

## CO-EVOLUTION OF SOCIAL NETWORKS AND CONTINUOUS ACTOR ATTRIBUTES

BY NYNKE M. D. NIEZINK\* AND TOM A. B. SNIJDERS

*University of Groningen*

*Abstract.* Social networks and the attributes of the actors in these networks are not static; they may develop interdependently over time. The stochastic actor-oriented model allows for statistical inference on the mechanisms driving this co-evolution process. In earlier versions of this model, dynamic actor attributes are assumed to be measured on an ordinal categorical scale. We present an extension of the stochastic actor-oriented model that does away with this restriction, using a stochastic differential equation to model the evolution of continuous actor attributes. We estimate the parameters by a procedure based on the method of moments. The proposed method is applied to study the dynamics of a friendship network among the students at an Australian high school. In particular, we model the relationship between friendship and obesity, focusing on body mass index as a continuous co-evolving attribute.

**1. Introduction.** Social actors on all levels, whether they are individuals, companies or countries, are embedded in social structures. Networks are a useful tool to represent these structures. They are defined by a particular relation, e.g., friendship, collaboration or trade, on a set of actors who are both shaping and shaped by the network they are embedded in. For example, people may change their attitudes and behaviors based on those of their friends (social influence). Simultaneously they may select their friends based on these same attitudes and behaviors (social selection). Christakis and Fowler (2007), based on the analysis of a social network among 12,067 people, claimed that obesity spreads through social ties. Their study and the many reactions it received illustrate the scientific and societal interest in social influence processes, and the intricate nature of the relation of these processes with social selection. Empirical (Cohen-Cole and Fletcher, 2008) and theoretical (Shalizi and Thomas, 2011) rebuttals emphasized that influence and selection are generally confounded. Causal claims about these processes based on observational data are to be made with care. Distin-

---

\*Supported by the Netherlands Organization for Scientific Research (NWO grant 406-12-165).

*Keywords and phrases:* social networks, longitudinal data, Markov model, stochastic differential equations, method of moments

guishing selection and influence requires strong assumptions on the absence of latent causal factors and on the parametrization of the underlying social process (Shalizi and Thomas, 2011). In this article, we consider statistical models for network-attribute co-evolution processes that aim to deal with this complexity (Steglich, Snijders and Pearson, 2010). In particular, we develop a model for the case that the actors' attributes are expressed on a continuous scale.

Co-evolution of networks and actor attributes is a continuous-time process. However, available longitudinal data are often only the discrete-time manifestations of this process at a few time points. To study the dynamics of networks based on such data, Holland and Leinhardt (1977) proposed the use of a continuous-time Markov chain model, with all possible networks on a specific actor set as its state space. They illustrated this approach through dyad-based models, which assume that the relations between pairs of actors (dyads) in a network evolve independently. The popularity model by Wasserman (1980) extends the approach. However, neither of these models takes into account the many complex dependency structures that characterize social networks (e.g., triadic structures such as 'a friend of a friend is my friend'). The stochastic actor-oriented model is a model in the tradition of Holland and Leinhardt (1977) that can take into account these complex structural mechanisms (Snijders, 2001; Snijders, Koskinen and Schweinberger, 2010). This model has been extended for the statistical analysis of the co-evolution of networks and actor attributes (Snijders, Steglich and Schweinberger, 2007; Steglich, Snijders and Pearson, 2010).

The stochastic actor-oriented model is used, for example, to study the spread of behaviors and attitudes through social networks and to explain why related actors often behave and think similarly. The latter phenomenon, called network autocorrelation, can be caused by influence processes (actors becoming more similar to those to whom they are related) or by homophilous selection (actors becoming related to similar others). The stochastic actor-oriented model facilitates the disentanglement of these processes. The model is widely applied, for example to study the role of peers in weapon-carrying and delinquency among adolescents (Dijkstra et al., 2010; Weerman, 2011) or the job satisfaction and interpersonal trust relationships in organizations (Agneessens and Wittek, 2008). De la Haye et al. (2011) applied the model to explain the existence of clusters of obese adolescents in friendship networks.

In these applications the model by Snijders, Steglich and Schweinberger (2007) is used, which assumes the attributes of network actors to be measured on an ordinal categorical scale, and models the evolution processes of the network relations and the actor attributes jointly as a continuous-time

Markov chain. Although ordinal discrete variables occur in many applications, this assumption has been experienced as restrictive in several others. Researchers needed to discretize their continuous actor variables before analyzing them in the stochastic actor-oriented modeling framework. For example, Dijkstra et al. (2012) transformed scales for self-reported aggression and victimization to a 4-point scale and De la Haye et al. (2011) expressed body mass index (weight divided by height squared) on a 16-point scale. When there are no substantive grounds for discretization, as in these examples, selecting a particular discretization is hard. Moreover, discretization may lead to loss of information and substantive conclusions may differ between discretizations.

This article presents an extension of the stochastic actor-oriented model for the study of the co-evolution of networks and actor attributes that are measured on a continuous scale. We model the evolution of the continuous attributes by stochastic differential equations. This is a standard approach in econometrics (Bergstrom, 1984, 1988) and finance (Black and Scholes, 1973; Merton, 1990), but has also been proposed for panel data in the social sciences generally (Hamerle, Singer and Nagl, 1993; Oud and Jansen, 2000). Although stochastic differential equation models have clear advantages over discrete-time models (Voelkle et al., 2012), social science applications other than financial ones are rare. Moreover, in almost all applications observation units are assumed to be independent: the idea that the units might be interconnected has received little to no attention. An exception is the work by Oud et al. (2012), who account for the spatial proximity of observation units. In their model interconnection is induced by geographic location and treated as something that needs to be controlled for; it is assumed to be static and is part of the model's error process. In this article, interconnection is assumed to be a dynamic phenomenon and is object of study itself.

The article is organized as follows. In Section 2, we propose the model for continuous attribute evolution. Section 3 first discusses the stochastic actor-oriented model for network evolution. Then the two models are integrated and a simulation algorithm for the co-evolution process is outlined. It is by combining these two model components that selection and influence processes can be studied. Section 4 describes a method of moments procedure for parameter estimation. The performance of this method is evaluated in the simulation study in Section 6. The setup of the simulation study is inspired by the application of the method in Section 5, in which we study the effects of peer influence and social selection on body mass index in adolescent friendship networks. In the dataset we analyze, we do not find support for either of these effects. Section 7 concludes with a discussion.

1.1. *Notation and data structure.* A social network on a given set of actors  $\mathcal{I} = \{1, \dots, n\}$  can be modelled as a directed graph, in which the nodes correspond to the actors and the set of directed ties to a specific social relation between them. The directed graph can be represented by an adjacency matrix  $x = (x_{ij}) \in \{0, 1\}^{n \times n}$ , where  $x_{ij} = 1$  and  $x_{ij} = 0$  respectively indicate the presence and absence of a tie from actor  $i$  to actor  $j$ . Ties are assumed to be directed, so  $x_{ij}$  and  $x_{ji}$  are not necessarily equal, and to be nonreflexive, so  $x_{ii} = 0$  for  $i \in \mathcal{I}$ .

This article considers data structures consisting of repeated observations of relations on a fixed set of actors and the attributes of these actors. These attributes are assumed to be continuous and measured on an interval scale. We consider the same  $p$  attributes for each actor. The attribute values of all actors are summarized in matrix  $z = (z_{ih}) \in \mathbb{R}^{n \times p}$ , where  $z_{ih}$  denotes the value of actor  $i$  on attribute  $h$ . Vector  $z_i = (z_{i1}, \dots, z_{ip})$  contains all attribute values of actor  $i$  (the  $i$ th row of  $z$ ).

The network and actor attributes are observed at a finite number of time points  $t_1 < \dots < t_M$ , resulting in observations  $x(t_m)$  and  $z(t_m)$ , where  $m = 1, \dots, M$  and  $M \geq 2$ . These observations are assumed to be realizations of stochastic networks  $X(t_m)$  and attributes  $Z(t_m)$ , embedded in a continuous-time stochastic process  $(X(t), Z(t))$ , where  $t_1 \leq t \leq t_M$ . This process may also depend on non-stochastic individual covariates  $v = (v_1, \dots, v_n)$  and dyadic covariates  $w = (w_{ij}) \in \mathbb{R}^{n \times n}$ . For notational simplicity, the covariates will mostly be treated implicitly. The entire process, including the covariates, is denoted by  $Y(t)$ .

**2. Continuous attribute evolution.** We model the evolution of the  $p$  attributes  $Z_i(t)$  of actor  $i$  by a linear stochastic differential equation (e.g., Steele, 2001). For the period between two observation moments  $t_m$  and  $t_{m+1}$ , the model is given by

$$(1) \quad dZ_i(t) = \tau_m [A Z_i(t) + B u_i(t)] dt + \sqrt{\tau_m} G dW_i(t),$$

where we condition on the initial observation  $Z_i(t_m) = z_i(t_m)$  of that period. Note that the only period-specific parameter in this model is  $\tau_m$ . The meaning of this parameter is elaborated in Section 2.1. The linearity of the differential equation makes for easy simulation of the attribute evolution process: Section 2.2 describes how the corresponding transition density can be expressed analytically. Here, we will first take a closer look at model (1) and at how an actor's embeddedness in a social network may affect his attribute evolution.

