

University of Groningen

Contemporaneous Statistics for Estimation in Stochastic Actor-Oriented Co-evolution Models

Amati, Viviana; Schoenenberger, Felix; Snijders, Tom A. B.

Published in:
 Psychometrika

DOI:
[10.1007/s11336-019-09676-3](https://doi.org/10.1007/s11336-019-09676-3)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
 Publisher's PDF , also known as Version of record

Publication date:
 2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Amati, V., Schoenenberger, F., & Snijders, T. A. B. (2019). Contemporaneous Statistics for Estimation in Stochastic Actor-Oriented Co-evolution Models. *Psychometrika*, *84*(4), 1068-1096.
<https://doi.org/10.1007/s11336-019-09676-3>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy


If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure). <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

CONTEMPORANEOUS STATISTICS FOR ESTIMATION IN STOCHASTIC ACTOR-ORIENTED CO-EVOLUTION MODELS

VIVIANA AMATI  AND FELIX SCHÖNENBERGER

ETH ZÜRICH

TOM A. B. SNIJDERS 

UNIVERSITY OF GRONINGEN

UNIVERSITY OF OXFORD

Stochastic actor-oriented models (SAOMs) can be used to analyse dynamic network data, collected by observing a network and a behaviour in a panel design. The parameters of SAOMs are usually estimated by the method of moments (MoM) implemented by a stochastic approximation algorithm, where statistics defining the moment conditions correspond in a natural way to the parameters. Here, we propose to apply the generalized method of moments (GMoM), using more statistics than parameters. We concentrate on statistics depending jointly on the network and the behaviour, because of the importance of their interdependence, and propose to add contemporaneous statistics to the usual cross-lagged statistics. We describe the stochastic algorithm developed to approximate the GMoM solution. A small simulation study supports the greater statistical efficiency of the GMoM estimator compared to the MoM.

Key words: generalized method of moments, networks, behaviour, panel data, stochastic actor-oriented model, stochastic approximation.

1. Introduction

Stochastic actor-oriented models (SAOMs) are models for network panel data, that is observations of a network and a nodal variable referred to as “behaviour” at several time points t_1, \dots, t_M (Snijders, 2017a). In their first formulation (Snijders, 1996, 2001), SAOMs were developed to explain the network evolution, i.e. the changes over time. A recent extension (Snijders et al., 2007; Steglich et al., 2010) allows modelling the co-evolution of networks and behaviours, where behaviours are changing covariates of the actors. Examples of networks and behaviours co-evolving over time are friendship networks and delinquent behaviour (Burk et al., 2008; Haynie et al., 2014), advice networks and work climate perception (Schulte et al., 2012), and alliance networks and reputation (Ebbens and Wijnberg, 2010), to name a few. A full list of publications using co-evolution SAOMs is available at <http://www.stats.ox.ac.uk/~snijders/sienal/>.

The term “co-evolution” indicates that behaviours and networks are inter-dependent: ties and behaviours can indeed change as a function of both the structure of the network and the

Part of this research has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007–2013)/ERC Grant Agreement No. 319209.

Correspondence should be made to Viviana Amati, Social Networks Lab, Department of Humanities, Social and Political Sciences, ETH Zurich, Weinbergstrasse 109, 8092 Zurich, Switzerland. Email: viviana.amati@gess.ethz.ch

behaviours of the actors. For instance, actors might have a tendency to create ties to actors that occupy a certain position in the network or have a certain behaviour. In a similar way, actors may change their behaviour to conform to the behaviour of the actors to whom they are connected. Thus, tie changes can affect behavioural changes and, vice versa, behavioural changes can affect tie changes. This suggests that the tie and the behavioural dynamics closely intertwine and cannot be studied separately.

Co-evolution SAOMs have become widely used since the analysis of the co-evolution of networks and behaviours allows researchers to address questions related to selection and influence processes. In co-evolution SAOMs the network and the behaviour are jointly the two dependent variables. The dependence of network dynamics on the network-behavioural configuration is referred to as the selection process, whereas the dependence of behavioural dynamics on the network-behavioural configuration is referred to as the influence process. While the selection process concerns the way actors link to others in the network, the influence process relates to the way actors adjust their behaviour to the behaviour of the actors to whom they are tied. Influence and selection may result in the same outcome, namely the fact that connected actors show a similar behaviour, also known as homophily (McPherson et al., 2001). Homophily can indeed stem from the propensity to connect to actors who have a similar behaviour (selection) and/or from the proclivity for adjusting the behaviour to the behaviour of the other actors in the network (influence).

The simultaneous modelling of selection and influence processes in SAOMs allows distinguishing between their role in the co-evolution of networks and behaviours and estimating separate parameters for these two sub-processes. The parameters of these models are usually estimated using the method of moments (MoM) (Snijders, 2001). Snijders et al. (2007) proposed to distinguish parameters modelling selection—where the dependent variable is the network and the explanatory variables include the behaviour—from those modelling influence—where the dependent variable is the behaviour and the explanatory variables include the network—by using time-lagged statistics where the dependent variable is taken at time t_m while the explanatory variables at time t_{m-1} . These are called cross-lagged statistics and described extensively in Steglich et al. (2010). Contemporaneous statistics, defined by taking the two (or more) dependent variables at the same time point, t_m , are never used for this MoM estimation. However, they do carry information on the simultaneous observation of network and behavioural changes. Here, we present an estimator based on the generalized method of moments (GMoM), using more statistics than parameters, by including contemporaneous statistics in the estimation procedure. This work extends the development by Amati et al. (2015) of a GMoM estimator for SAOMs for dynamics of networks without behaviour.

The remainder of the paper is organized as follows. Section 2 describes the formulation of the SAOM and parameter estimation using MoM as well as GMoM. Section 3 presents contemporaneous statistics and the stochastic algorithm used to compute the GMoM estimator for the co-evolution of networks and behaviour. A small simulation study analysing the finite-sample properties of the GMoM estimator and investigating its efficiency relative to the MoM estimator is reported in Section 4. Finally, an illustrative example of the use of GMoM is reported in Section 5. The paper finishes in Section 6 with a discussion.

2. Stochastic Actor-Oriented Models

2.1. Data

For network and behaviour panel data, we use the following notation. Let $\mathcal{N} = \{1, \dots, n\}$ be a set of entities, hereafter referred to as actors, over which a binary relation $\mathcal{R}: \mathcal{N} \times \mathcal{N} \rightarrow \{0, 1\}$

is defined. This relation can be regarded as a digraph on the node set \mathcal{N} and will be called the network. It is identified with the adjacency matrix $X_{n \times n}$, whose cell X_{ij} takes the value $\mathcal{R}(i, j)$. The relation is assumed to be non-reflexive, thereby implying that $X_{ii} = 0$, and directed, so that $X_{ij} = 1$ does not imply that $X_{ji} = 1$. When $X_{ij} = 1$, we say there is a tie from i to j , or also that i is tied to j . Further, let Z be an n -dimensional vector with Z_i describing the behaviour of actor i . Until recently, the formulation of SAOMs has required that the behaviour Z is an ordinal variable taking integer values in $[1, c]$, $c \in \mathbb{N}_{>1}$. A recent development of the model allows the behaviour Z to be also a continuous variable (Nieznik and Snijders, 2017; Nieznik et al., 2019). The current paper assumes a discrete ordinal behaviour variable.

We assume that the network X and the behaviour Z are stochastic and are subject to change over time. The dependence on time t is denoted by $(X, Z)_t$. Networks and behaviour are observed at $M > 1$ distinct time points t_1, \dots, t_M . Observations are denoted by $x(t_m)$ and $z(t_m)$, $m = 1, \dots, M$. Generally, it is our convention that random variables are denoted by upper case letters and fixed quantities or realizations of random variables are denoted by lower case letters.

Actor- and dyadic-level variables might also be collected besides network and behavioural variables. The former describe constant or time-dependent characteristics of the actors and are denoted by the array $v = (v_1, \dots, v_M)$ where $v_m \in \mathbb{R}^{n \times p}$ is the standard data table at time t_m . The latter describe constant or time-dependent characteristics of dyads and therefore are recorded for each pair of actors. Dyadic variables are represented by the sequence of arrays $w = (w_1, \dots, w_M)$ where $w_m \in \mathbb{R}^{n \times n \times q}$ is an $n \times n \times q$ array whose element w_{ijq} is the value of the q th dyadic covariate over the pair (i, j) at time t_m . Monadic and dyadic covariates are considered to be constant and non-stochastic between two observational periods. Thus, they will always be denoted by lower case letters.

The time series of networks and behaviours $\{(X, Z)_{t_1}, \dots, (X, Z)_{t_M}\}$ together with monadic and dyadic covariates v and w constitute the network panel data. These are assumed to be the outcome of a stochastic process $\{(X, Z)_t, v, w\}$, $t \in \mathcal{T}$. When convenient, monadic and dyadic covariates will be omitted so that the process and its observed states will be simply denoted by $(X, Z)_t$ and $(x, z)_t$, respectively.

2.2. Model Definition

Following the approach originally proposed by Holland and Leinhardt (1977), in SAOMs networks and behaviours are jointly modelled as a continuous-time Markov chain. Given the discrete outcome space for $(X, Z)_t$, this assumption implies that the co-evolution of networks and behaviours can be thought of as the result of a series of unobserved changes happening in continuous time between two consecutive observations $(x, z)_{t_{m-1}}$ and $(x, z)_{t_m}$. Thus, \mathcal{T} is an interval on the real line. In line with Holland and Leinhardt (1977) these steps are assumed to be the smallest possible steps, in which only one variable changes by the smallest possible amount, i.e. $+1$ or -1 , so that the simultaneous occurrence of network and behavioural changes is precluded. The observed change from time t_m to the next observation t_{m+1} is the net result of all these small steps. A convenient interpretation of the model is a sequence of choices by the actors, where at stochastically determined time points, one selected actor has the opportunity to change either one of his outgoing ties or his behaviour. Due to the Markov property, these changes are probabilistically determined only by the current network-behavioural configuration, without any influence deriving from the previous configurations.

Each step of the continuous-time Markov chain defining SAOMs is referred to as a *micro-step* and is defined by the time point a change is going to happen and the change itself. The change process can be therefore decomposed into two sub-processes, the change opportunity process and the change determination process.

The change opportunity process models the waiting time between two consecutive opportunities for a change for an actor i and determines whether a network or a behavioural micro-step is going to take place. Due to the Markov assumption, the waiting time until the next opportunity for a change follows an exponential distribution. The model supposes that network and behavioural changes may happen at different frequencies across the intervals $[t_{m-1}, t_m]$ with rates $\rho_m^{[X]}$ and $\rho_m^{[Z]}$, respectively. In the simplest specification these rates are assumed to be constant over the actors, thereby implying that actors get the opportunities for changing according to a homogeneous Poisson process with rates depending only on the time period $[t_{m-1}, t_m]$. In more complex models the rates might depend on actor variables or their position in the network through a link function [for more details see Snijders et al. (2010a) and Snijders (2017a)]. For this reason the rates are usually referred to as *rate functions*.

The type of the next micro-step is determined in the following way. At each point in time $t \in [t_{m-1}, t_m]$, the waiting times for all the actors are generated according to exponential distributions with parameters $\rho_m^{[X]}$ and $\rho_m^{[Z]}$. The minimum of these waiting times determines the time at which the next micro-step is happening, the type of the micro-step and the actor who is going to get the opportunity to make a change.

