of distribution in multinomial discrete models. *Econometrica* 51:225–228
- Schweinberger M (2007) Chapter 4: statistical modeling of network panel data: goodness of fit. PhD thesis, University of Groningen, Groningen

- Schweinberger M, Snijders T (2007) Markov models for digraph panel data: Monte carlo-based derivative estimation. *Comput Stat Data Anal* 51(9):4465–4483
- Snijders T (2001) The statistical evaluation of social network dynamics. In: Sobel M, Becker M (eds) *Sociological methodology*. Blackwell, London pp 361–395
- Snijders T (2002) Markov chain Monte Carlo estimation of exponential random graph models. *J Soc Struct* 3:1–40
- Snijders T (2005) Models for longitudinal network data. In: Carrington P, Scott J, Wasserman S (eds) *Models and methods in social network analysis*. Cambridge University Press, Cambridge, pp 215–247
- Snijders T (2009) Longitudinal methods of network analysis. In: Meyers B (ed) *Encyclopedia of complexity and system science*. Springer, Berlin, pp 5998–6013
- Snijders T, Steglich C, Schweinberger M (2007) Modeling the co-evolution of networks and behavior. In: van Montfort K, Oud H, Satorra A (eds) *Longitudinal models in the behavioral and related sciences*. Lawrence Erlbaum, Hillsdale, pp 41–71
- Snijders T, van de Bunt C, Steglich C (2010a) Introduction to actor-based models for network dynamics. *Soc Netw* 32:44–60
- Snijders TA, Koskinen J, Schweinberger M (2010b) Maximum likelihood estimation for social network dynamics. *Ann Appl Stat* 4:567–588
- Steglich C, Sinclair P, Holliday J, Moore L (2010) Actor-based analysis of peer influence in a stop smoking in schools trial (ASSIST). *Soc Netw* (in press)
- Strauss D (1986) On a general class of models for interaction. *SIAM Rev* 28:513–527
- Van De Bunt GG, Van Duijn MAJ, Snijders T (1999) Friendship networks through time: An actor-oriented dynamic statistical network model. *Comput Math Organ Theory* 5(2):167–192
- Wooldridge J (2002) *Econometric analysis of cross section and panel data*. MIT Press, Cambridge
- Yatchew A, Griliches Z (1985) Specification error in probit models. *Rev Econ Stat* 67:134–139
- Zweig M, Campbell G (1993) Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 39(8):561–577