The matrix  $A \in \mathbb{R}^{p \times p}$  in the stochastic differential equation is called the drift and specifies the feedback relationships among the  $p$  attributes. The

elements of input vector  $u_i(t) \in \mathbb{R}^r$  are called effects. Effects are functions of the state  $Y(t)$  of the co-evolution process. They can, for example, depend on time-constant actor covariates or on network-related characteristics of actor  $i$ . Network-related effects lead to a dependence of the attribute evolution on the network. They are what turns the attribute evolution into a network-attribute co-evolution process. Some examples of network-related effects are given in Table 1. All these effects specify a differential drift based on a network-related characteristic of actor  $i$ . The isolate effect, for example, reflects the effect of having no incoming relations (i.e., being unpopular). The average alter effect can be used to model social influence. This effect represents the dependence of the attributes  $Z_i(t)$  of actor  $i$  on the attributes of the actors to whom  $i$  has a relation at time  $t$ . For discrete attribute evolution in the context of stochastic actor-oriented models, many effects have already been defined (Ripley et al., 2016). Most of these can be generalized straightforwardly for continuous attributes.

Matrix  $B \in \mathbb{R}^{p \times r}$  contains parameters indicating the strength of the effects in  $u_i(t)$  on the attribute dynamics. If unit variable 1 is an element of the input vector, the corresponding parameters in  $B$  serve as the intercept. The  $p$ -dimensional Wiener process  $W_i(t)$  is responsible for the stochastic nature of  $Z_i(t)$ . The Wiener process  $\{W_i(t), t \geq t_1\}$  has the property that  $W_i(t)$  is normally distributed with mean  $0_p$  and covariance  $(t - t_1)I_p$ , where  $0_p$  is the  $p$ -dimensional all-zero vector and  $I_p$  is the  $p \times p$  identity matrix. Matrix  $G \in \mathbb{R}^{p \times p}$  transforms this process into a Wiener process with an arbitrary covariance matrix; it indicates the strength of the error process. The  $W_i(t)$  are independent for the actors  $i \in \mathcal{I}$ .

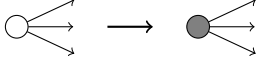
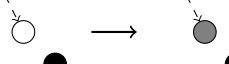
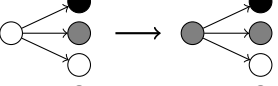
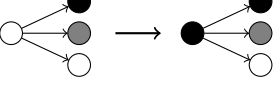
2.1. *Period dependence.* The periods between consecutive observation moments  $t_m$  and  $t_{m+1}$  can have any duration. The period-specific parameter  $\tau_m$  is included in model (1) to take this into account. In our network-attribute co-evolution simulation scheme (Section 3.2), necessary for parameter estimation, we model each period to have unit duration. The discrepancy between ‘model time’ and ‘real time’ is captured by  $\tau_m$ . This can be seen as follows.

Suppose for the moment that  $\tau_m = 1$  in model (1), removing it from the equation. Let  $t = \tau s$ , where  $s$  denotes the ‘model time’, running between 0 and 1, and  $t$  denotes the ‘real time’. This results in the following model in terms of  $s$ :

$$(2) \quad dZ_i(s) = \tau[A Z_i(s) + B u_i(s)] ds + G dW_i(\tau s),$$

where the first factor  $\tau$  stems from  $dt/ds = \tau$ . The effect of time scaling in the stochastic part of the differential equation differs from that in

TABLE 1  
*Selection of effects for modeling attribute evolution.*

Effect name	Effect formula <sup>a</sup>	Effective changes <sup>b</sup>
Outdegree effect	$\sum_j x_{ij} (= x_{i+})$	
Isolate effect	$1 - \max_j(x_{ji})$	
Average alter effect of the $k$ th attribute (defined as 0 in case $x_{i+} = 0$ )	$\sum_j x_{ij}(z_{jk} - \bar{z}_k)/x_{i+}$	
Maximum alter effect of the $k$ th attribute	$\max_j(x_{ij}z_{jk})$	

<sup>a</sup>Time-dependence is omitted for brevity.  $\bar{z}_k$  denotes the observed mean of attribute  $k$ .

<sup>b</sup>Darker colors represent higher values of the attribute. Dotted arrows represent absent relationships. Illustrations are not exhaustive.

the deterministic part. Wiener processes have the following scaling property: given a standard Wiener process  $\{W(t), t \geq 0\}$ , for each  $\tau > 0$ ,  $\{1/\sqrt{\tau}W(\tau t), t \geq 0\}$  is also a standard Wiener process (e.g., Steele, 2001, p. 40). Consequently, changing the time scale by  $t = \tau s$  transforms the standard Wiener process as  $W(t) = W(\tau s) = \sqrt{\tau}W_\tau(s)$ , where  $W_\tau(s)$  is again a standard Wiener process. This explains the way  $\tau_m$  appears in model (1). The parameter absorbs the consequences of time scaling and allows us to assume that in ‘model time’ each period has unit duration.

2.2. *Exact discrete model.* Stochastic differential equation (1) is a convenient way to express the integral equation

$$(3) \quad Z_i(t) - z_i(t_m) = \int_{t_m}^t A Z_i(s) + B u_i(s) ds + \int_{t_m}^t G dW_i(s),$$

in which the second integral is a stochastic integral in the sense of Itô (e.g., Steele, 2001). For many stochastic differential equations there is no analytic expression for their transition density (i.e., how an observation of the modeled variables at a certain moment reflects the accumulation of their dynamics since an earlier point in time). We will show that for our purpose such an expression does exist.

Let ‘vec’ denote the operation of stacking all rows of a matrix into one column vector, ‘ivec’ the inverse of this operation and  $\otimes$  the Kronecker

product. The solution to equation (3) is given by

$$(4) \quad Z_i(t) = e^{A(t-t_m)} z_i(t_m) + \int_{t_m}^t e^{A(t-s)} B u_i(s) ds + \int_{t_m}^t e^{A(t-s)} G dW_i(s)$$

(Arnold, 1974, p.129-130), where the last term is (multivariate) normally distributed with mean  $0_p$  and covariance

$$(5) \quad \text{ivec}[(A \otimes I_p + I_p \otimes A)^{-1} (e^{A(t-t_m)} \otimes e^{A(t-t_m)} - I_p \otimes I_p) \text{vec}(GG^\top)].$$

This is true under the assumption that  $A$  has nonzero eigenvalues and nonzero sums of eigenvalue pairs (e.g., Oud and Jansen, 2000).

Note that equation (4) still contains an integral that depends on input vector  $u_i(s)$ . If this vector includes network-related effects, its value is highly variable, given that the network is dynamic. However, if we consider a small time interval  $[t, t + \Delta t)$  in which the network is constant and assume  $u_i(s)$  to be constant on this interval (see also Section 3.2), observations at time points  $t$  and  $t + \Delta t$  exactly satisfy a system of stochastic difference equations. This system is referred to as the exact discrete model (Bergstrom, 1984; Singer, 1996; Oud and Jansen, 2000). Let  $z_{i,t}$  denote the value of the attribute variables and  $u_{i,t}$  the input vector values of actor  $i$  at time  $t$ . The exact discrete model for model (1) is given by

$$(6) \quad z_{i,t+\Delta t} = A_{\Delta t} z_{i,t} + B_{\Delta t} u_{i,t} + w_{i,\Delta t},$$

where  $w_{i,\Delta t}$  is (multivariate) normally distributed with mean  $0_p$  and covariance matrix  $Q_{\Delta t}$ . The continuous-time parameters in (1) are linked to the discrete-time parameters in (6) by the identities

$$(7) \quad \begin{aligned} A_{\Delta t} &= e^{\tau_m A \Delta t} \\ B_{\Delta t} &= A^{-1} (A_{\Delta t} - I_p) B \\ Q_{\Delta t} &= \text{ivec}[(A \otimes I_p + I_p \otimes A)^{-1} (A_{\Delta t} \otimes A_{\Delta t} - I_p \otimes I_p) \text{vec}(GG^\top)]. \end{aligned}$$

These follow directly from expressions (4) and (5).

**2.3. Identifiability.** While we can use model (1) to simulate attribute trajectories, the model is not identifiable. It contains several redundant parameters. Since matrix  $G$  only enters the exact discrete model through  $GG^\top$  in  $Q_{\Delta t}$ , we can multiply  $G$  by an orthogonal matrix  $L$  without changing  $Q_{\Delta t}$ , i.e.,  $(GL)(GL)^\top = GLL^\top G^\top = GG^\top$ . Moreover, we can multiply parameters  $\tau_m$  by a constant and divide the entries of  $A$  and  $B$  by the same constant

and those of  $G$  by its square root, without changing the stochastic differential equation. To enforce identifiability, we therefore assume  $G$  to be a lower triangular matrix with positive diagonal entries. In this way  $GG^\top$  is uniquely linked to  $G$  through Cholesky decomposition. We also set the upper left entry of  $G$  equal to 1. As a consequence, the scale of the first attribute variable will in practice greatly affect the  $\tau_m$  values. Note, however, that we could have fixed any other parameter instead.

**3. Co-evolution model.** In this section, we discuss how the evolution of a social network is represented by the stochastic actor-oriented model (Snijders, 2001; Snijders, 2005). The model for network evolution process  $X(t)$  can be decomposed into two stochastic subprocesses. The first process models the speed by which the network changes, i.e., the rate by which each actor in the network gets the opportunity to change one of his outgoing tie variables. The second models the mechanisms that determine which particular tie is changed, when the opportunity arises. Together with the attribute model, the network model forms a continuous-time Markov process  $Y(t) = (X(t), Z(t))$ . Section 3.2 presents a simulation procedure for this network-attribute co-evolution process.

*3.1. Network evolution.* In the stochastic actor-oriented model, network evolution is modeled in a continuous-time Markov chain, defined on the space of all possible network configurations (Snijders, 2001). Changes in the network are modeled as choices made by actors: at random moments actors may choose to create or dissolve one of their outgoing ties. This happens under the constraints that only one change may occur at a time and that actors act conditionally independent of each other at any time  $t$ , given the current state of the process  $Y(t)$ . The latter implies that no enforced connection between two actors' decisions is possible. These assumptions allow for the evolution process to be modeled in terms of smallest possible steps, an approach first proposed by Holland and Leinhardt (1977).