The change determination process models the probabilities of changing and thus defines the transition matrix of the process. In SAOMs the probability of the choice made by the actor i who gets the opportunity for a change is modelled by a multinomial logistic regression model where the linear combination of statistics and explanatory variables is determined by the so-called *evaluation function*. For the random variable $Y \in \{X, Z\}$, the evaluation function is defined as

$$f_i^{[Y]}(\beta, x, z) = \sum_{k=1}^K \beta_k^{[Y]} s_{ik}^{[Y]}(x, z), \quad (1)$$

where K is the number of explanatory variables, β_k is a statistical parameter and $s_{ik}^{[Y]}(x, z)$ are suitable explanatory variables, referred to as effects, representing local configuration of ties seen from the point of view of actor i . Table 1 shows some of the most common explanatory variables. A complete list can be found in Ripley et al. (2019). For more than two waves, it is usually assumed that the parameters β_k are constant over time, though interactions with time-dependent variables can be incorporated to account for time variation.

When a network micro-step happens, the actor i who has the opportunity for a change has the alternatives of either changing one of his outgoing ties (i.e. create a non-existing tie or terminate an existing tie), or maintaining the current configuration of his outgoing ties. Thus, given the current state of the network (x, z) , the next configuration of the network is either (x', z) , stemming from the decision of i to change the tie variable X_{ij} into $1 - X_{ij}$, or (x, z) as the result of doing nothing.

When a behavioural micro-step takes place, the actor i who gets the opportunity for a change has the option of either increasing or decreasing the level of the behaviour, or maintaining its current level. Since SAOMs assume that changes in the behaviour are gradual, actors can increase or decrease the level of their behaviour only by one unit at a time. Thus, given the opportunity for a behavioural change, the next state of the chain is (x, z') when actor i changes the behavioural variable Z_i into $Z_i \pm 1$ or (x, z) as the result of doing nothing.

The probability of moving from (x, z) to either (x', z) or (x, z') , conditional on whether a network or a behavioural micro-step is taking place, is then computed as

$$P_{(x', z)}^{[X]} = \frac{\exp\left(f_i^{[X]}(\beta, x', z)\right)}{\sum_{x''} \exp\left(f_i^{[X]}(\beta, x'', z)\right)} \quad \text{and} \quad P_{(x, z')}^{[Z]} = \frac{\exp\left(f_i^{[Z]}(\beta, x, z')\right)}{\sum_{z''} \exp\left(f_i^{[Z]}(\beta, x, z'')\right)}, \quad (2)$$

where the sum in the denominators is over all possible options in this micro-step. Those probabilities are parameterized as a multinomial logit model, and therefore, the parameters β_k can be interpreted as conditional log-odds ratios. Positive (negative) values of the parameter β_k indicate that the actors have a tendency towards (against) the creation or maintenance of ties leading to local network configurations of type $s_{ik}^{[Y]}(x, z)$.

One intuitive interpretation of the probabilities in (2) derives from rational choice theory (Train, 2009), which attempts to explain how decision makers choose among a set of alternatives. According to this approach, the actors in a network can be thought of as decision makers and their utility is the sum of a deterministic term, which is the evaluation function, and a random term. If the random term follows an Extreme Type I distribution, then the probability that an actor i takes a certain action is expressed in a closed form by the multinomial logit model (Luce and Suppes, 1965; McFadden, 1973). This means that, when actors get the opportunity to make either a network or a behavioural change, they choose the new network-behavioural configuration providing them the highest utility. The utility is then construed as the satisfaction of the actors with the new network-behavioural configuration. The interpretation then is that the effects characterize the network configurations that the actors endeavour to form, for positive parameter values, and to avoid, for negative parameter values.

The above defines the probability distribution of the stochastic process $(X, Z)_t$, as a continuous-time Markov chain with discrete state space. The transition distribution of such processes is characterized by the intensity matrix (Norris, 1997), which can be understood as the instantaneous rate at which a given state (x, z) changes into another state (x', z') . Due to the assumptions stated above and the specification of the waiting times and the transition matrix described in the previous lines, the intensity matrix of this process is given by

$$Q((x, z), (x', z'), \theta) = \begin{cases} \rho_m^{[X]} p_{(x', z')}^{[X]} & \text{if } z' = z \\ \rho_m^{[Z]} p_{(x, z')}^{[Z]} & \text{if } x' = x \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where θ is the vector of parameters of a SAOM, and (x, z) and (x', z') differ in exactly one variable x_{ij} or z_i .

2.3. Parameter Estimation

The estimation of the parameter of SAOMs is not straightforward since the complete series of micro-steps is not observed. Several estimation methods have been developed for panel network data: the method of moments (Snijders, 2001; Snijders et al., 2007) and its extension the generalized method of moments (Amati et al., 2015), Bayesian procedures (Koskinen and Snijders, 2007) and maximum likelihood methods (Snijders et al., 2010a). In the following, we consider only the method of moments and the generalized method of moments.

2.3.1. Method of Moments Let $\theta = (\rho^{[X]}, \rho^{[Z]}, \beta^{[X]}, \beta^{[Z]})$ be the p -dimensional vector of parameters of a SAOM. Estimation is performed conditional on the first observation $(x, z)_{t_1}$. Consequently, there is no need to formulate assumptions concerning the mechanisms that may have generated the first observation, and the estimated parameter concerns only the evolution of the network and behaviour during the observation period $[t_1, t_M]$. In the following notation, we omit the conditioning on the first observation.

The method of moments (MoM) for estimation in SAOMs (Snijders, 2001) requires choosing a suitable set of p statistics $s(X, Z) = (s_1(X, Z), \dots, s_p(X, Z))$, and to estimate θ with the value $\hat{\theta}$ such that the expected value of the vector of the statistics $E_{\theta}[s(X, Z)]$ is equal to its sample counterpart $s(x, z)$, i.e.

$$E_{\theta}[s(X, Z)] = s(x, z). \tag{4}$$

The statistics are chosen in a heuristic way since criteria, such as reduction to sufficiency, cannot be applied due to the complexity of the model. Therefore, the statistics should be sensitive to the parameter value in the sense that

$$\frac{\partial}{\partial \theta_p} E_{\theta}[s_p(X, Z)] > 0.$$

In other words, higher values of the parameter θ_p should lead to higher values of the corresponding statistic $s_p(X, Z)$.

The rate parameter models the frequency at which the actors get the opportunity to make a change in either the network or the behaviour across the interval $[t_{m-1}, t_m]$. Consequently, sensitive statistics for the parameters $\rho_m^{[X]}$ and $\rho_m^{[Z]}$ are

$$s_{\rho}^{[X]}(X(t_m), X(t_{m-1})) = \sum_{ij} |X_{ij}(t_m) - X_{ij}(t_{m-1})| \tag{5}$$

$$s_{\rho}^{[Z]}(Z(t_m), Z(t_{m-1})) = \sum_i |Z_i(t_m) - Z_i(t_{m-1})|, \tag{6}$$

respectively, and are counts of the number of network and behavioural changes between two consecutive observations. The corresponding moment equations in (4) are

$$E_{\theta}[s_{\rho}^{[X]}(X(t_m), X(t_{m-1})) | (X, Z)_{t_{m-1}}] = \sum_{ij} |x_{ij}(t_m) - x_{ij}(t_{m-1})| \quad m = 2, \dots, M$$

$$E_{\theta}[s_{\rho}^{[Z]}(Z(t_m), Z(t_{m-1})) | (X, Z)_{t_{m-1}}] = \sum_i |z_i(t_m) - z_i(t_{m-1})| \quad m = 2, \dots, M. \tag{7}$$

For the parameters $\beta_k^{[Y]}$ in the network ($Y = X$) and behaviour ($Y = Z$) evaluation functions, higher values of the parameters suggest that configurations of type $s_k^{[Y]}(x, z)$ are desirable for the actors. For instance, a positive value for the reciprocity parameter indicates that actors have a preference for creating ties that reciprocate an existing incoming tie. Thus, suitable statistics for the parameters of the evaluation function have the form

$$s_k^{[Y]}(X, Z) = \sum_i s_{ik}^{[Y]}(X, Z) \tag{8}$$

and are counts of local configurations representing the mechanisms that might have lead to the observed networks and behaviours. For more than two waves, since the assumption is made that parameters are constant over time (else interactions with time-dependent variables can be

incorporated), and because of the Markov assumption, this is combined over several waves by the moment equation

$$\sum_{m=2}^M E_{\theta} \left[s_{\chi}^{[Y]}(X, Z)_{t_m} \mid (X, Z)_{t_{m-1}} \right] = \sum_{m=2}^M s_{\chi}^{[Y]}(x, z)_{t_m}. \quad (9)$$

This procedure does not work, however, when the same effect $s_{\chi}(X, Z)$ is used to model both the dynamics of the network and the dynamics of the behaviour. An example in Table 1 is the outdegree effect $z_i \sum_j x_{ij}$ on the behaviour which also occurs as the covariate ego effect on the network when Z is taken as the covariate. Also it will not work well when closely related effects are used, such as, in the same table, the average similarity and the covariate similarity effects for covariate Z which are linearly dependent. Intuitively, in these situations, we are using the same information for estimating two parameters. Formally, if the same statistic is used to estimate two different parameters, the resulting system of moment equations has more unknowns than moment equations. Such a system either does not have any solution or has infinitely many solutions (Strang, 1976; Meyer, 2000). This implies that the parameter cannot be estimated uniquely. In a similar way, if two statistics are linearly dependent, the corresponding moment equations are redundant and one of them should be discarded from the system. Thus, the system of moment equations has again more unknowns than moment equations. To deal with this issue, Snijders et al. (2007) proposed to distinguish effects modelling selection (dependent variable: network) from those modelling influence (dependent variable: behaviour) by using time-lagged variables: the dependent variable is taken at time t_m , whereas the explanatory variables are taken at time t_{m-1} . It follows that the combination of values of X at time t_m and Z at time t_{m-1} are used to define the statistics for the estimation of the parameters of the network evaluation function

$$s_{\chi}^{[X]}(X(t_m), Z(t_{m-1})) = \sum_i s_{i\chi}^{[X]}(X(t_m), Z(t_{m-1})), \quad (10)$$

while the combination of values of Z at time t_m and X at time t_{m-1} is used to define the statistics for the estimation of the parameters of the behavioural evaluation function

$$s_{\chi}^{[Z]}(X(t_{m-1}), Z(t_m)) = \sum_i s_{i\chi}^{[Z]}(X(t_{m-1}), Z(t_m)). \quad (11)$$

The solution of the system composed of the moment equations in (7) and (9) cannot be computed analytically since the expected values of the statistics do not admit a closed form, except for some trivial cases as shown by Snijders and van Duijn (1997). Therefore, the value of the MoM estimator for θ is approximated using a stochastic algorithm (Snijders, 2001) based on a Monte Carlo approximation of the expected values and the Robbins–Monro step (Robbins and Monro, 1951). A detailed description of the algorithm as it was originally implemented can be found in Snijders (2001). Subsequent changes are described in Ripley et al. (2019) and Snijders (2017b).

2.3.2. The Generalized Method of Moments In the MoM, the moment conditions represent the information used to estimate the unknown parameters of a model. In general, it is desirable to account for as much information as possible in the estimation process. However, the MoM does not allow to include a number q of moment equations that is larger than the number of parameters p . Using $q > p$ moment equations results indeed in an over-identified system of equations which

TABLE 1.

A list of some common effects used to specify the network and behavioural evaluation functions in stochastic actor-oriented models.