At stochastically determined moments actors receive the opportunity to change one of their outgoing ties. Since the process is assumed to be Markovian, the waiting times between these opportunities are exponentially distributed. In general, the rate parameter for actor  $i$  is given by a so-called rate function  $\lambda_i(y, m)$ , that may depend on the time period  $m$ , given by  $\{t \mid t_m \leq t < t_{m+1}\}$ , and the current state of the process  $Y(t) = y$  (Snijders, 2001). However, here we assume the rate parameters to be equal for all actors:  $\lambda_m$  in period  $m$ . This implies that in period  $m$  the waiting time until the next network change by any actor is exponentially distributed with rate  $n\lambda_m$ . The probability that it is actor  $i$  who will receive the opportunity to



make a change is  $1/n$ . The rate parameters  $\lambda_m$  play the same role as the scale parameters  $\tau_m$  in model (1). They account for heterogeneity in period length and allow us to model each period as having unit duration.

Suppose actor  $i$  has received the opportunity to make a network change and the current state of the network is  $x$ . The actor may choose either to maintain the status quo or to change a tie variable to one of the other actors. The set of network configurations  $\mathcal{A}_i(x)$  to which he may change therefore is given by  $\mathcal{A}_i(x) = \{x\} \cup \mathcal{A}_i^1(x)$ , where

$$(8) \quad \mathcal{A}_i^1(x) = \bigcup_{j:j \neq i} \{\tilde{x} \mid \tilde{x}_{ij} = 1 - x_{ij} \text{ and } \tilde{x}_{hk} = x_{hk} \text{ for } (h, k) \neq (i, j)\}.$$

Other definitions of  $\mathcal{A}_i(x)$  are possible; actors may be obliged to make a change if they receive the opportunity to do so, or ties may not be allowed to dissolve once they are created. The conditional probability that actor  $i$  changes the network  $x$  to  $\tilde{x} \in \mathcal{A}_i(x)$  is given by

$$(9) \quad p_i(\tilde{x} \mid x, z) = \begin{cases} \exp(f_i(\tilde{x}, z)) / \sum_{x' \in \mathcal{A}_i(x)} \exp(f_i(x', z)) & \text{if } \tilde{x} \in \mathcal{A}_i(x), \\ 0 & \text{if } \tilde{x} \notin \mathcal{A}_i(x). \end{cases}$$

This multinomial logit model can be interpreted as representing an actor's utility maximizing behavior (McFadden, 1974).<sup>1</sup> In this case, the utility actor  $i$  attaches to a specific new network configuration  $\tilde{x}$  is the sum of an objective function  $f_i(\tilde{x}, z)$  and a random term with standard Gumbel distribution. Note that expression (9) may also depend on constant actor or dyadic covariates. For notational simplicity, these are not mentioned explicitly. Function  $f_i(x, z)$  is given by a linear combination of effects  $s_{ik}(x, z)$ ,

$$(10) \quad f_i(x, z) = \sum_k \beta_k s_{ik}(x, z).$$

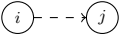
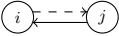
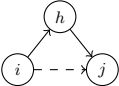
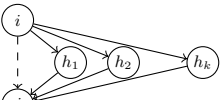
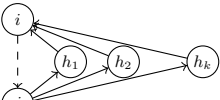
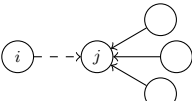
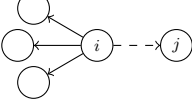

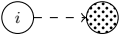

These effects reflect the mechanisms that play a role in relationship formation. They may depend purely on the network structure as experienced by actor  $i$ , as is the case for all but the last three effects in Table 2. The transitivity effect, for example, indicates network closure ('befriending friends of friends'). For large networks, the extension of this effect that uses geometrically weighted triad statistics as in Hunter (2007) leads to better convergence and better fitting models.

---

<sup>1</sup>Although expression (9) may be reminiscent of an exponential random graph model (ERGM), note that it represents the choice probabilities over  $n$  possible network changes instead of a probability distribution over the set of  $2^{n \times (n-1)}$  directed graphs on  $n$  actors. In this respect, our model does not suffer from the computational complexity involved in ERGM estimation.

Effects may also depend on the actors' attributes or on covariates. The covariate-related effects in Table 2 can similarly be defined for the co-evolving attributes. In this way, we can model homophilous selection, i.e., the propensity for actors to initiate relations to similar others, by including an attribute similarity effect. Any similarity measure  $\text{sim}(v_i, v_j)$  can be used in this effect. See Ripley et al. (2016) for an overview of all effects implemented to model network evolution in a stochastic actor-oriented model and for some guidelines on the practice of selecting network effects.

TABLE 2  
Selection of effects for modeling network evolution.

Effect name	Effect formula	Network representation <sup>a</sup>
Outdegree	$\sum_j x_{ij}$	
Reciprocity	$\sum_j x_{ij}x_{ji}$	
Transitivity	$\sum_{j,h} x_{ij}x_{ih}x_{jh}$	
Transitivity (gwersp) <sup>b</sup>	$\sum_k e^\alpha \{1 - (1 - e^{-\alpha})^k\} T_{ik}$	
Cyclicality (gwersp) <sup>b</sup>	$\sum_k e^\alpha \{1 - (1 - e^{-\alpha})^k\} C_{ik}$	
Indegree popularity	$\sum_{j,h} x_{ij}x_{hj}$	
Outdegree activity	$(\sum_j x_{ij})^2$	
Covariate ego <sup>c</sup>	$\sum_j x_{ij}(v_i - \bar{v})$	
Covariate alter <sup>c</sup>	$\sum_j x_{ij}(v_j - \bar{v})$	
Covariate similarity <sup>c</sup>	$\sum_j x_{ij}\text{sim}(v_i, v_j)$	

<sup>a</sup>Dotted arrows represent the effective network change.

<sup>b</sup> $T_{ik}$  denotes the number of actors  $j \in \mathcal{I}$  for whom  $i \rightarrow j$  exists and there are exactly  $k$  actors  $h$  such that  $i \rightarrow h \rightarrow j$  (replace this for  $C_{ik}$  by  $i \leftarrow h \leftarrow j$ ).

<sup>c</sup>The covariate effects can similarly be defined for the dynamic actor attributes.

3.2. *Network-attribute co-evolution scheme.* Suppose that the network-attribute state at a specific time  $t$  is  $y = (x, z)$  and we are modeling the process in period  $m$ . Let  $T(t) = \Delta t$  denote the waiting time until the next network change after time  $t$ , given this state. This waiting time is exponentially distributed with rate  $n\lambda_m$ .

The evolution of the attributes of each of the actors is governed by stochastic differential equation (1). For simulating the model we make the approximation that within the  $\Delta t$  period the input vector  $u_i(t)$  is constant, so the exact discrete model can be used to express the distribution of the actors' attributes at time  $t + \Delta t$  analytically. This approximation is exact if  $u_i(t)$  does not include functions that depend on  $Z_j(t)$  ( $j \neq i$ ), such as the average alter effect defined in Section 2. If  $u_i(t)$  does include such functions, the attribute evolution trajectories of the actors are as related as the actors themselves. In practice, however, the time  $\Delta t \sim \mathcal{O}(1/(n\lambda_m))$  is so short that the effects of the approximation are negligible (see Appendix A).

Under the assumption that  $u_i(t)$  is constant between  $t$  and  $t + \Delta t$ , the attributes  $Z_i(s)$  of different actors  $i \in \mathcal{I}$  evolve independently during the  $\Delta t$  period. The exact discrete model (6) yields

$$(11) \quad (Z_i(t+s) \mid Y(t) = (x, z)) \sim \mathcal{N}(A_s z_{i,t} + B_s u_{i,t}, Q_s),$$

for  $0 < s \leq \Delta t$ , where the matrices  $A_s$ ,  $B_s$  and  $Q_s$  are specified as in (7). After waiting time  $\Delta t$ , a change in the network may occur. The probability that the next network is  $\tilde{x}$  is given by

$$(12) \quad P(X(t+\Delta t) = \tilde{x} \mid T(t) = \Delta t, X(t) = x, Z(t+\Delta t) = z) = \frac{1}{n} \sum_i p_i(\tilde{x} \mid x, z).$$

Algorithm 1 can be used to simulate the stochastic process  $Y(t)$  and is derived directly from the above specification.

The expected number of tie changes in a single period  $m$  is  $n\lambda_m$ , and for each tie change  $n$  options have to be considered and the attributes of  $n$  actors need to be computed. Consequently, the time complexity of simulating a co-evolution process for all  $M - 1$  observed periods is  $\mathcal{O}(n^2 \sum_{m=1}^{M-1} \lambda_m)$ .

**4. Parameter estimation.** Stochastic actor-oriented models are in general too complicated for likelihood functions or estimators to be expressed in a computable form. Nevertheless, they can be used as data simulation models and the expected values of functions of the data can be easily estimated for given parameter values. Therefore, parameters in these models are usually estimated by the method of moments (Snijders, 2001). This

---

**Algorithm 1** Simulating the network-attribute co-evolution in period  $m$ .

---

**Input:**  $x(t_m)$ ,  $z(t_m)$ , covariates and parameter values.**Output:** Simulated network  $x$  and attributes  $z$ .

- 1: Set  $t = 0$ ,  $x = x(t_m)$  and  $z_i = z_i(t_m)$ ,  $u_i = u_i(x(t_m), z(t_m))$  for all  $i \in \mathcal{I}$ .
  - 2: Sample  $\Delta t$  from an exponential distribution with rate  $n\lambda_m$ .
  - 3: **while**  $t + \Delta t < 1$  **do**
  - 4: For all  $i \in \mathcal{I}$ : sample  $c_i$  from  $\mathcal{N}(A_{\Delta t}z_i + B_{\Delta t}u_i, Q_{\Delta t})$  and set  $z_i = c_i$ .
  - 5: Select  $i \in \mathcal{I}$  with probability  $1/n$ .
  - 6: Select  $\tilde{x} \in \mathcal{A}_i(x)$  according to probabilities  $p_i(\tilde{x} | x, z)$ .
  - 7: Set  $t = t + \Delta t$  and  $x = \tilde{x}$ .
  - 8: For all  $i \in \mathcal{I}$ : update  $u_i = u_i(x, z)$ .
  - 9: Sample  $\Delta t$  from an exponential distribution with rate  $n\lambda_m$ .
  - 10: **end while**
  - 11: For all  $i \in \mathcal{I}$ : sample  $c_i$  from  $\mathcal{N}(A_{(1-t)}z_i + B_{(1-t)}u_i, Q_{(1-t)})$  and set  $z_i = c_i$ .
  - 12: Set  $t = 1$ .
- 

method has recently been extended to a generalized method of moments procedure (Amati, Schöenberger and Snijders, 2015). Bayesian (Koskinen and Snijders, 2007) and maximum likelihood (Snijders, Koskinen and Schweinberger, 2010) estimation methods have also been proposed for stochastic actor-oriented models, but are computationally much more intensive than the method of moments procedures.

Here we extend the method of moments procedure described by Snijders (2001) to simultaneously estimate the parameters in the stochastic differential equation model (1) and the network evolution model. Let  $\theta = (\theta_k)$  denote the parameter vector containing all parameters in the model. For each parameter  $\theta_k$ , we specify a statistic whose expected value is sensitive to changes in  $\theta_k$ . The method of moments estimator  $\hat{\theta}$  is given by those parameter values for which the expected values of all selected statistics  $S(Y)$  are equal to the observed values  $S(y)$ ,

$$(13) \quad \mathbb{E}_{\hat{\theta}}\{S(Y)\} = S(y).$$

Equation (13) is referred to as the moment equation. Given the panel data structure and the assumption of a Markov process, we use a conditional method of moments procedure: the statistics are functions of the conditional distribution of  $Y(t_{m+1})$  given  $Y(t_m) = y(t_m)$ , for  $m = 1, \dots, M - 1$ . For each parameter  $\theta_k$  a real-valued function  $S_k(Y(t_m), Y(t_{m+1}))$  is selected, that tends to become larger as  $\theta_k$  increases. The latter is motivated by the fact that the stochastic monotonicity property, stating that for given  $y(t_m)$ ,

$$(14) \quad \frac{\partial}{\partial \theta_k} \mathbb{E}_{\theta}\{S_k(Y(t_m), Y(t_{m+1})) \mid Y(t_m) = y(t_m)\} > 0,$$

ensures good convergence properties for the estimation algorithm.

The network change rate parameter  $\lambda_m$  and the parameter  $\tau_m$  in the stochastic differential equation (1) only influence the stochastic network-attribute evolution process in a specific period, from  $t_m$  to  $t_{m+1}$ . For these parameters, the function  $S_k(Y(t_m), Y(t_{m+1}))$  itself is a suitable statistic for the moment equation:

$$(15) \quad E_{\theta}\{S_k(Y(t_m), Y(t_{m+1})) \mid Y(t_m) = y(t_m)\} = S_k(y(t_m), y(t_{m+1})).$$

Parameters that are assumed to be constant over the entire evolution process are estimated based on statistics of the form

$$(16) \quad S_k^+(Y) = \sum_{m=1}^{M-1} S_k(Y(t_m), Y(t_{m+1}))$$

and for these parameters the moment equation is given by

$$(17) \quad \sum_{m=1}^{M-1} E_{\theta}\{S_k(Y(t_m), Y(t_{m+1})) \mid Y(t_m) = y(t_m)\} = S_k^+(y).$$

It follows from the delta method (e.g., Lehmann, 1999, p. 315) that we can approximate the covariance matrix of  $\hat{\theta}$  by

$$(18) \quad \text{cov}(\hat{\theta}) \approx D_{\theta}^{-1} \text{cov}_{\theta}(S)(D_{\theta}^{-1})^{\top},$$

where  $D_{\theta}$  is the matrix of partial derivatives of the statistics  $S(Y)$  with respect to the parameters  $\theta$  and  $\text{cov}_{\theta}(S)$  is the covariance matrix of  $S(Y)$ . The latter two matrices are approximated based on simulated data (Schweinberger and Snijders, 2007); they are evaluated at the estimate  $\hat{\theta}$  to obtain  $\text{cov}(\hat{\theta})$  (see Appendix B).

The moment equations (15) and (17) cannot be solved analytically, because except for some trivial cases the expected values in these equations cannot be calculated explicitly. Instead, we estimate  $\theta$  using a multivariate Robbins-Monro stochastic approximation algorithm (Robbins and Monro, 1951; Kushner and Yin, 2003). See Snijders (2001) for a full description of the estimation procedure and Ripley et al. (2016) for a discussion of the convergence criteria.

4.1. *Statistics for network evolution parameters.* A natural statistic for estimating the period-dependent rate parameter  $\lambda_m$  is the amount of network change between  $t_m$  and  $t_{m+1}$ ,

$$(19) \quad \sum_{i,j} |X_{ij}(t_{m+1}) - X_{ij}(t_m)|.$$

This statistic satisfies the stochastic monotonicity property. The motivation for the statistics for the parameters  $\beta_k$ , corresponding to the effects  $s_{ik}(Y(t))$  in the objective function (10), is of a heuristic nature (Snijders, 2001). These statistics are of the form (16), where the function  $S_k(Y(t_m), Y(t_{m+1}))$  is

$$(20) \quad \sum_i s_{ik}(X(t_{m+1}), Z(t_m)).$$

Here the combination of  $X(t_{m+1})$  and  $Z(t_m)$  represents the relation of selection that effect  $s_{ik}(Y(t))$  may represent, i.e., how the network relations are affected by (an earlier state of) the actor attributes (Snijders, Steglich and Schweinberger, 2007).

4.2. *Statistics for attribute evolution parameters.* Statistic (19) represents the overall amount of network change within a period  $m$ . Similarly, we define the following statistic for estimating the period-dependent parameters  $\tau_m$  in the stochastic differential equation model (1):

$$(21) \quad \sum_{i,h} [Z_{ih}(t_{m+1}) - Z_{ih}(t_m)]^2.$$

We assume that the parameter matrices  $A$ ,  $B$  and  $G$  are constant over all periods, so their statistics for the moment equations are of the form (16). In the following, we specify the functions  $S_k(Y(t_m), Y(t_{m+1}))$  for the case that the input  $u_i(t)$  is constant. We will use these function also for the general case, because of their intuitive appeal.

Consider model (1) for the first period,  $t_1$  to  $t_2$ , and suppose that the input  $u_i(t)$  is constant over this period and that  $\tau_1 = 1$ . Then the exact discrete model yields

$$(22) \quad Z_i(t_2) = \tilde{A}z_i(t_1) + \tilde{B}u_i(t_1) + w_i,$$

where the  $w_i$  are normally distributed with mean  $0_p$  and covariance  $\tilde{Q}$ , and  $\tilde{A} = A_{t_2-t_1}$ ,  $\tilde{B} = B_{t_2-t_1}$  and  $\tilde{Q} = Q_{t_2-t_1}$  as defined in (7). For an exponential family distribution, such as model (22), maximum likelihood estimation and method of moments estimation are equivalent, when the sufficient statistics for the distribution are used as statistics in the moment equation. The sufficient statistics for model (22) are

$$(23) \quad \sum_i Z_i(t_2)z_i(t_1)^\top, \quad \sum_i Z_i(t_2)u_i(t_1)^\top \quad \text{and} \quad \sum_i Z_i(t_2)Z_i(t_2)^\top.$$

We can use these to estimate parameters  $\tilde{A}$ ,  $\tilde{B}$  and  $\tilde{Q}$ . Under certain conditions, equations (7) uniquely link  $\tilde{A}$ ,  $\tilde{B}$  and  $\tilde{Q}$  to the continuous-time

parameters  $A$ ,  $B$  and  $G$  in model (1), as shown in Lemma 1. As a consequence, expressions (23) can also be used in the estimation of  $A$ ,  $B$  and  $G$ . Conditions 2 and 3 in the lemma have been set earlier in this article.

LEMMA 1. Suppose that 1) matrix  $\tilde{A}$  has no zero or negative real eigenvalues, 2) the eigenvalues of  $A$  are nonzero, and 3)  $G$  is a lower triangular matrix with strictly positive diagonal elements. Then parameters  $A$ ,  $B$  and  $G$  can be uniquely expressed in terms of  $\tilde{A}$ ,  $\tilde{B}$  and  $\tilde{Q}$ .