Effect	$s_{ik}^{[X]}(X, Z)$
<i>Network</i>	
Density	$\sum_j x_{ij}$
Reciprocity	$\sum_j x_{ij} x_{ji}$
Transitive triplets	$\sum_{j,h} x_{ij} x_{ih} x_{hj}$
Transitive ties	$\sum_j x_{ij} \max_h (x_{ih} x_{hj})$
3-cycle	$\sum_{j,h} x_{ij} x_{jh} x_{hi}$
Outdegree popularity (sqrt)	$\sum_j x_{ij} \sqrt{\sum_h x_{jh}}$
Covariate ego	$v_i \sum_j x_{ij}$
Covariate alter	$\sum_j x_{ij} v_j$
Covariate similarity	$\sum_j x_{ij} (\text{sim}_j^p - \widehat{\text{sim}}_v^p)$
Effect	$s_{ik}^{[Z]}(X, Z)$
<i>Behaviour</i>	
Linear shape	z_i
Quadratic shape	z_i^2
Outdegree	$z_i \sum_j x_{ij}$
Indegree	$z_i \sum_j x_{ji}$
Average similarity	$(\sum_j x_{ij})^{-1} \sum_j x_{ij} (\text{sim}_j^2 - \widehat{\text{sim}}_v^2)$

The quantities in the covariate and average similarity effects are the similarity scores $\text{sim}_j^p = \frac{\Delta - |v_j - v_j|}{\Delta}$, $\Delta = \max_{ij} |v_i - v_j|$, and its mean $\widehat{\text{sim}}_v^p$.

does not admit a solution, unless some of these equations are linearly dependent (Strang, 1976; Meyer, 2000).

The generalized method of moments (GMM) is an extension of the MoM that allows to deal with over-identified systems of moment equations. The GMM (Hall, 2005; Mátyás, 1999) was developed in the econometric literature (Hansen, 1982; Burguete et al., 1982) as an alternative to maximum likelihood estimation when: (i) the probability distribution cannot be correctly specified (e.g. Hansen and Singleton, 1982); (ii) the regularity conditions under which the maximum likelihood estimators have optimal asymptotic properties are not met (e.g. Kim and Frees, 2007; Bollen et al., 2014); (iii) the computation of the maximum likelihood estimation is computationally onerous (e.g. Gallant et al., 1997).

The basic idea of the GMM is to weight the moment conditions to obtain an asymptotically optimal estimator. Instead of looking for the value of θ satisfying all the moment conditions, the GMM minimizes a quadratic form of the moment conditions. Formally, let θ be a p -dimensional vector of parameters which are to be estimated and $s^*(X, Z)$ a vector of $q > p$ statistics. The GMM estimator for θ is the value

$$\hat{\theta} = \underset{\theta}{\text{argmin}} E_{\theta}[s^*(X, Z) - s^*(x, z)]' W E_{\theta}[s^*(X, Z) - s^*(x, z)], \quad (12)$$

where $W_{q \times q}$ is a matrix of weights assigning importance to each moment condition. An optimal choice for W is the inverse of the variance-covariance matrix of $E_{\theta}[s^*(X, Z)]$ (Hansen, 1982), so

that the moment conditions are combined by accounting for the different variances of the statistics and for their correlations.

It must be noted that when the number of moment conditions equals the number of unknown parameters and W is the identity matrix, the GMM reduces to the regular MoM. Moreover, since W is a positive semi-definite matrix, the quadratic form $E_{\theta}[s^*(Y) - s^*(y)]' W E_{\theta}[s^*(Y) - s^*(y)]$ is a meaningful distance measure and thus the formulation of the GMM follows the principle of the Chi-squared method of Neyman and Pearson (Pearson, 1900; Neyman and Pearson, 1928), which is indeed a statistical antecedent of the GMM.

In SAOMs the information for the parameter estimation is provided by the statistics described in the previous section. However, several other statistics sensitive to the parameter of SAOMs are available and their use might improve the estimation. For instance, for the parameters of the network evaluation function related to the effects that depends only on the network (e.g. reciprocity and transitivity), the statistics account only for the observation of the network at time t_m . Amati et al. (2015) argued that this set of statistics can be supplemented by statistics considering both observations of the network at time t_m and t_{m-1} , and showed that the use of the GMM to account for these statistics might improve the estimation.

In the following section, we extend the work of Amati et al. (2015) and concentrate on effects that depend on both network and behaviour. We motivate and define a new set of statistics for modelling selection and influence mechanisms and describe the algorithm that is used to obtain the estimates.

3. GMM for the Co-evolution of Networks and Behaviours

3.1. Contemporaneous Statistics

The distinction between effects modelling selection and those modelling influence lies in the use of cross-lagged statistics, where the dependent variable is taken at time t_m and the independent variable at time t_{m-1} , as shown in Equations (10) and (11). Contemporaneous statistics, where all the dependent variables (i.e. networks and behaviours) are taken at the same time point t_m , are never used in this estimation. However, they do carry valuable information on the simultaneous observation of network and behaviour, which is supplementary to the information provided by the cross-lagged statistics. We clarify this by way of an example.

Let us consider the following statistics for the network and the behavioural evaluation function

$$\text{Covariate ego } s_{\text{ego}}^{[X]}(X, Z) = \sum_{ij} X_{ij}(t_m) Z_i(t_{m-1}) \quad (13)$$

$$\text{Same covariate } s_{\text{sam}}^{[X]}(X, Z) = \sum_{ij} X_{ij}(t_m) \mathbb{I}\{Z_i(t_{m-1}) = Z_j(t_{m-1})\} \quad (14)$$

$$\text{Outdegree } s_{\text{out}}^{[Z]}(X, Z) = \sum_{ij} X_{ij}(t_{m-1}) Z_i(t_m). \quad (15)$$

where $\mathbb{I}\{Z_i(t_{m-1}) = Z_j(t_{m-1})\}$ is an indicator function taking value 1 when i and j have the same level of the behaviour and 0 otherwise. The covariate ego and the same covariate effects are used to model selection. While the former expresses the mechanism that actors having a higher level of the behaviour are more likely to create ties than others, the latter expresses the tendency of actors to create ties with those that are similar to themselves with respect to the behaviour. The outdegree effect is used to model influence and refers to the propensity of actors having a given outdegree to increase or decrease their level of the behaviour.

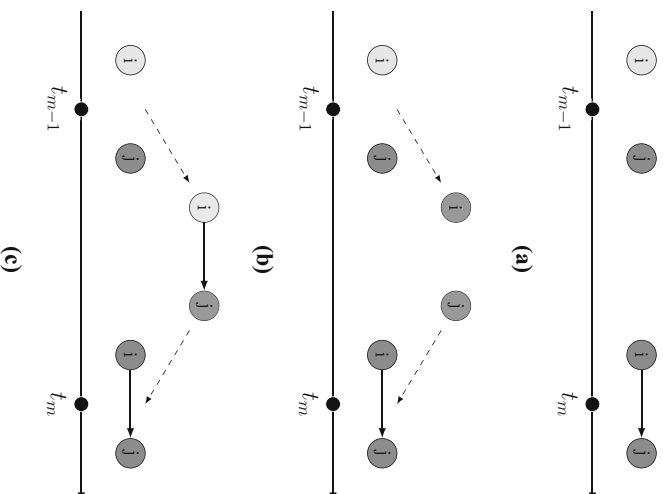


FIGURE 1.

Concurrent change in a network. **a** Actor i made a network and a behavioural change between times t_{m-1} and t_m , **b** selection mechanism: first actor i changed his behaviour and then created the tie towards actor j , **c** influence mechanism: first actor i created the tie towards actor j and then adapted his behaviour to that of actor j .

Let us now consider the scenario presented in Figure 1a, depicting the observation of the dyad (i, j) at times t_{m-1} and t_m . The colour of the nodes represents the behaviour: the darker the colour, the higher the level of the behaviour. At time t_{m-1} , i and j are disconnected and show a different value of the behaviour. At time t_m , i and j are connected and have the same level of the behaviour. This means that, according to the process assumed by SAOMs, actor i changed both one of his outgoing ties and his behaviour during the observation period $[t_{m-1}, t_m]$: he created the tie towards j and adjusted the level of his behaviour to that of actor j . Changes in both the behaviour and the set of outgoing ties between two consecutive observations of a network and a behaviour are hereafter referred as to “concurrent changes.”

Depending on the order in which these changes took place, the observation at time t_m might be the result of a selection or an influence process. The former would have happened if first i had changed his behaviour and then created the tie towards j (Figure 1b), the latter if these changes had taken place in the opposite order (Figure 1c). Since the complete series of micro-steps was not observed, we cannot establish whether the resulting configuration of the network and behaviour for actor i was due to a selection or an influence process. However, only the behaviour ego statistic defined in Equation (13) accounts for these situations. The contribution to the same behaviour statistic in (14) is indeed null since i and j show a different level of the behaviour at time t_{m-1} and thus $\mathbb{I}\{Z_i(t_{m-1}) = Z_j(t_{m-1})\} = 0$. Similarly, the contribution to the outdegree statistic in Equation (15) is also null since there was no tie from i to j at time t_{m-1} and thus $X_{ij}(t_{m-1}) = 0$.

The same reasoning can be applied to any statistic that depends on both network and behaviour (e.g. the similarity and the total similarity statistics) and is therefore affected by concurrent changes of the network and the behaviour.

The example above shows that when an actor i makes concurrent changes of network and behaviour within two consecutive panel observations, the cross-lagged statistics might not account for them. It follows that in situations where concurrent changes of network ties and behavioural levels happen on a regular basis, additional information can be gained by considering the contemporaneous statistics.

In this context, contemporaneous statistics that consider the network and the behaviour at the same time point t_m , and which are defined as

$$s_k^{[XZ]}(X(t_m), Z(t_m)) = \sum_i s_{ik}^{[XZ]}(X(t_m), Z(t_m)), \quad (16)$$

provide indeed a way to include the information on concurrent changes in the estimation procedure. In the following, we denote the contemporaneous statistics by indicating both the network and behavioural variable at the superscript of the statistic.

Contemporaneous statistics corresponding to the statistics in Equations (13)–(15) are therefore denoted by

$$\begin{aligned} s_{\text{ego}}^{[XZ]}(X, Z) &= \sum_{ij} X_{ij}(t_m) Z_i(t_m) \\ s_{\text{sam}}^{[XZ]}(X, Z) &= \sum_{ij} X_{ij}(t_m) \mathbb{I}\{Z_i(t_m) = Z_j(t_m)\} \\ s_{\text{out}}^{[XZ]}(X, Z) &= \sum_{ij} X_{ij}(t_m) Z_i(t_m). \end{aligned} \quad (16)$$

The joint use of cross-lagged and contemporaneous statistics for the estimation of the parameter θ of the SAOMs leads to an over-identified system of moment equations which requires the use of the GMM to be solved. As already mentioned, the expected values of the statistics do not admit a closed form and thus the GMM solutions for the parameter θ are approximated using the procedure proposed by Amati et al. (2015), which is an adaptation of the stochastic approximation algorithm developed by Snijders (2001) used to approximate the MoM solution.

3.2. Stochastic Approximation Algorithm for the GMM

Let θ be a p -dimensional vector of parameters corresponding to a certain specification of SAOMs and $s^*(X, Z)$ a vector of $q > p$ statistics, including the contemporaneous statistics defined in the previous section. The GMM estimator for θ is the value defined in Equation (12). For more than two waves this expression could be elaborated similar to how (7) and (9) are elaborations of (4); to keep the notation transparent, we stick to (12).