PROOF. Consider the equations (7) with  $\tau_1 = 1$ . Assume without loss of generality that  $t_2 - t_1 = 1$ . Because of condition 1, the equation  $\tilde{A} = e^A$  is uniquely identified by the principal logarithm  $\ln \tilde{A}$ . If  $\lambda$  is an eigenvalue of  $A$ ,  $e^\lambda$  is an eigenvalue of  $e^A$ . Therefore, condition 2 implies that none of the eigenvalues of  $e^A$  are equal to 1 and none of the eigenvalues of  $e^A - I_p$  are zero, so  $e^A - I_p$  is invertible. Finally, the eigenvalues of  $M_1 \otimes M_2$  are all the products of the pairs of eigenvalues of  $M_1$  and  $M_2$ . Therefore,  $e^A \otimes e^A - I_p \otimes I_p$  is also invertible, and

$$\begin{aligned} A &= \ln \tilde{A} \\ B &= (\tilde{A} - I_p)^{-1} (\ln \tilde{A}) \tilde{B} \\ GG^\top &= \text{ivc} [(\tilde{A} \otimes \tilde{A} - I_p \otimes I_p)^{-1} (\ln \tilde{A} \otimes I_p + I_p \otimes \ln \tilde{A}) \text{vec } \tilde{Q}] \end{aligned}$$

forms a well-defined set of solutions to (7). As a consequence of condition 3, matrix  $G$  can be retrieved from  $GG^\top$  through Cholesky decomposition.  $\square$

Input  $u_i(t)$  in model (1) is usually not constant, as many interesting research questions require the attribute evolution of actor  $i$  to depend on the attributes of actors  $j \neq i$  or on the network. However, as we do not observe the change in  $u_i(t)$  between measurement moments, we select the following functions  $S_k(Y(t_m), Y(t_{m+1}))$  for the statistic (16):

$$(24) \text{ for } a_{hk} \text{ in } A: \quad \sum_i Z_{ih}(t_{m+1}) Z_{ik}(t_m),$$

$$(25) \text{ for } b_{hk} \text{ in } B: \quad \sum_i Z_{ih}(t_{m+1}) u_{ik}(t_m),$$

$$(26) \text{ for } g_{hk} \text{ in } G: \quad \sum_i [Z_{ih}(t_{m+1}) - Z_{ih}(t_m)] [Z_{ik}(t_{m+1}) - Z_{ik}(t_m)].$$

These functions are linear (bijective) transformations of expressions (23) and thus yield maximum likelihood estimates in case  $u_i(t)$  is constant. We only use the functions (26) corresponding to the lower triangular (i.e., nonzero)

and non-fixed entries of  $G$ . The combination  $Z_{ih}(t_{m+1})$  and  $u_{ik}(t_m)$  in function (25) shows the relation that this function is sensitive to: the effect of  $u_{ik}(t)$  on the attributes of actor  $i$ .

**5. Application: Co-evolution of friendship and BMI.** As an illustration of the method proposed above, we re-analyze a data set collected by De la Haye et al. (2011) to study how the evolution of adolescent friendships is affected by their body mass index, and vice versa. Body mass index, or BMI, is defined as the ratio of weight (kg) to squared height (m<sup>2</sup>). Clusters of obese students have repeatedly been observed in friendship networks (Christakis and Fowler, 2007), and using these data we explore possible causes of this phenomenon. On the one hand, adolescents might select their friends based on their BMI. On the other hand, friends might get similar BMI values, e.g., because they serve as each other’s ‘weight referents’ or engage in similar health-related behavior. We will test these competing hypotheses of social selection and social influence using the stochastic actor-oriented model, as was done by De la Haye et al. (2011). However, in this study we analyze BMI as a continuous co-evolving attribute.

Four waves of data were collected among a cohort of students in their first two years at an Australian high school. Students were asked to nominate their friends and to provide information about attributes associated with friendship formation. In addition, their BMI was measured. Here we consider only the data from the first three waves of data collection, as for the last wave only rounded BMI scores were available. See Table 3 for some descriptives. We center BMI scores by gender to account for natural differences between boys and girls. Gender and home group co-membership are included as covariates. Of the 156 participating students, 117, 121 and 123 were present at the first three waves.<sup>2</sup>

We study the data in two models. The first model was specified to closely resemble the model presented by De la Haye et al. (2011). In its objective function (10), modeling the friendship dynamics, we include the effects of outdegree, reciprocity, transitivity, the gender of the friendship nominator (‘ego’), the gender of the friendship nominee (‘alter’) and gender similarity. We control for home group co-membership. We include BMI ego, alter and similarity effects, the latter to test our social selection hypothesis. Finally, as De la Haye et al. (2011), we include the interaction of reciprocity and

---

<sup>2</sup>We impute missing network data by the approach discussed in Ripley et al. (2016). The missing BMI data is imputed stochastically based on available BMI data and gender. The imputed values are only used for simulation purposes. For the calculation of the statistics in the method of moments, any terms in (19)–(21) and (24)–(26) that refer to missing variables are left out.



TABLE 3  
*Descriptive statistics of the friendship network and BMI data.*

	Wave 1	Wave 2	Wave 3
Average degree (number of friendship ties)	7.8 (459)	7.8 (483)	7.8 (487)
Proportion of friendship ties reciprocated	0.50	0.49	0.54
Clustering coefficient	0.37	0.37	0.38
BMI boys – median (MAD)	19.9 (2.5)	20.5 (3.4)	20.3 (3.8)
BMI girls – median (MAD)	18.9 (2.4)	19.0 (2.1)	19.1 (2.5)
<i>Compared to the previous wave:</i>			
Number of stable friendship ties	–	236	260
Number of new friendship ties	–	201	184
Number of dissolved friendship ties	–	180	173
Change in BMI values – median (MAD)	–	0.38 (0.58)	0.12 (0.59)

BMI similarity.

Later in this section, we will see that the first model does not capture the network structure well. The second model controls for more endogeneous network effects, defined in Table 2. We include some interaction effects, which are defined as the product of the summands of two effects summed over all actors (e.g.,  $\sum_j x_{ij}x_{ji}\text{sim}(v_i, v_j)$  for the interaction of reciprocity with covariate similarity). In general, controlling for network evolution mechanisms, whether these are related to covariates or purely structural, is necessary to accurately assess the effects of BMI.

We model the BMI dynamics by a simple stochastic differential equation, including a BMI average alter effect to test our social influence hypothesis:

$$(27) \quad dZ_i(t) = \tau_m [a Z_i(t) + b_0 + b_1 \sum_j X_{ij}(t)(Z_j(t) - \bar{z})/X_{i+}(t)] dt + \sqrt{\tau_m} dW_i(t).$$

In case an actor has no friends to be influenced by ( $X_{i+}(t) = 0$ ), the contribution of the average alter effect is 0. Note that, as we consider friendship networks to be non-reflexive ( $X_{ii}(t) = 0$ ), the value  $Z_i(t)$  only affects its own change through feedback proportional to parameter  $a$ .

Table 4 shows the results of the two models. The substantive conclusions that we can draw from model 1 are very similar to the results of De la Haye et al. (2011). We find that students tend to reciprocate friendships and to befriend the friends of their friends. They prefer friendships with students of their own gender and in their own home group. Female students initiate fewer friendships to male students than vice versa. BMI does not significantly affect the tendency to nominate friends or to be nominated as friend. Unlike De la Haye et al. (2011), we find that the effect of BMI similarity on

friendship formation is not significant, although the sign of the effect is positive, reflecting homophilous choices, as expected. Our hypothesis of social selection based on BMI is thus not supported. However, BMI similarity has a significant effect on the reciprocation of this tie, implying that the more similar in BMI the students are, the less likely the reciprocation of this tie. The non-significant average alter effect indicates that there is no evidence that social influence plays a role in the BMI dynamics. This is contrary to the findings by Christakis and Fowler (2007). For substantive discussion of this result we refer to De la Haye et al. (2011) who, in their original study, also did not find evidence of peer effects on BMI.

TABLE 4

*Stochastic actor-oriented models for friendship and BMI dynamics: model 1 resembles De la Haye et al. (2011), model 2 controls for more endogeneous network effects.*

	Model 1		Model 2	
	estimate	(s.e.)	estimate	(s.e.)
<i>Friendship dynamics</i>				
rate period 1	8.02	(0.76)	8.05	(0.76)
rate period 2	6.55	(0.65)	6.67	(0.58)
outdegree	-3.28*	(0.09)	-3.39*	(0.19)
reciprocity	1.85*	(0.15)	3.59*	(0.24)
transitivity	0.45*	(0.03)		
cyclicality (gwesp)			-0.39*	(0.14)
transitivity (gwesp)			2.68*	(0.15)
transitivity (gwesp) $\times$ reciprocity			-1.65*	(0.23)
indegree popularity			-0.09*	(0.02)
outdegree activity			-0.04*	(0.01)
same home group	0.31*	(0.13)	0.52*	(0.14)
same home group $\times$ reciprocity			-0.73*	(0.25)
female ego	-0.32*	(0.13)	-0.37*	(0.14)
female alter	0.29*	(0.11)	0.24*	(0.12)
same gender	0.77*	(0.10)	0.60*	(0.10)
BMI ego	-0.019	(0.055)	-0.023	(0.046)
BMI alter	-0.029	(0.049)	-0.034	(0.042)
BMI similarity	0.99	(0.53)	0.80	(0.62)
BMI similarity $\times$ reciprocity	-3.78*	(1.44)	-2.17	(1.22)
<i>BMI dynamics</i>				
scale period 1 $\tau_1$	0.065	(0.007)	0.065	(0.009)
scale period 2 $\tau_2$	0.063	(0.017)	0.063	(0.008)
feedback $a$	-0.09	(0.21)	-0.10	(0.24)
intercept $b_0$	1.10*	(0.35)	1.10*	(0.34)
average alter $b_1$	-0.42	(0.94)	-0.39	(0.60)

\* $p$ -value  $< 0.05$ .

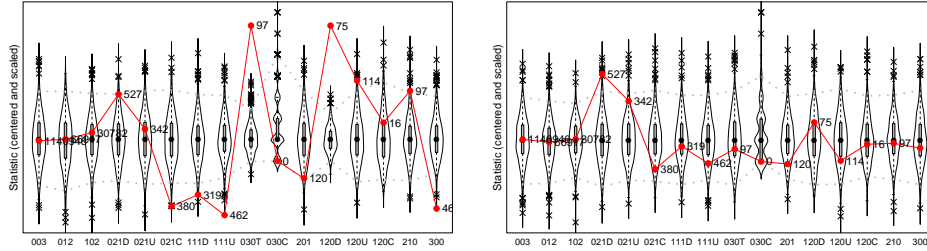


Figure 1: Triad census goodness of fit for model 1 (left) and model 2 (right).

We assess the fit of model 1 by checking how well it represents features of the observed network data that are not directly modeled. The left panel of Figure 1 shows how well the observed triad census (superimposed points connected by line segments) is fit. The triad census is the count of all possible network configurations on three actors and represents local network structure (Wasserman and Faust, 1994). The violin plots show the distributions of the different configurations in the triad census based on 1000 simulations under the estimated model. Clearly, the triadic configurations are not well represented. In model 2, we replace the transitivity effect by a more elaborate set of structural effects, which drastically improves the fit (Figure 1, right).

The substantive conclusions drawn from models 1 and 2 are similar, but not the same. We find a stronger and significant effect of sharing a home group on friendship formation. The shared home group context is not important when it comes to the reciprocation of a friendship tie. Moreover, the interaction effect of BMI similarity and reciprocity is reduced by 57% and not significant in model 2. Accounting for a wider range of network effects makes the BMI-related effects less prominent. Figure 2 shows that the combination of the network and BMI data is well represented by model 2.

**6. Simulation study.** In this section we analyze simulated data similar to the data studied in the application. We study two repeated observations on 156 actors. The first observed network and BMI values as well as the distribution of the covariates are identical to their first observed values in the De la Haye data. We generated 1000 networks and BMI values for the second observation time. Based on these, we re-estimated the parameters. The simulation model is a simplified version of model 1 in the previous section. The data-generating parameter values are rounded numbers close to the estimates obtained for model 1. They are given in Table 5, together

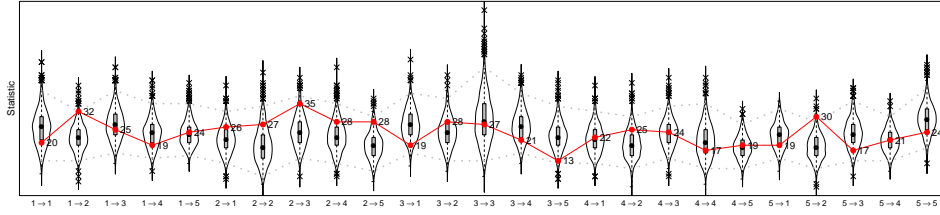


Figure 2: Goodness of fit of the behavior distribution on pairs of related actors (model 2). For group  $i \rightarrow j$  the sender’s BMI value is in the  $i$ th 20% of the observed BMI distribution and the receiver’s value in the  $j$ th 20%.

with the average estimates, the root mean square errors (standard errors of estimation), the rejection rates for testing the data-generating value of the parameter as the null hypothesis (estimating type-I error rates), and the rejection rates for testing that the parameter equals 0 (estimating power). The tests were two-sided tests based on the  $t$ -ratio for the estimated parameters (5% significance level).

Table 5 shows that the parameters are re-estimated well. The estimated type-I error rates do not deviate much from the nominal value (0.05). It appeared that the standard errors and the estimates of the scale parameter  $\tau_1$  were correlated and that the test based on the  $t$ -ratio was not valid here. A log-transformation reduced the correlation (from  $r = 0.373$  to  $-0.051$ ) and the type-I error rate (from 0.096 to 0.077). The last column shows that especially the BMI similarity effect and the average alter effect are hard to detect. This is in line with the general difficulty of disentangling selection and influence effects. Also, in the simulation BMI has only a weak effect on friendship formation and the sample is not large.

**7. Discussion.** Selection and influence are two very different social processes that may yield the same result: a network in which related actors are similar. Network-attribute co-evolution models can help unravel this picture. In this article, we present a model for the co-evolution of social networks and actor attributes that are measured on a continuous scale. This extends the stochastic actor-oriented model (Snijders, Steglich and Schweinberger, 2007; Steglich, Snijders and Pearson, 2010), of which the earlier version assumed actor attributes to be ordinal categorical variables. The model has many potential application areas. Examples include health-related studies, such as the one discussed in Section 5, that explore the effect of social interaction on health-related behaviors, studies on the effect of positive (e.g., helping)

TABLE 5  
*Simulation results: average estimates ( $\hat{\theta}$ ), root mean squared errors (rmse), estimated type-I error rate ( $\alpha$ ), estimated power ( $\beta$ ).*

	$\theta$	$\hat{\theta}$	rmse	$\alpha$	$\beta$
<i>Friendship dynamics</i>					
rate period 1	7.0	6.93	0.55	0.057	
outdegree	-3.3	-3.32	0.14	0.038	1.00
reciprocity	1.5	1.50	0.15	0.035	1.00
transitivity	0.4	0.39	0.045	0.029	1.00
same home group	0.3	0.31	0.14	0.048	0.61
female ego	-0.3	-0.31	0.15	0.043	0.55
female alter	0.3	0.31	0.14	0.041	0.57
same gender	0.8	0.82	0.15	0.044	1.00
BMI similarity	0.3	0.32	0.30	0.035	0.15
<i>BMI dynamics</i>					
scale $\tau_1$	0.1	0.10	0.013	0.077	
feedback $a$	-0.1	-0.11	0.08	0.037	0.24
intercept $b_0$	1.1	1.14	0.31	0.040	0.98
average alter $b_1$	-0.4	-0.43	0.25	0.044	0.37

or negative (e.g., bullying) relations on students' performance, and studies about the formation of partnerships between organizations and their effect on organizational performance.

To model the evolution of continuous variables in continuous time, we use a linear stochastic differential equation. Since linearity is assumed, there exists an analytic expression for the corresponding discrete-time model: the exact discrete model (Bergstrom, 1984). The linear differential equation is conceptually very similar to the regular linear regression model. An advantage of the availability of a model for continuous rather than ordinal discrete actor attributes is that its basis in models for multivariate normal distributions may allow further elaboration exploiting the many known properties for normal distribution models. An example of this is the fact that in the boundary case of a constant network, the moment estimator is the same as the maximum likelihood estimator. Another possibility may be an extension to a random effects model to represent variability among actors.

With respect to substantive conclusions of our example in Section 5 and the low power obtained for testing the two main parameters in the simulation study of Section 6, it should be noted that the social influence of friends on body weight, and the effects of body weight on the selection of friends, if they exist, must be expected to be rather weak. One cannot expect any statistical method to have a reasonably high power for a sample of only 156

adolescents.

In this article, we estimate model parameters using a method of moments procedure. However, other methods of parameter estimation are possible. The other estimation procedures mentioned in Section 4 could be extended to simultaneously estimate the parameters in the continuous attribute evolution model. An extension of the maximum likelihood estimator (Snijders, Koskinen and Schweinberger, 2010), for example, would increase statistical efficiency, and make the model better applicable for data sets containing little information, e.g., for small networks.

We propose an alternative to the model introduced by Snijders, Steglich and Schweinberger (2007) for the case that actor attributes are measured on a continuous scale. Further investigation into the differences between analyzing continuous and discretized actor attribute data is desirable, given that several studies have been conducted with discretized attribute variables. We plan to assess the effects of the loss of information, for various discretization schemes, in a simulation study.

**8. Acknowledgements.** We would like to thank two anonymous referees, the Associate Editor and the Editor for their constructive comments. We also thank Kayla de la Haye for permission to use the data analyzed in Section 5.

#### APPENDIX A: JUSTIFYING THE APPROXIMATION IN SECTION 3.2

The co-evolution scheme of Section 3.2 assumes the effects  $u_{i,t}$  to be constant and the attributes of all actors to evolve independently on the time interval  $[t, t + \Delta t)$ . These assumptions are violated if we include social influence in the attribute evolution model. In this appendix, we justify the approximation occurring in analyses with a social influence effect and discuss a practical issue we would run into without the assumptions.

We can operationalize social influence as an effect of the attributes of the actors to whom an actor is related (his/her alters), on the evolution of his/her own attributes. An example of such an effect is the average alter effect  $\sum_j x_{ij} z_{jk} / x_{i+}$  of attribute  $k$ .<sup>3</sup> We could include the average alter effects of all attributes  $k = 1, \dots, p$  on the evolution of the attributes of actor  $i$ . If these are the only effects in the input  $u_i(t)$ ,  $\tau = 1$  and the differential equation is deterministic ( $G = 0$ ), the stochastic differential equation (1)

---

<sup>3</sup>For notational simplicity, this effect is not centered like the average alter effect proposed in Section 2.

reduces to

$$(28) \quad \frac{dZ_i(t)}{dt} = AZ_i(t) + B \sum_j X_{ij}(t)Z_j(t)/X_{i+}(t).$$

**Example 1.** Consider a constant network  $\mathcal{X}$  on three actors. This network and its corresponding adjacency matrix  $X$  and row-normalized adjacency matrix  $Q$  are given by

$$\mathcal{X} = \begin{array}{c} \textcircled{1} \\ \swarrow \quad \searrow \\ \textcircled{3} \quad \textcircled{2} \\ \longleftarrow \quad \longrightarrow \end{array}, \quad X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

Suppose we consider for the actors the evolution of two attributes  $Z_i(t) = (Z_{i1}, Z_{i2})(t)$ , defined by equation (28) with

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Parameter  $b_{ij}$  corresponds to the effect on an actor's attribute  $i$  of the average value on attribute  $j$  among the actor's alters. If we assume the network to be constant over time and let  $\tilde{Z}(t) = (Z_{11}, Z_{12}, Z_{21}, Z_{22}, Z_{31}, Z_{32})(t)$ , we can combine the differential equations (28) for  $i = 1, 2, 3$  into

$$(29) \quad \frac{d\tilde{Z}(t)}{dt} = \begin{pmatrix} a_{11} & a_{12} & b_{11} & b_{12} & 0 & 0 \\ a_{21} & a_{22} & b_{21} & b_{22} & 0 & 0 \\ 0 & 0 & a_{11} & a_{12} & b_{11} & b_{12} \\ 0 & 0 & a_{21} & a_{22} & b_{21} & b_{22} \\ b_{11}/2 & b_{12}/2 & b_{11}/2 & b_{12}/2 & a_{11} & a_{12} \\ b_{21}/2 & b_{22}/2 & b_{21}/2 & b_{22}/2 & a_{21} & a_{22} \end{pmatrix} \tilde{Z}(t).$$