The quadratic form $E_\theta[s^*(X, Z) - s^*(x, z)]' W E_\theta[s^*(X, Z) - s^*(x, z)]$ may be assumed to be convex since W , being the inverse of the covariance matrix of the statistics, is a symmetric and positive semi-definite matrix.¹ Thus, the critical point of the quadratic form is a minimum and the estimate for θ is the value such that

$$\Gamma' W E_\theta [s^*(X, Z) - s^*(x, z)] = B E_\theta [s^*(X, Z) - s^*(x, z)] = 0, \quad (17)$$

¹Because of the difficulty of these network models this is not something we can prove, but it is supported by all practical evidence.

with $\Gamma = \frac{\partial}{\partial \theta} E_{\theta} [s^*(X, Z) - s^*(x, z)]'$ the $p \times q$ matrix of the first-order derivatives of the statistics, and $B = \Gamma W$. Equation (17) is a system of p equations about expected values in p unknown variables just like (4). Thus, its solution $\hat{\theta}$ can be obtained by adapting the stochastic algorithm used to solve Equation (4) for the MoM (Snijders, 2001).

The stochastic algorithm is based on the Robbins–Monro method for stochastic approximation (Robbins and Monro, 1951) which is characterized by the iteration step

$$\hat{\theta}_{r+1} = \hat{\theta}_r - \alpha_r D^{-1} B [s^*(x^{(r)}, z^{(r)}) - s^*(x, z)], \tag{18}$$

where $s^*(x^{(r)}, z^{(r)})$ is a simulation of $s^*(X, Z)$ from the SAOM with parameter $\hat{\theta}_r$, D is a suitable $p \times q$ matrix and α_r is a sequence of numbers converging to 0.

The algorithm consists of three phases, and it is only sketched here. A detailed description is provided in the Appendix.

The first phase is a preliminary phase used to approximate the quantities involved in the Robbins–Monro step: the matrix of first-order derivatives Γ , the matrix of weights W and the matrix $D = B \Gamma'$ of the first-order derivatives of the function $B E_{\theta} [s^*(X, Z) - s^*(x, z)]$. We denote these approximations by Γ_0 , W_0 and D_0 , respectively.

Given an initial guess θ_0 for the parameter θ , n_1 trajectories of the co-evolution of networks and behaviours are simulated. Then, the vectors $s^*(x^{(h)}, z^{(h)})$, $h = 1, \dots, n_1$ are computed and used to estimate the weighting matrix W and the matrix of first-order derivatives Γ' . The matrix of weights W is approximated by

$$W_0 = \hat{\Sigma}^{-1},$$

where $\hat{\Sigma}$ is a Monte Carlo estimate of the covariance matrix of the statistics. Compared to the algorithm presented in Amati et al. (2015), the matrix W is computed using the full covariance matrix of the vector of the statistics, rather than a block diagonal structure determined by the rates and the evaluation function parameters. The reason is that this choice gives more accurate estimates for the rate parameters and more stability to the algorithm.

The matrix Γ is approximated using the score function method (Schweinberger and Snijders, 2007) so that

$$\Gamma_0 = \frac{1}{n_1} \sum_{h=1}^{n_1} U_h [s^*(x^{(h)}, z^{(h)}) - s^*(x, z)],$$

where $U_h = \frac{\partial}{\partial \theta} \log [p_h(X, Z)]$ is the score function for the h th simulation. Phase 1 ends with the approximation of the matrix B , given by

$$B_0 = \Gamma_0 W_0,$$

and the update of the value of θ using one partial step of the Newton–Raphson algorithm

$$\hat{\theta} = \theta_0 - \alpha D_0^{-1} B_0 [\bar{s}^*(x, z) - s^*(x, z)],$$

with $\bar{s}^*(x, z)$ the vector of the averages of the simulated statistics, $D_0 = B_0 \Gamma_0'$, and α a number between 0 and 1.

The value $\hat{\theta}$ is the first rough estimate of θ and is subsequently refined in the second phase of the algorithm. This phase is divided into L sub-phases characterized by a decreasing value of α_t (Pflug, 1990) and the use of the diagonal matrix obtained from the matrix D (Ruppert, 1988; Polyak, 1990). In the t th sub-phase, n_{2t} steps are performed. At each step, one trajectory of the co-evolution of network and behaviour is simulated and the value of $\hat{\theta}_t$ is updated via the Robbins–Monro step. At the end of the sub-phase α_t is halved and the estimate for θ is the average of the values $\hat{\theta}_t$ (Ruppert, 1988; Polyak, 1990). This estimate is used as the initial value of θ for the next sub-phase. The average of the values $\hat{\theta}_t$ during the last sub-phase is the final estimate for θ .

In the third phase, n_3 trajectories of the network and behaviour co-evolution are simulated and used to approximate the standard error of the GMoM estimator and to evaluate the convergence of the algorithm. Application of the delta method leads to the following expression for the variance–covariance matrix of the GMoM estimator

$$\Sigma_{\hat{\theta}_{\text{GMoM}}} = (\hat{B} \hat{\Gamma})^{-1} (\hat{B} \hat{\Sigma} \hat{B}') ((\hat{B} \hat{\Gamma})^{-1})',$$

where \hat{B} , $\hat{\Gamma}$ and $\hat{\Sigma}$ are approximations of the matrices B , Γ and Σ computed as in Phase 1.

The convergence of the algorithm is evaluated using t -ratios for convergence, defined as follows. Given Equation (17), the differences between the expected value of the statistics and their sample counterparts weighted by the matrix B should be ideally 0. However, since the algorithm yields only an approximation, these weighted differences will be small but cannot be expected to be exactly equal to 0. To test whether these weighted differences are close enough to 0, the ratios between their mean values and their standard deviation are taken. For parameter p , this is

$$t\text{-ratio}_p = \frac{B(\bar{s}_p^*(x, z) - s_p^*(x, z))}{\sqrt{(B \hat{\Sigma} B')^{pp}}}.$$

As a rule of thumb, when these t -ratios are smaller than 0.1, then convergence is satisfactory.

4. Simulation

Simulations of the network and behaviour co-evolution can be used to investigate the relative efficiency of the GMoM to the MoM estimators. In general, we would expect that the statistical efficiency of the GMoM estimator will be greater since more information is considered. In the following, we present a small simulation study providing some evidence for the finite-sample properties of the GMoM estimator. The results are limited to the example considered, and they cannot be generalized.

The simulated network panel data comprise a directed network with $n = 30$ actors (which is a rather low value) and a behaviour with 5 ordinal categories, both simulated at 4 points in time, as well as a constant binary actor covariate. The first observation of the network and behaviour was generated in the following way: (i) a random network with an expected average degree of 3 and a random behaviour with a uniform distribution were generated; (ii) a simulation from the SAOM specified by the effects and the parameter values described in Table 2 but with rates sufficiently large (15 for the network and 3 for the behavioural rate parameters) was used to go from the totally random network and behaviour to the first observation. Data for the times t_2 , t_3 and t_4 were generated using a SAOM specified by the effects and the parameter values described in Table 2.

The GMoM estimates were obtained by including the contemporaneous statistic related to the similarity effects of the network and the behavioural function defined as

$$s^*(X, Z) = \sum_m \sum_{ij} x_{ij}(t_m) \left(\frac{\Delta - |z_i(t_m) - z_j(t_m)|}{\Delta} - \widehat{\text{sim}}_z \right),$$

where $\Delta = 4$ is the range of the behaviour and $\widehat{\text{sim}}_z$ is the mean of the similarity score ($\Delta - |z_i(t_m) - z_j(t_m)|$)/ Δ .

Two sets of simulations were performed. They differ in the values of the network and behavioural rate functions and consequently in the percentage of concurrent changes between two consecutive observations of the network and behaviour. We simulated 500 network panel data. For each simulation, the MoM and the GMoM estimates were computed twice so that we could account for both the variability deriving from the data simulations and the variability of the estimates due to the stochastic approximations. To compare the relative efficiency of the GMoM to the MoM, we tested the equality of the root-mean-squared errors (RMSEs) by means of the Wilcoxon signed-rank test with the alternative hypothesis—the RMSEs of the GMoM are lower than the RMSEs of the MoM. The results are presented in Table 2 which reports the mean parameter estimates, the mean standard errors, the average of the RMSEs and the p values attained by the Wilcoxon signed-rank test.

The first set of simulation was characterized by lower rates. The percentage of actors taking concurrent changes of network and behaviour between any two observations varied between 14% and 34% with an average of 25% overall the simulations. In general, the averages of the MoM and the GMoM estimates are close, but the mean of the estimated standard error is lower for the GMoM than for the MoM. The comparison of the RMSE of the two estimators indicates that the GMoM outperforms the MoM for many of the effects in the model as indicated by the significance of the Wilcoxon signed-rank test, reported in Table 2. The gain in efficiency mainly affects the parameters related to the outdegree effects (outdegree and outdegree popularity), and the behaviour and the average similarity effects for which a contemporaneous statistic was considered. This is illustrated in Figure 2 which plots the ratios between the average of the RMSEs of the MoM and the GMoM for each statistic. The grey line has intercept 1, and therefore, a dot above this line indicates that the RMSE of the MoM is greater than the RMSE of the GMoM. As we can observe, the ratios of the RMSEs are larger for the outdegree, the outdegree popularity, and the behaviour and the average similarity effects.

The gain in efficiency is due to both the lower bias and the smaller variance of the GMoM estimates as depicted in Figures 3 and 4. These pictures contrast the absolute deviations and the estimated standard errors of the GMoM (y axis) and the MoM (x axis) estimates for each simulation. Dots under the diagonal indicate that the GMoM estimates have a lower absolute deviation or a lower estimated standard error compared to the MoM estimates. For each effect in the model specification, the cloud concentrates around the diagonal and mainly lies under the diagonal for the estimates of the outdegree, outdegree popularity, behaviour similarity and average similarity effects.

The second set of simulations was generated using higher values of the rate parameters. The percentage of actors taking concurrent changes of network and behaviour between any two observations were between 36% and 68% with an average of 51% overall the simulations. The results of the Wilcoxon signed-rank test in Table 2 indicate that the GMoM outperforms the MoM for most of the statistics in the model. Again, the gain in efficiency mainly affects the parameters related to the outdegree effects (i.e. outdegree and outdegree popularity), as well as the behaviour and the average similarity effects as illustrated in Figure 2. Figures 5 and 6 suggest that the gain in efficiency is due to both lower bias and smaller variance.

TABLE 2.

Simulation results, 3 periods for 30 actors: true value of the parameter(θ), mean parameter estimates (Est.), mean estimated standard errors (SE) and the average of the root-mean-squared errors (RMSEs).

	θ	MOM			GMM			Sig.
		Est.	SE	RMSE	Est.	SE	RMSE	
<i>Network dynamics</i>								
Rate (period 1)	3.500	3.228	0.481	0.723	3.239	0.479	0.727	
Rate (period 2)	3.500	3.523	0.558	0.745	3.534	0.555	0.750	
Rate (period 3)	4.000	4.001	0.653	0.865	4.022	0.652	0.875	
Outdegree (density)	-2.000	-1.842	0.463	0.617	-1.932	0.436	0.493	***
Reciprocity	1.800	1.813	0.236	0.318	1.777	0.228	0.310	*
Transitive triplets	0.250	0.203	0.109	0.157	0.209	0.106	0.152	*
3-Cycle	-0.270	-0.216	0.171	0.239	-0.243	0.165	0.222	***
Transitive ties	0.670	0.721	0.236	0.320	0.702	0.232	0.313	
Outdegree popularity (sqrt)	-0.600	-0.673	0.281	0.365	-0.597	0.262	0.294	***
Covariate alter	0.190	0.240	0.233	0.310	0.204	0.222	0.279	***
Covariate ego	0.480	0.485	0.199	0.262	0.483	0.193	0.255	**
Same covariate	0.610	0.594	0.176	0.241	0.590	0.173	0.240	
Behaviour similarity	3.000	2.894	0.816	1.112	2.923	0.680	0.934	***
<i>Behaviour dynamics</i>								
Rate (period 1)	1.000	0.781	0.306	0.482	0.790	0.308	0.486	
Rate (period 2)	2.000	2.023	0.906	1.240	2.051	0.910	1.248	
Rate (period 3)	2.000	2.280	1.178	1.469	2.300	1.165	1.459	
Linear shape	0.050	0.070	0.194	0.288	0.068	0.195	0.291	
Quadratic shape	0.100	0.142	0.107	0.150	0.146	0.107	0.151	
Average similarity	6.000	5.492	2.676	3.518	5.740	2.451	3.208	***
<i>Network dynamics</i>								
Rate (period 1)	7.000	6.634	1.141	1.579	6.638	1.097	1.572	
Rate (period 2)	7.500	7.807	1.478	1.983	7.802	1.428	1.960	**
Rate (period 3)	8.000	8.269	1.573	2.130	8.255	1.520	2.111	
Outdegree (density)	-2.000	-1.965	0.439	0.559	-1.987	0.396	0.446	***
Reciprocity	1.800	1.819	0.205	0.274	1.790	0.197	0.270	*
Transitive triplets	0.250	0.225	0.065	0.090	0.227	0.062	0.086	*
3-Cycle	-0.270	-0.233	0.126	0.174	-0.243	0.122	0.161	***
Transitive ties	0.670	0.711	0.198	0.269	0.677	0.193	0.260	*
Outdegree popularity (sqrt)	-0.600	-0.639	0.234	0.303	-0.603	0.219	0.247	***
Covariate alter	0.190	0.205	0.207	0.270	0.191	0.195	0.245	***
Covariate ego	0.480	0.491	0.171	0.225	0.490	0.157	0.211	***
Same covariate	0.610	0.621	0.149	0.200	0.616	0.139	0.189	***
Behaviour similarity	3.000	2.920	0.963	1.247	2.971	0.698	0.940	***
<i>Network dynamics</i>								
Rate (period 1)	2.000	1.731	0.401	0.638	1.745	0.401	0.633	
Rate (period 2)	2.000	2.045	0.999	1.323	2.088	0.997	1.319	
Rate (period 3)	2.500	2.591	1.407	1.748	2.660	1.423	1.755	
Linear shape	0.050	0.060	0.180	0.263	0.059	0.182	0.263	
Quadratic shape	0.100	0.139	0.097	0.139	0.141	0.097	0.140	
Average similarity	6.000	5.252	2.602	3.415	5.680	2.378	3.099	***

The two blocks represent two set of simulations differing in the values of the network and behavioural rate functions. The last column (Sig.) reports the significance of the Wilcoxon signed-rank test comparing the RMSE of the two estimators (* p value < 0.1; ** p value < 0.05; *** p value < 0.01).

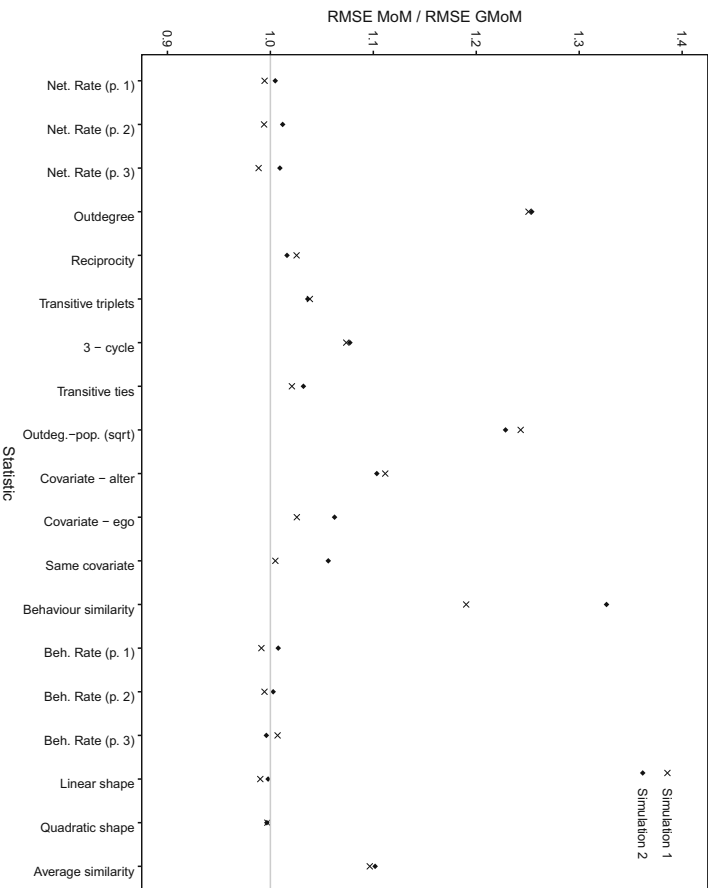


FIGURE 2.

Ratio between the RMSEs of the MoM and the RMSEs of the GMoM for the two blocks of simulations reported in Table 2. The grey line has intercept 1, which is the ratio under the null hypothesis that the GMoM is as efficient as the MoM.

A comparison between the two simulation studies with respect to the standard errors as reported in Figures 4 and 6 and with respect to the root-mean-squared errors presented in Table 2, shows that for the second set of simulations the efficiency gain is larger, suggesting that as the number of concurrent changes increases, the efficiency gain of the GMoM relative to the MoM is greater.

5. Example

To demonstrate the use of the GMoM we considered data collected by Mitchell and West (1996) under the “Teenage Friends and Lifestyle Study.” The dataset was collected on a cohort of 160 students followed over a two year period starting in February 1995, when the pupils were aged 13, and ending in January 1997. The friendship network of the pupils was observed at three time points. Pupils were asked to name up to six friends and provide information on their socio-demographic characteristics along with the use of substances, such as tobacco and alcohol consumption. This data set was also analysed in Steglich et al. (2010). Here, we use the data extended with information about the distance between the houses where the students lived and a different model specification. The set of statistics was chosen according to sociological theories concerning friendship network and behavioural changes as well as the analysis of the goodness of fit of the model. The data is downloadable from the website <http://www.stats.ox.ac.uk/~snijders/siena/>.

In the following, we analyse the co-evolution of the friendship network composed of the 129 pupils who were present at all three observations—through methods for treating composition

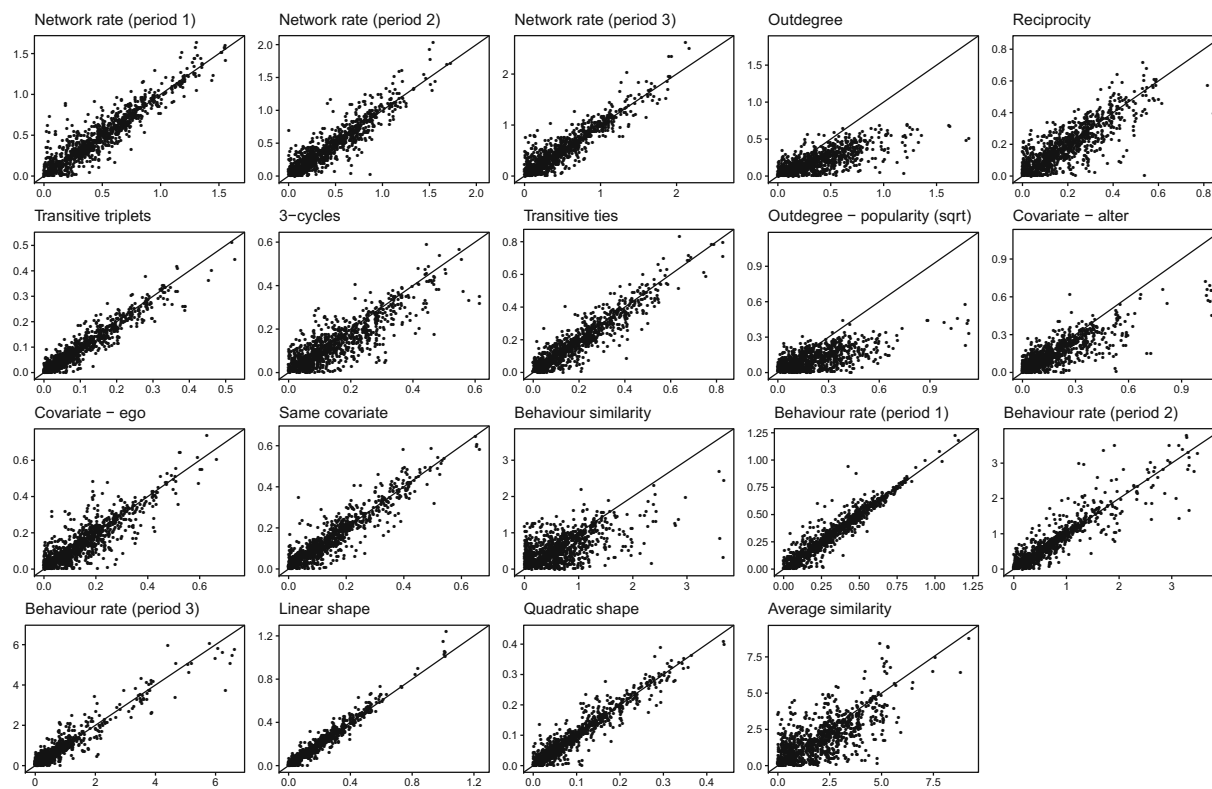


FIGURE 3.
Comparison between the absolute deviations of the MoM (x axis) and the GMoM (y axis) for the first block in Table 2.

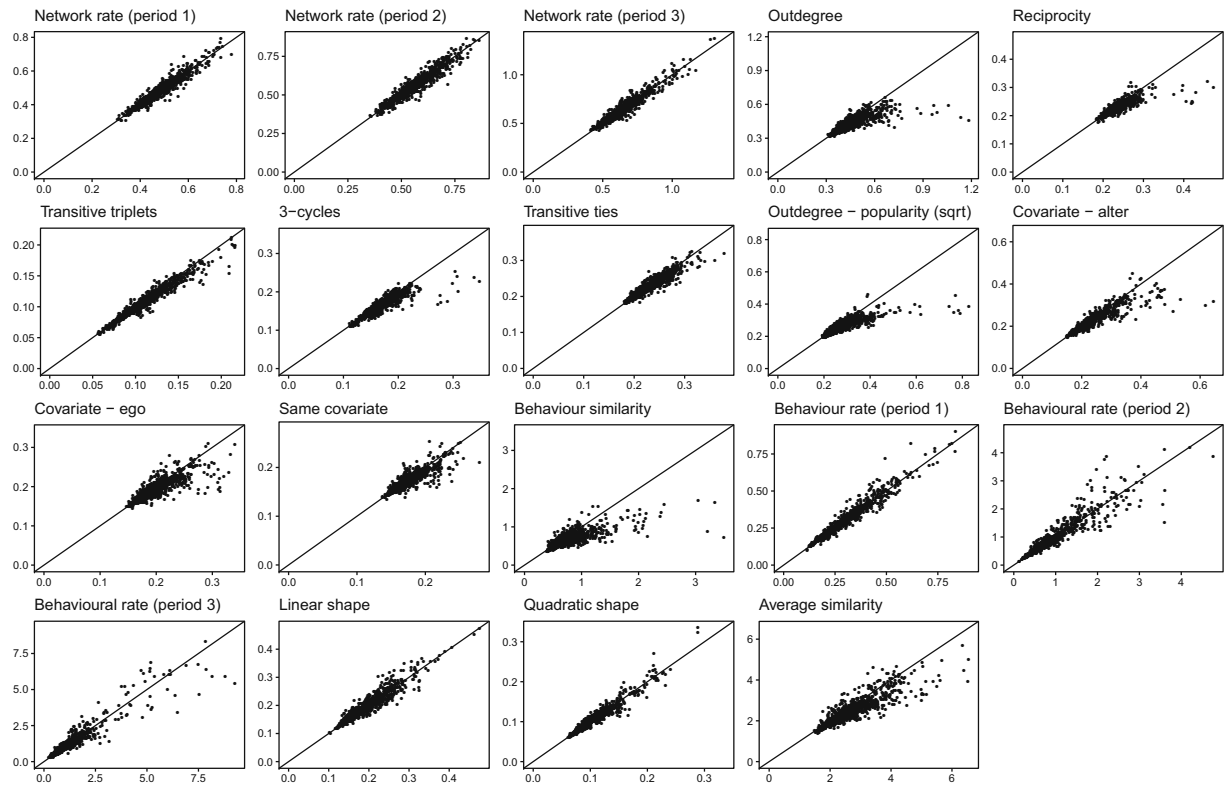


FIGURE 4.
 Comparison between the estimated standard errors of the parameters of the MoM (x axis) and the GMoM (y axis) for the first block in Table 2.