In general, if  $Q(t)$  denotes the row-normalized version of  $X(t)$  and  $\tilde{Z}(t)$  aggregates the  $Z_i(t)$  in one vector, equations (28) for  $i = 1, \dots, n$  reduce to

$$(30) \quad \frac{d\tilde{Z}(t)}{dt} = [I_n \otimes A + Q(t) \otimes B] \tilde{Z}(t).$$

Using this idea, we can model an influence effect in a stochastic differential equation without violating the assumptions discussed earlier. The drift matrix in the new equation is given by the  $np \times np$  matrix  $\tilde{A} = I_n \otimes A + Q(t) \otimes B$ . Using this formulation, the exact discrete model can be applied exactly, without approximation. However, if the number of actors in a study is large, the repeated evaluation of this exact discrete model is computationally very intensive, as it involves the computation of, e.g.,  $e^{\tilde{A}t}$  and  $\tilde{A}^{-1}$ . Fortunately, in practice there turns out to be little difference between modeling

the attribute evolution using equation (30) and its approximation. In the co-evolution scheme of Section 3.2, the time  $\Delta t$  between consecutive network and attribute updates is exponentially distributed with expected value  $E(\Delta t) = 1/(n\lambda_m)$ . If  $n\lambda_m$  is large, the  $\Delta t$  are small and so are the changes occurring in this interval. Therefore, the approximation error will be small. This is illustrated in the following example.

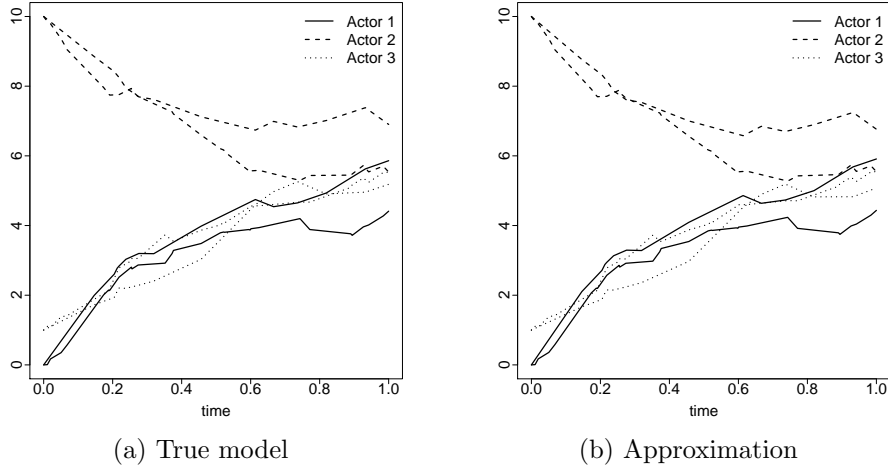


Figure 3: Comparison of the true model and the approximation for initial values  $Z(0) = (0, 10, 1)$  and  $E(\Delta t) = 0.05$ .

**Example 2.** Consider again network  $\mathcal{X}$ . We model the evolution of a (single) attribute  $Z_i(t)$  of actors  $i = 1, 2, 3$  in this constant network by

$$(31) \quad dZ_i(t) = [-2Z_i(t) + 6 + X_{ij}Z_j(t)/X_{i+}]dt + dW_i(t).$$

These equations can be reduced to one equation as in (30). We will refer to the latter as the true model and to the scheme of Section 3.2 as applied to the former as the approximation. We study the evolution processes on the time interval  $[0, 1]$ , using common random numbers in the generation of sample paths. We let the times between consecutive attribute updates  $\Delta t$  be exponentially distributed with specified  $E(\Delta t)$ . Figures 3a and 3b show two sample paths for each of the actors, for fixed initial values and  $E(\Delta t) = 0.05$ . The average absolute difference per actor between the values at  $t = 1$  for the true model and the approximation is 0.06. This is small compared to the mean absolute deviation  $\text{mad } Z(t_1)$  of the true values at  $t = 1$ , averaged over the two sets of sample paths, which is 0.51.

Figure 4 shows the average absolute difference per actor for different levels of  $E(\Delta t)$ . For each level, 100 true and approximated evolution processes are simulated, with initial attribute values sampled uniformly on  $[0, 10]$ . The figure shows



that the differences between the true and approximated processes at time  $t = 1$  increase with the  $E(\Delta t)$  level, as expected. Given the variation between the actor’s attributes values at  $t = 1$ , the level  $10^{-2}$  already yields a low within-actor approximation error. In practice, this value is often much smaller. For example, the value of  $E(\Delta t)$  in the application in this article is smaller than  $10^{-3}$ .

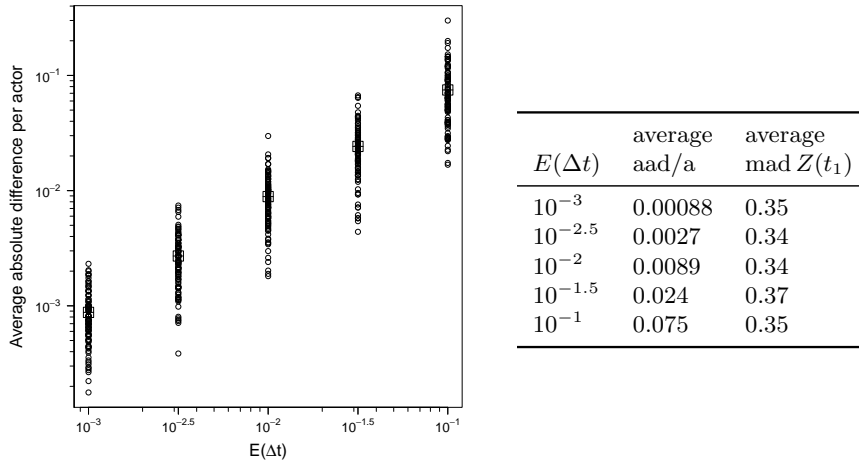


Figure 4: Comparison of the true model and the approximation: the average absolute difference per actor (aad/a) for different levels of  $E(\Delta t)$ . The means per level are indicated by the squares in the figure and are given in the table.

### APPENDIX B: COVARIANCE ESTIMATION

Estimating  $\text{cov}(\hat{\theta})$  by a bootstrap procedure is inconvenient, as each of the multiple estimation runs is time-consuming (Schweinberger and Snijders, 2007). Instead, we use the approximation given in equation (18). Monte Carlo estimation of  $\Sigma_\theta = \text{cov}_\theta S(Y)$  is straightforward. The issue is how to define an estimator of  $D_\theta = \frac{\partial}{\partial \theta} E_\theta S(Y)$ . Let  $J_\theta$  denote the score function of  $Y$ , i.e.,  $J_\theta = \frac{\partial}{\partial \theta} \log p_\theta(Y)$ . It can be shown that

$$(32) \quad D_\theta = E_\theta \left( S(Y) J_\theta^\top \right).$$

See Schweinberger and Snijders (2007) for more details, e.g., the use of control variates to reduce the variance in the estimation of  $D_\theta$ . They derive the score  $J_\theta$  with respect to the network evolution parameters. Below we obtain expressions for the score  $J_\theta$  with respect to the attribute evolution parameters. Give these score functions, we can estimate  $D_\theta$  from Monte Carlo simulations, based on (32).

We assume, as in our illustration in Section 5, that there is a single continuous attribute, the evolution of which we model by

$$(33) \quad dZ_i(t) = \tau_m [a Z_i(t) + b^\top u_i(t)] dt + \sqrt{\tau_m} dW_i(t),$$

where  $a \in \mathbb{R}$  and  $b \in \mathbb{R}^p$ . The calculations below can be generalized for higher-dimensional  $Z_i(t)$ , but each extra dimension brings along additional complexity. The log-likelihood  $\ell = \log p_\theta(z_{t+\Delta t})$  of one step of the corresponding exact discrete model for all  $n$  actors is

$$(34) \quad -\frac{n}{2} (\log 2\pi + \log \sigma_{\Delta t}^2) - \frac{1}{2\sigma_{\Delta t}^2} \sum_{i=1}^n (\epsilon_{i,t,\Delta t})^2,$$

where  $\epsilon_{i,t,\Delta t} = z_{i,t+\Delta t} - \mu_i(\Delta t, z_{i,t}, u_{i,t})$  is the random term with variance  $\sigma_{\Delta t}^2$  for actor  $i$  having evolved over a period  $\Delta t$  after time  $t$ , and

$$\begin{aligned} \mu_i(\Delta t, z_{i,t}, u_{i,t}) &= e^{a\tau_m \Delta t} z_{i,t} + \frac{1}{a} (e^{a\tau_m \Delta t} - 1) b^\top u_{i,t} \\ \sigma_{\Delta t}^2 &= \frac{(e^{2a\tau_m \Delta t} - 1)}{2a}. \end{aligned}$$

We will determine the score functions for this single step. The total score can be computed by adding the score components of the separate attribute evolution steps taken during the simulation procedure specified in Section 3.2. The score function with respect to  $b_k$  is

$$\frac{\partial \ell}{\partial b_k} = \frac{1}{\sigma_{\Delta t}^2} \sum_{i=1}^n \epsilon_{i,t,\Delta t} \times \frac{1}{a} (e^{a\tau_m \Delta t} - 1) (u_{i,t})_k = \frac{2}{(e^{a\tau_m \Delta t} + 1)} \sum_{i=1}^n \epsilon_{i,t,\Delta t} (u_{i,t})_k.$$

The score function with respect to  $a$  is

$$\begin{aligned} \frac{\partial \ell}{\partial a} &= \frac{n}{2a} - \frac{n}{2} \frac{2\tau_m \Delta t e^{2a\tau_m \Delta t}}{e^{2a\tau_m \Delta t} - 1} - \frac{\partial}{\partial a} \left[ \frac{1}{2\sigma_{\Delta t}^2} \sum_{i=1}^n (\epsilon_{i,t,\Delta t})^2 \right] \\ &= \frac{n}{2a} \left( 1 - \frac{\tau_m \Delta t e^{2a\tau_m \Delta t}}{\sigma_{\Delta t}^2} \right) - \sum_{i=1}^n (\epsilon_{i,t,\Delta t})^2 \frac{\partial}{\partial a} \frac{1}{2\sigma_{\Delta t}^2} + \frac{1}{\sigma_{\Delta t}^2} \sum_{i=1}^n \epsilon_{i,t,\Delta t} \frac{\partial \mu_i}{\partial a}, \end{aligned}$$

where

$$\begin{aligned} \frac{\partial}{\partial a} \frac{1}{2\sigma_{\Delta t}^2} &= \frac{(e^{2a\tau_m \Delta t} - 1) - 2a\tau_m \Delta t e^{2a\tau_m \Delta t}}{(e^{2a\tau_m \Delta t} - 1)^2} = \frac{1}{2a\sigma_{\Delta t}^2} - \frac{\tau_m \Delta t e^{2a\tau_m \Delta t}}{2a\sigma_{\Delta t}^4} \\ \frac{\partial \mu_i}{\partial a} &= \tau_m \Delta t e^{a\tau_m \Delta t} z_{i,t} + b^\top u_{i,t} \frac{a\tau_m \Delta t e^{a\tau_m \Delta t} - e^{a\tau_m \Delta t} + 1}{a^2} \\ &= \tau_m \Delta t \mu_i + b^\top u_{i,t} \left( \frac{\tau_m \Delta t}{a} - \frac{e^{a\tau_m \Delta t} - 1}{a^2} \right). \end{aligned}$$

The score function with respect to  $\tau_m$  is

$$\begin{aligned} \frac{\partial \ell}{\partial \tau_m} &= -\frac{na\Delta t e^{2a\tau_m\Delta t}}{e^{2a\tau_m\Delta t} - 1} - \frac{\partial}{\partial \tau_m} \left[ \frac{1}{2\sigma_{\Delta t}^2} \sum_{i=1}^n (\epsilon_{i,t,\Delta t})^2 \right] \\ &= -\frac{n\Delta t e^{2a\tau_m\Delta t}}{2\sigma_{\Delta t}^2} + \frac{\Delta t e^{2a\tau_m\Delta t}}{2\sigma_{\Delta t}^4} \sum_{i=1}^n (\epsilon_{i,t,\Delta t})^2 + \frac{1}{\sigma_{\Delta t}^2} \sum_{i=1}^n \epsilon_{i,t,\Delta t} \frac{\partial \mu_i}{\partial \tau_m}, \end{aligned}$$

where

$$\frac{\partial \mu_i}{\partial \tau} = a\Delta t e^{a\tau_m\Delta t} z_{i,t} + \Delta t b^\top u_{i,t} e^{a\tau_m\Delta t} = \Delta t(a\mu_i + b^\top u_{i,t}).$$

## REFERENCES

- AGNEESSENS, F. and WITTEK, R. (2008). Social capital and employee well-being: disentangling intrapersonal and interpersonal selection and influence mechanisms. *Revue Française de Sociologie* **49** 613–637.
- AMATI, V., SCHÖNENBERGER, F. and SNIJDERS, T. A. B. (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** 140–165.
- ARNOLD, L. (1974). *Stochastic differential equations*. Wiley, New York.
- BERGSTROM, A. R. (1984). Continuous time stochastic models and issues of aggregation over time. In *Handbook of econometrics, Volume II*. (Z. Griliches and M. D. Intriligator, eds.) 1146–1212. Elsevier Science Publishers BV, Amsterdam.
- BERGSTROM, A. R. (1988). The history of continuous-time econometric models. *Econometric Theory* **4** 356–383.
- BLACK, F. and SCHOLES, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy* **81** 637–654.
- CHRISTAKIS, N. A. and FOWLER, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine* **357** 370–379.
- COHEN-COLE, E. and FLETCHER, J. M. (2008). Is obesity contagious? Social networks vs. environmental factors in obesity epidemic. *Journal of Health Economics* **27** 1382–1387.
- DIJKSTRA, J. K., LINDENBERG, S., VEENSTRA, R., STEGLICH, C., ISAACS, J., CARD, N. A. and HODGES, E. V. E. (2010). Influence and selection processes in weapon carrying during adolescence: The roles of status, aggression, and vulnerability. *Criminology* **48** 187–220.
- DIJKSTRA, J. K., GEST, S. D., LINDENBERG, S., VEENSTRA, R. and CILLESSEN, A. H. N. (2012). The emergence of weapon carrying in peer context. Testing three explanations: the role of aggression, victimization, and friends. *Journal of Adolescent Health* **50** 371–376.
- HAMERLE, A., SINGER, H. and NAGL, W. (1993). Identification and estimation of continuous time dynamic systems with exogenous variables using panel data. *Econometric Theory* **9** 283–295.
- HOLLAND, P. W. and LEINHARDT, S. (1977). A dynamic model for social networks. *Journal of Mathematical Sociology* **5** 5–20.
- HUNTER, D. R. (2007). Curved exponential family models for social network. *Social Networks* **29** 216–230.
- KOSKINEN, J. H. and SNIJDERS, T. A. B. (2007). Bayesian inference for dynamic social network data. *Journal of Statistical Planning and Inference* **137** 3930–3938.

- KUSHNER, H. J. and YIN, G. G. (2003). *Stochastic approximation and recursive algorithms and applications*, 2 ed. Springer, New York.
- LEHMANN, E. L. (1999). *Elements of large sample theory*. Springer, New York.
- McFADDEN, D. (1974). Conditional logit analysis of qualitative choice behavior. In *Frontiers in Economics* (P. Zarembka, ed.) 105–142. Academic Press, New York.
- MERTON, R. C. (1990). *Continuous-time finance*. Basil Blackwell, Oxford UK.
- ODD, J. H. L. and JANSEN, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika* **65** 199–215.
- ODD, J. H. L., FOLMER, H., PATUELLI, R. and NIJKAMP, P. (2012). Continuous-time modeling with spatial dependence. *Geographical Analysis* **44** 29–46.
- DE LA HAYE, K., ROBINS, G., MOHR, P. and WILSON, C. (2011). Homophily and contagion as explanations for weight similarities among adolescent friends. *Journal of Adolescent Health* **49** 421–427.
- RIPLEY, R. M., SNIJDERS, T. A. B., BODA, Z., VÖRÖS, A. and PRECIADO, P. (2016). Manual for RSiena (version December 2016), Oxford: University of Oxford, Department of Statistics; Nuffield College.
- ROBBINS, H. and MONRO, S. (1951). A stochastic approximation method. *Annals of Mathematical Statistics* **22** 400–407.
- SCHWEINBERGER, M. and SNIJDERS, T. A. B. (2007). Markov models for digraph panel data: Monte Carlo-based derivative estimation. *Computational Statistics & Data Analysis* **51** 4465–4483.
- SHALIZI, C. R. and THOMAS, A. C. (2011). Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research* **40** 211–239.
- SINGER, H. (1996). Continuous-time dynamic models for panel data. In *Analysis of change: advanced techniques in panel data analysis* (U. Engel and R. J., eds.) 113–133. De Gruyter, Berlin.
- SNIJDERS, T. A. B. (2001). The statistical evaluation of social network dynamics. In *Sociological methodology* (M. Sobel and M. Becker, eds.) 361–395. Basil Blackwell, Boston and London.
- SNIJDERS, T. A. B. (2005). Models for longitudinal network data. In *Models and methods in social network analysis* (P. J. Carrington, J. Scott and S. S. Wasserman, eds.) Cambridge University Press, New York.
- SNIJDERS, T. A. B., KOSKINEN, J. and SCHWEINBERGER, M. (2010). Maximum likelihood estimation for social network dynamics. *The Annals of Applied Statistics* **4** 567–588.
- SNIJDERS, T. A. B., STEGLICH, C. E. G. and SCHWEINBERGER, M. (2007). Modeling the co-evolution of networks and behavior. In *Longitudinal models in the behavioral and related sciences* (K. van Montfort, H. Oud and A. Satorra, eds.) 41–71. Cambridge University Press, New York.
- STEELE, J. M. (2001). *Stochastic calculus and financial applications*. Springer-Verlag, New York.
- STEGLICH, C. E. G., SNIJDERS, T. A. B. and PEARSON, M. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological Methodology* **40** 329–392.
- VOELKLE, M. C., ODD, J. H. L., DAVIDOV, E. and SCHMIDT, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods* **17** 176–192.
- WASSERMAN, S. S. (1980). Analyzing social networks as stochastic processes. *Journal of the American Statistical Association* **75** 280–294.
- WASSERMAN, S. and FAUST, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press, Cambridge UK.

WEERMAN, F. (2011). Delinquent peers in context: A longitudinal network analysis of selection and influence effects. *Criminology* **49** 253–286.

N. M. D. NIEZINK  
T. A. B. SNIJDERS  
UNIVERSITY OF GRONINGEN  
GROTE ROZENSTRAAT 31  
9712 TG GRONINGEN  
THE NETHERLANDS  
E-MAIL: n.m.d.niezink@rug.nl  
t.a.b.snijders@rug.nl