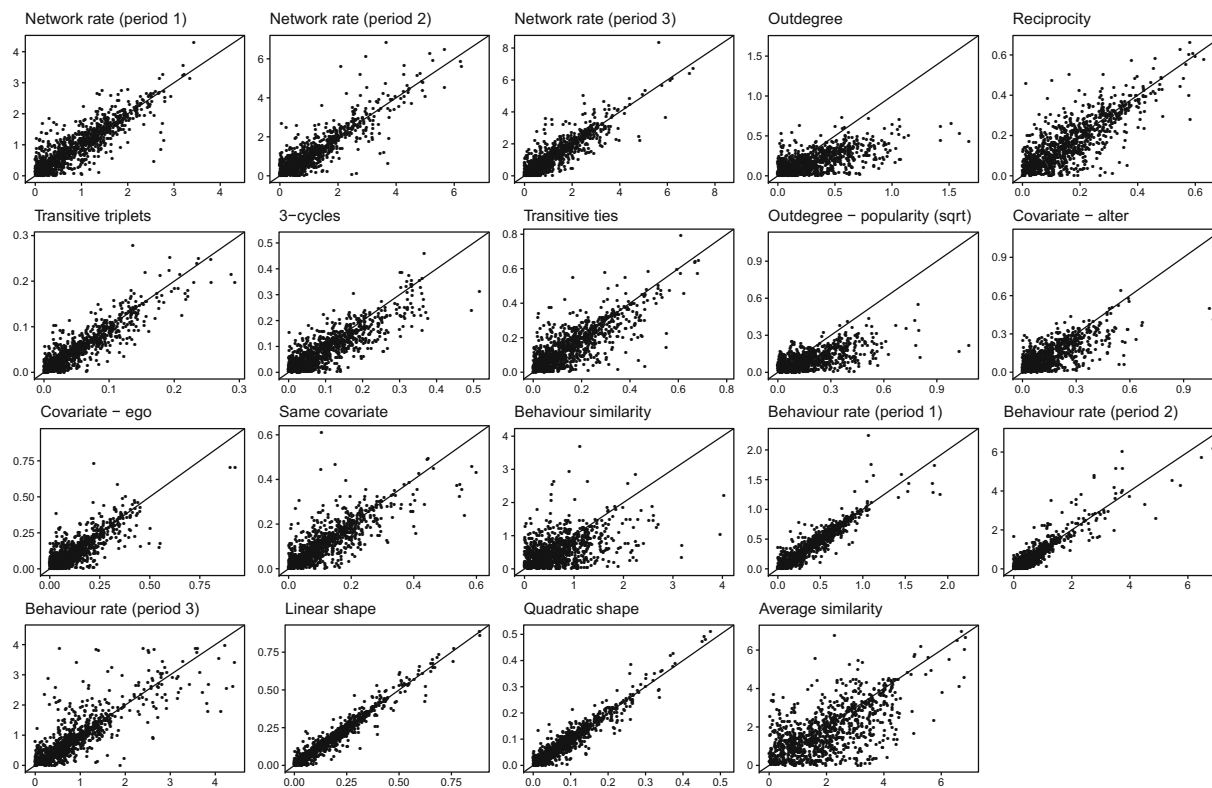


FIGURE 5.

Comparison between the absolute deviations of the parameters estimated with the MoM (x axis) and the GMoM (y axis) for the second block in Table 2.

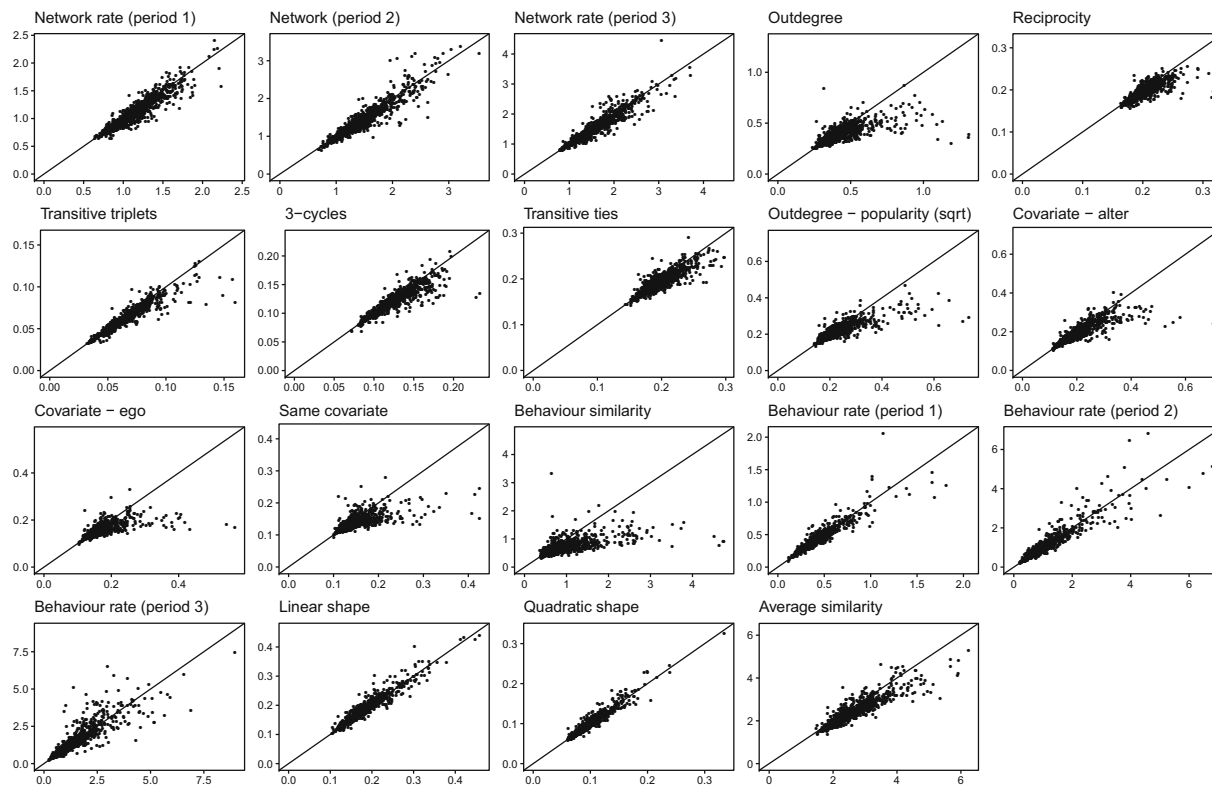


FIGURE 6.
Comparison between the estimated standard errors of the MoM (x axis) and the GMoM (y axis) for the second block in Table 2.

changes in the actor set have been developed (Ripley et al., 2019)—and the drinking behaviour of the pupils measured on a scale of 5 ordinal categories (never, once or twice a year, once a month, once a week and more than once a week). In the analysis we also considered gender as a nodal covariate and the distance between the postcode regions where two pupils lived as a dyadic covariate. We confirmed the model specification to what may be considered current best practice, as laid down in Snijders et al. (2010b) and Ripley et al. (2019), together with Block (2015) and Snijders and Lomi (2019). Network closure was represented by geometrically weighted transitive and cyclic statistics (Hunter, 2007), shortly GWESP statistics, and by the interaction term of transitivity and reciprocity effects (Block, 2015). Given the focus of this paper, we give no further detailed explanation of the model specification and refer to Ripley et al. (2019).

The model was estimated using both the MoM and the GMoM. Three contemporaneous statistics were included in the estimation procedure, and namely:

- The contemporaneous covariate ego statistic

$$s_{\text{ego}}^{[XZ]} = \sum_m \sum_{i_j} X_{ij}(t_m) Z_i(t_m)$$
- The contemporaneous covariate-squared ego statistic

$$s_{\text{egoSq}}^{[XZ]} = \sum_m \sum_{i_j} X_{ij}(t_m) Z_i^2(t_m)$$
- The contemporaneous average alter statistic

$$s_{\text{avAlt}}^{[XZ]} = \sum_m \sum_i \frac{Z_i}{\sum_j X_{ij}(t_m) \sum_j X_{ij}(t_m) Z_j(t_m)}$$

Table 3 reports the parameter estimates obtained by the two methods. The standard errors were computed using 10,000 simulations during phase three of the algorithm. The comparison between the estimated values of the parameters indicates that the MoM and the GMoM estimates are similar. A considerable difference between the estimates is observed for the parameters related to the effects for which the contemporaneous statistics were used, i.e. the drinking ego and the squared drinking ego effects, the drinking ego \times alter effect, and the drinking average alter effects. Comparison between the standard errors shows that most standard errors of the GMoM are smaller than those of the MoM. In agreement with the results of the simulation study, the differences in standard errors are larger for the effects of the network evaluation function related to the drinking behaviour and the behavioural evaluation function.

In spite of the differences in some of the estimated coefficients, the substantive conclusions concerning the mechanisms explaining the co-evolution of friendship and drinking behaviour in this cohort are the same. For testing parameters in these models, it is customarily assumed that the distribution of the estimators is approximately normal; this is supported by extensive simulation results, although not by proofs of asymptotic properties. With this assumption, the parameters are tested using the Wald test, i.e. by computing the ratio of the parameter estimate to its standard error, and testing this in a standard normal distribution.

We first comment on the parameters of the network dynamics. The network rate parameters suggest that on average an actor got 11 and 8 opportunities to change one of his outgoing ties, respectively in the first and second observation period.

TABLE 3.

Estimates (Est.) and standard errors (SE) obtained using the method of moments (MoM) and the generalized method of moments (GMoM) for the data collected by Mitchell and West (1996) under the “Teenage Friends and Lifestyle Study.”

	MoM		GMoM	
	Est.	SE	Est.	SE
<i>Network rate function</i>				
Rate friendship (period 1)	10.899*	1.119	10.920*	1.061
Rate friendship (period 2)	8.840*	0.801	8.872*	0.819
<i>Network evaluation function</i>				
Outdegree (density)	-3.524*	0.291	-3.302*	0.289
Reciprocity	3.354*	0.309	3.238*	0.300
Transitivity GWESP ($\alpha = 0.30$)	3.009*	0.268	2.924*	0.265
Cyclicly GWESP ($\alpha = 0.30$)	0.520*	0.128	0.513	0.123
Indegree popularity	-0.031	0.020	-0.033	0.021
Outdegree activity	0.068	0.043	0.053	0.040
Reciprocal degree activity	-0.131	0.075	-0.099	0.071
Indegree activity	-0.181*	0.048	-0.194*	0.045
Log(distance)	-0.210*	0.042	-0.207*	0.042
Gender alter	-0.072	0.103	-0.072	0.099
Gender ego	-0.037	0.111	-0.032	0.107
Same gender	1.036*	0.172	0.948*	0.167
Reciprocity \times transitivity GWESP ($\alpha = 0.30$)	-1.383*	0.225	-1.403*	0.215
Same gender \times transitivity GWESP ($\alpha = 0.30$)	-0.948*	0.266	-0.827*	0.265
Drinking alter	0.037	0.052	0.029	0.049
Drinking squared alter	-0.103	0.057	-0.091	0.054
Drinking ego	-0.108	0.059	-0.082	0.047
Drinking squared ego	0.030	0.063	-0.003	0.041
Drinking ego \times drinking alter	0.185*	0.052	0.148*	0.042
<i>Behavioural rate function</i>				
Rate drinking (period 1)	1.497*	0.250	1.476*	0.242
Rate drinking (period 2)	2.262*	0.383	2.228*	0.377
<i>Behavioural evaluation function</i>				
Linear shape	0.436*	0.395	0.488*	0.294
Quadratic shape	-0.341*	0.116	-0.259*	0.078
Indegree	0.074	0.087	0.064	0.071
Outdegree	-0.107	0.154	-0.117	0.106
Average alter	0.816*	0.285	0.568*	0.195

An asterisk denotes the significance of the parameter at level 0.05.

The parameters of the network evaluation function indicate preferences of actors to form certain type of ties. The negative outdegree parameter suggests that actors are reluctant to form ties that are not part of any other network configuration, while they are inclined to create ties that reciprocate an existing incoming tie. They also have a tendency to close a transitive or a cyclic triad as indicated by the positive values of the reciprocity and both the GWESP effects. However, the negative parameter of the interaction term between reciprocity and the transitivity GWESP effect suggests that transitivity is less strong if the tie is reciprocated. The negative parameter of the indegree activity shows that popular pupils tend to nominate fewer friends.

For the nodal covariate effect we observed that there is a tendency towards gender homophily, as denoted by the positive values of the same gender parameter. This indicates that there is a strong tendency of pupils to choose friends that have the same gender. Transitivity between pupils of the

same gender is less strong than between pupils of different gender, as attested to by the negative parameter of the interaction term between the same gender and the GWESP effects. The negative value of the parameter related to the distance suggests that pupils tend to be friends with other pupils who live nearby.

The effect of the drinking behaviour on the network evolution is significant only for the drinking ego \times alter covariate. The corresponding coefficient is positive, indicating that pupils with high drinking behaviour tend to relate with other pupils who also drink more. This suggests that the drinking behaviour plays a role in the network evolution and homophily with respect to alcohol consumption is (partially) explained by selection mechanisms whereby pupils choose their friends according to their drinking behaviour.

We now consider the behaviour dynamics. The behaviour rate parameters suggest that on average a pupil got 1 and 2 opportunities to change his drinking behaviour, respectively in the first and second observation period. Thus, the mean number of opportunities for a change of the behaviour is much lower than that of a network change. The linear and quadratic shape parameters of the behavioural evaluation function indicate that the distribution of the drinking behaviour of the pupils tends to concentrate around the middle category (once a month) with an upward trend. The significant positive parameter of the average alter effect suggests that there is a tendency of the pupils to adjust their behaviour to the average behaviour of the other pupils to whom they are tied. This indicates that homophily with respect to the drinking behaviour is also due to influence mechanisms whereby pupils tend to adjust their drinking behaviour to that of their friends.

6. Conclusions

In this paper, we developed the generalized method of moments (GMoM) estimation method for stochastic actor-oriented models (SAOMs) for the co-evolution of networks and behaviour. The GMoM is a generalization of the MoM allowing to use more statistics than parameters, and thus including more information in the estimation process. The GMoM estimation was proposed for the SAOM for networks-only dynamics by Amati et al. (2015). As an extension, this paper focused on the statistics depending on both the network and the behaviour, because the efficiency gain expected from the GMoM may be especially pronounced for this case. These statistics are important because they are used to determine the role of selection and influence mechanisms when network-behaviour panel data are analysed.

The statistics that were proposed in Snijders et al. (2007) for modelling selection and influence processes are characterized by a time lag between the dependent variables (network for selection and behaviour for influence mechanisms) and the explanatory variables (behaviour for selection and network for influence mechanisms). Such cross-lagged statistics are less sensitive for the occurrence of concurrent changes of network and behaviour made by the same actor between two consecutive observations. Since the complete series of changes between observation points is not known, those concurrent changes might be the result of either an influence or a selection process and provide information that is supplementary to the information in the cross-lagged statistics. To account for this supplementary information, we defined a new set of contemporaneous statistics considering both dependent variables at the same time point.

The new statistics lead to the definition of additional moment conditions so that the system of moment equations is over-identified and the application of the GMoM is required to compute the estimates of the parameter of the SAOM. The GMoM estimator results from the minimization of a quadratic function based on the difference between the expected values of the statistics and their sample counterparts. Those equations are combined using a matrix of weights which accounts for different variances of the statistics and their correlations, so that the unique and

shared information provided by the statistics are combined in an optimal way. This is particularly relevant in the context of network models, where high correlations between statistics are usual.

The GMoM estimates for the parameter of SAOMs cannot be analytically computed as the expected values of the statistics cannot be expressed in a closed form. Thus, the solution is calculated using an adaptation of the stochastic algorithm proposed by Snijders (2001) for the MoM and later extended to the GMoM in Amati et al. (2015). Here, we described a more stable version of the algorithm obtained by considering the full matrix of weights, instead of a block matrix accounting for the covariances between block of parameters determined by the network and the behavioural rate and evaluation functions. As noted in Amati et al. (2015), the algorithm for the GMoM is slower than the one for the MoM, for two main reasons: (i) a higher number of simulations during both phase 1 and phase 3 is necessary to properly estimate the matrix of weights B and the variance covariance matrix of the estimator $\Sigma_{\hat{\theta}_{\text{GMoM}}}$; and (ii) a higher number of sub-phases in phase 2 is usually required to reach proper convergence.

We presented a small simulation study to investigate the finite-sample properties of the GMoM and the relative efficiency of the GMoM and the MoM estimators. Although limited to a small case study, the results show that the GMoM outperforms the MoM and the efficiency of the GMoM increases with the number of concurrent changes between observations. The greater efficiency of the GMoM estimators derives from both lower bias and smaller variances of the estimates. The efficiency gain is especially pronounced for the parameters representing effects of the network on the behaviour (“influence”), and of the behaviour on the network (“selection”). We then demonstrated the use of the GMoM by analysing the co-evolution of a friendship network and drinking behaviour in a cohort of pupils. The results of the study were coherent with those of the simulations in that the parameter estimates differed between the GMoM and the MoM but not dramatically so, and almost all standard errors were lower for the GMoM than for the MoM.

In this paper, we have not addressed the question of which additional statistics should be used. In general, it can be argued that the higher the number of moment conditions, the more (asymptotically) efficient the resulting GMoM estimator will be. However, several studies (Hansen, 1982; Breusch et al., 1999; Newey and Windmeijer, 2009; Hall, 2005 and the references therein) have proven that additional moment conditions improve the asymptotic efficiency of the GMoM estimator only when they contribute extra information to that already provided by the other moment conditions. Therefore, the choice of the contemporaneous statistics that should be used for estimation of parameters in SAOMs is a noteworthy point, and further investigation is needed. The development of guidelines for the selection of the additional statistics, and a more extensive comparison of the GMoM with the MoM and MLE in SAOMs are objects of future work.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix

We describe in details the algorithm used to approximate the value of the GMoM and outline the additional steps that need to be performed compared to the “regular” algorithm for approximating the value of the MoM estimator (Snijders, 2001). In the following, we refer to the Robbins–Morro procedure for the MoM as the “MoM algorithm,” and to its modified version for the GMoM as the “GMoM algorithm.” To keep the notation transparent, we stick to the notation used in Section 3.2. The GMoM estimate for θ is the value $\hat{\theta}$ such that

$$B E_{\theta} [s^*(X, Z) - s^*(x, z)] = 0,$$

with $B = \frac{\partial}{\partial \theta} E_{\theta} [s^*(X, Z) - s^*(x, z)]' W = \Gamma W$. The additional steps in the GMoM algorithm compared to the MoM algorithm are related to the approximation of the matrix B .

Let θ_0 be an initial value of the parameter θ . For the network and behavioural rate parameters, the default initial values are computed as a function of the number of network changes and the number, direction and size of the behavioural changes. For the evaluation functions, the default initial values are obtained by setting all the parameters, but the outdegree and the linear shape parameters, equal to 0. For the outdegree parameter, the initial value is computed as a function of the log-odds of the probability of a tie being present given the observed data (Snijders, 2005). For the linear shape parameter, the initial value is computed as a function of the absolute mean and the variance of the behaviour over the observation period. For relatively simple models, the default initial values usually work fine. For more complex models, the GMoM algorithm might converge slowly and better initial values can be obtained by fitting the regular MoM to the data and set θ_0 to the MoM estimate.

The estimation algorithm consists of three phases.

1. Phase 1

The first phase is a preliminary phase used to approximate the quantities involved in the Robbins–Monro step, namely the matrix Γ of first-order derivatives of the statistics, the matrix of weights W and the matrix D of the first-order derivatives of the function $B E_{\theta} [s^*(X, Z) - s^*(x, z)]$. We denote these approximations by Γ_0 , W_0 and D_0 , respectively.

- 1.1. Given the initial value θ_0 , simulate n_1 network–behaviour co-evolution trajectories.
- 1.2. Approximate the generic element of Γ_0 and defined as

$$\gamma_{pk} = \frac{\partial}{\partial \theta_p} E_{\theta} [s_k^*(X, Z) - s_k^*(x, z)]$$

by averaging the product of the simulated statistics and the score functions [score function method, Schweinberger and Snijders (2007)] or by averaging difference quotients using random numbers [finite difference method, Snijders (2001)]. The former method is preferred since it is more computationally efficient for complex models and leads to the following approximation

$$\hat{\gamma}_{pk} = \frac{1}{n_1} \sum_{h=1}^{n_1} U_{ph} [s_k^*(x^{(h)}, z^{(h)}) - s_k^*(x, z)],$$

where $s_k^*(x^{(h)}, z^{(h)})$ and $U_{ph} [p(x^{(h)}, z^{(h)})] = \frac{\partial}{\partial \theta_p} \log [p(x^{(h)}, z^{(h)})]$ are the values of the statistics and the p th component of the score function vector computed for the h th simulation.

- 1.3. Compute the sample covariance matrix of the simulated statistics, denoted by $\hat{\Sigma}$. The generic cell of this matrix is computed in the following way

$$\hat{\Sigma}_{ku} = \frac{1}{n_1} \sum_{h=1}^{n_1} [s_k^*(x^{(h)}, z^{(h)}) - \bar{s}_k^*(x, z)] [s_u^*(x^{(h)}, z^{(h)}) - \bar{s}_u^*(x, z)],$$

with $\bar{s}_k^*(x, z)$ the sample average of the values of the simulated statistic $s_k(x, z)$, i.e.

$$\bar{s}_k^*(x, z) = \frac{1}{n_1} \sum_{h=1}^{n_1} s_k^*(x^{(h)}, z^{(h)}).$$

- 1.4. Approximate the weight matrix W by computing $W_0 = \hat{\Sigma}^{-1}$.
- 1.5. Compute $B_0 = \Gamma_0 W_0$.
- 1.6. Divide all elements of B_0 by their row sums.
- 1.7. Approximate the matrix

$$D = \frac{\partial}{\partial \theta} B E_{\theta} [s^*(X, Z) - s^*(x, z)]$$

by using $D_0 = B_0 \Gamma_0'$.

- 1.8. Determine the first rough estimate of θ using one Newton–Raphson step

$$\hat{\theta} = \theta_0 - \alpha D_0^{-1} B_0 [\bar{s}^*(x, z) - s^*(x, z)],$$

with $\bar{s}^*(x, z)$ the vector of the averages of the simulated statistics and α a number between 0 and 1.

In comparison with Phase 1 of the MoM algorithm, the GMoM algorithm requires the additional steps 1.3–1.6 and the computation of the larger matrix Γ of first-order derivatives of the statistics. Thus, a higher number of simulations $n_1 = 100 + (7 \times q)$ is required.

In a previous version of the GMoM algorithm (Amati et al., 2015), the matrix W was computed as a block diagonal matrix, where the blocks correspond to the parameter of the rate functions and those of the evaluation functions. Here, we use a full matrix of weights W since we noticed that this choice gives more accurate estimates for the rate parameters and more stability to the algorithm.

2. Phase 2

The second phase carries out the estimation and follows the MoM algorithm except for the inclusion of the matrix B in the Robbins–Monro step. During the second phase, the matrices D and B are kept fixed at D_0 and B_0 .

The second phase of the algorithm is divided into L sub-phases characterized by a decreasing value of α_r and the use of the diagonal matrix obtained from the matrix D (Ruppert, 1988; Polyak, 1990; Pflug, 1990).

Each of the L (advice: $L = 5$ sub-phases) sub-phases comprises the following steps:

- 2.1. Set $\hat{\theta}_1 = \hat{\theta}$, α (advice: $\alpha = 0.2$ for the first sub-phase), and $n_2 = n_{2\ell}$, with ℓ being the number of the current sub-phase.
- 2.2. For $r = 1, \dots, n_{2\ell} - 1$, simulate one co-evolution trajectory for the current value of the parameter $\hat{\theta}_r$ and compute the corresponding value of the statistics $s^*(x^{(r)}, z^{(r)})$. Update θ by the Robbins–Monro step

$$\hat{\theta}_{r+1} = \hat{\theta}_r - \alpha D_0^{-1} B_0 [s^*(x^{(r)}, z^{(r)}) - s^*(x, z)]$$

2.3. Update the value of θ by

$$\hat{\theta} = \frac{1}{n_{2\ell}} \sum_{r=1}^{n_{2\ell}} \hat{\theta}_r$$

The average of the values $\hat{\theta}_r$ during the last sub-phase is the final estimate for θ . Thus, we define the GMoM estimate of $\hat{\theta}$ as the value $\hat{\theta}_1$ derived from the last sub-phase of Phase 2.

2.4 Set $\alpha = \alpha/2$, $n_{2\ell} = (7 + p) \times (2.52)^\ell + 1$.

3. Phase 3

This phase computes the standard errors of the estimates and evaluate the convergence of the algorithm.

- 3.1. Simulate n_3 co-evolution trajectories using the GMoM estimate $\hat{\theta}$.
- 3.2. Update the approximation of the matrices B , Γ and Σ as in steps (1.2)–(1.6) of Phase 1. We denote these estimates by \hat{B} , $\hat{\Gamma}$ and $\hat{\Sigma}$.
- 3.3. Compute the covariance matrix of the GMoM estimator as

$$\Sigma_{\hat{\theta}_{\text{GMoM}}} = (\hat{B} \hat{\Gamma})^{-1} (\hat{B} \hat{\Sigma} \hat{B}') ((\hat{B} \hat{\Gamma})^{-1})'$$

3.4. Calculate the t -ratios for convergence

$$t\text{-ratio}_p = \frac{\hat{B}(\bar{s}_p^*(x, z) - s_p^*(x, z))}{\sqrt{(\hat{B} \hat{\Sigma} \hat{B}')_{pp}}}$$

As a rule of thumb, when these t -ratios are smaller than 0.1, then convergence is satisfactory.

Compared to the MoM algorithm, the GMoM algorithm requires the additional estimation of B , a larger matrix of derivatives Γ , and a larger covariance matrix Σ (step 3.2). Thus, a higher number of simulations is required. For the simulation study in Section 4 and the example in Section 5, we used $n_3 = 10,000$.

References

- Amati, V., Schöenberger, F., & Snijders, T. A. (2015). Estimation of stochastic actor-oriented models for the evolution of networks by generalized method of moments. *Journal de la Société Française de Statistique*, 156(3), 140–165.
- Block, P. (2015). Reciprocity, transitivity, and the mysterious three-cycle. *Social Networks*, 40, 163–173.
- Bollen, K. A., Kolenikov, S., & Barlady, S. (2014). Model-implied instrumental variable—generalized method of moments (MIV-GMM) estimators for latent variable models. *Psychometrika*, 79(1), 20–50.
- Breusch, T., Qian, H., Schmidt, P., & Wyhowski, D. (1999). Redundancy of moment conditions. *Journal of Econometrics*, 91(1), 89–111.
- Burgueño, J. F., Ronald Gallant, A., & Souza, G. (1982). On unification of the asymptotic theory of nonlinear econometric models. *Econometric Reviews*, 1(2), 151–190.
- Burk, W. J., Kerr, M., & Statin, H. (2008). The co-evolution of early adolescent friendship networks, school involvement, and delinquent behaviors. *Revue française de sociologie*, 49(3), 499–522.
- Ebbers, J. J., & Wijnberg, N. M. (2010). Disentangling the effects of reputation and network position on the evolution of alliance networks. *Strategic Organization*, 8(3), 235–275.
- Gallant, A. R., Hsieh, D., & Tauchen, G. (1997). Estimation of stochastic volatility models with diagnostics. *Journal of econometrics*, 81(1), 159–192.
- Hall, A. R. (2005). *Generalized method of moments*. Oxford: Oxford University Press.
- Hansen, L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50, 1029–1054.

- Hansen, L. P., & Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica*, 50(5), 1269–1286.
- Haynie, D. L., Doogan, N. J., & Soltter, B. (2014). Gender, friendship networks, and delinquency: A dynamic network approach. *Criminology*, 52(4), 688–722.
- Holland, P. W., & Leinhardt, S. (1977). A dynamic model for social networks. *Journal of Mathematical Sociology*, 5(1), 5–20.
- Hunter, D. R. (2007). Curved exponential family models for social networks. *Social Networks*, 29, 216–230.
- Kim, J.-S., & Fries, E. W. (2007). Multilevel modeling with correlated effects. *Psychometrika*, 72(4), 505–533.
- Koskinen, J. H., & Snijders, T. A. B. (2007). Bayesian inference for dynamic social network data. *Journal of Statistical Planning and Inference*, 13, 3930–3938.
- Luce, R., & Suppes, P. (1965). Preference, utility, and subjective probability. *Handbook of Mathematical Psychology*, 3, 249–410.
- Mátys, L. (1999). *Generalized method of moments estimation*. Cambridge: Cambridge University Press.
- McFadden, D. (1973). *Conditional logit analysis of qualitative choice behavior*. Oakland: Institute of Urban and Regional Development, University of California.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415–444.
- Meyer, C. D. (2000). *Matrix analysis and applied linear algebra*. Philadelphia: SIAM.
- Michell, L., & West, P. (1996). Peer pressure to smoke: The meaning depends on the method. *Health Education Research*, 11(1), 39–49.
- Newey, W., & Windmeijer, F. (2009). Generalized method of moments with many weak moment conditions. *Econometrica*, 77(3), 687–719.
- Neyman, J., & Pearson, E. S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference: Part II. *Biometrika*, 20, 263–294.
- Nieznik, N. M. D., & Snijders, T. A. B. (2017). Co-evolution of social networks and continuous actor attributes. *The Annals of Applied Statistics*, 11(4), 1948–1973.
- Nieznik, N. M. D., Snijders, T. A. B., & van Duijn, M. A. J. (2019). No longer discrete: Modeling the dynamics of social networks and continuous behavior. *Sociological Methodology*. <https://doi.org/10.1177/0081175019842263>.
- Norris, J. R. (1997). *Markov chains*. Cambridge: Cambridge University Press.
- Pearson, K. (1900). On the criterion that a given system of deviations 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. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302), 157–175.
- Pflug, G. C. (1990). Non-asymptotic confidence bounds for stochastic approximation algorithms with constant step size. *Manuscripte für Mathematik*, 110(3–4), 297–314.
- Polyak, B. T. (1990). A new method of stochastic approximation type. *Automation and Remote Control*, 51, 937–946.
- Ripley, R. M., Snijders, T. A. B., Boda, Z., András, V., & Paulina, P. (2019). *Manual for Stiena*. Groningen: ICS, Department of Sociology, University of Groningen.
- Robbins, H., & Monro, S. (1951). A stochastic approximation method. *The Annals of Mathematical Statistics*, 22, 400–407.
- Ruppert, D. (1988). *Efficient estimations from a slowly convergent Robbins–Monro process*. Technical report, Cornell University Operations Research and Industrial Engineering.
- Schulte, M., Cohen, N. A., & Klein, K. J. (2012). The coevolution of network ties and perceptions of team psychological safety. *Organization Science*, 23(2), 564–581.
- Schweinberger, M., & Snijders, T. A. B. (2007). Markov models for digraph panel data: Monte Carlo-based derivative estimation. *Computational Statistics & Data Analysis*, 51(9), 4465–4483.
- Snijders, T. A. B. (1996). Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology*, 21(1–2), 149–172.
- Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. *Sociological Methodology*, 31(1), 361–395.
- Snijders, T. A. B. (2005). Models for longitudinal network data. In P. J. C. Conte, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 215–247). Cambridge: Cambridge University Press.
- Snijders, T. A. B. (2017a). Stochastic actor-oriented models for network dynamics. *Annual Review of Statistics and Its Application*, 4, 343–363.
- Snijders, T. A. B. (2017b). *Stiena algorithms*. Technical report, University of Groningen. University of Oxford. http://www.stats.ox.ac.uk/~snijders/stiena/Stiena_algorithms.pdf.
- Snijders, T. A. B., Koskinen, J., & Schweinberger, M. (2010a). Maximum likelihood estimation for social network dynamics. *The Annals of Applied Statistics*, 4(2), 567–588.
- Snijders, T. A. B., & Lomi, A. (2019). Beyond homophily: Incorporating actor variables in statistical network models. *Network Science*, 7(1), 1–19.
- Snijders, T. A. B., Steglich, C. E. G., & Schweinberger, M. (2007). Modeling the co-evolution of networks and behavior. In K. van Montfort, H. Oud, & A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences* (pp. 41–71). Mahwah, NJ: Lawrence Erlbaum.
- Snijders, T. A. B., Van de Bunt, G. G., & Steglich, C. E. G. (2010b). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1), 44–60.
- Snijders, T. A. B., & van Duijn, M. A. J. (1997). Simulation for statistical inference in dynamic network models. In R. Conte, R. Hegselmann, & P. Terna (Eds.), *Simulating social phenomena* (pp. 493–512). Berlin: Springer.
- Steglich, C. E. G., Snijders, T. A. B., & Pearson, M. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological Methodology*, 40(1), 329–393.

Strang, G. (1976). *Linear algebra and its applications*. New York: Academic Press.
Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge: Cambridge University Press.

Manuscript Received: 29 MAR 2018

Final Version Received: 24 JUN 2019

Published Online Date: 15 JUL 2019