Forms of Dependence

Block, Per; Stadtfeld, Christoph; Snijders, Tom A. B.

Published in:
Sociological Methods & Research

DOI:
10.1177/0049124116672680

Article

Sociological Methods & Research
2019, Vol. 48(1) 202-239

DOI: 10.1177/0049124116672680
journals.sagepub.com/home/smr
SAGE

Forms of Dependence: Comparing SAOMs and ERGMs From Basic Principles

Per Block[1], Christoph Stadtfeld[1], and Tom A. B. Snijders[2,3]

Abstract

Two approaches for the statistical analysis of social network generation are widely used; the tie-oriented exponential random graph model (ERGM) and the stochastic actor-oriented model (SAOM) or Siena model. While the choice for either model by empirical researchers often seems arbitrary, there are important differences between these models that current literature tends to miss. First, the ERGM is defined on the graph level, while the SAOM is defined on the transition level. This allows the SAOM to model asymmetric or one-sided tie transition dependence. Second, network statistics in the ERGM are defined globally but are nested in actors in the SAOM. Consequently, dependence assumptions in the SAOM are generally stronger than in the ERGM. Resulting from both, meso- and macro-level properties of networks that can be represented by either model differ substantively and analyzing the same network employing ERGMs and SAOMs can lead to distinct results. Guidelines for theoretically founded model choice are suggested.

[1] Chair of Social Networks, ETH Zürich, Zürich, Switzerland
[2] Department of Sociology, University of Groningen, Groningen, the Netherlands
[3] Nuffield College, University of Oxford, Oxford, UK

Corresponding Author:
Per Block, Chair of Social Networks, ETH Zürich, Clausiusstr. 52, Zürich, 8092, Switzerland.
Email: per.block@gess.ethz.ch

Keywords
social networks, statistical network modeling, ERGM, SAOM, Siena

Introduction

The statistical analysis of social networks has made considerable advances over the last decade and is by now firmly established in many fields of the social and psychological sciences. Among its main focuses is the analysis of the different processes that give rise to empirically observed networks and network structures (Rivera, Soderstrom, and Uzzi 2010). There are two generally used types of models that enable the simultaneous analysis of the diverse mechanisms underlying network formation that can flexibly account for the endogenous dependencies of social networks: The exponential random graph model (ERGM; Lusher, Koskinen, and Robins 2013), which dominates the literature on cross-sectional networks, and the stochastic actor-oriented model (SAOM; Snijders, van de Bunt, and Steglich 2010), mainly used for longitudinal network data.

The ERGM is the older of the two models with its probabilistic and statistical basis being developed since the 1950s.[1] The ERGM is a probability model for networks. As network ties are interdependent, and consequently standard statistical methods cannot be applied, it is assumed that all dependence between ties can be captured by local configurations (Robins, Pattison, et al. 2007). This is in line with the idea that social networks emerge from local processes. Intuitively, the presence of a tie in a social network will be conditionally independent from ties that are "far away" in social space.[2] Thus, the ERGM is based on assumptions of conditional independence of ties (Frank and Strauss 1986; Wasserman and Pattison 1996). Statistical dependence between ties is expressed in terms of local configurations of ties in small connected substructures, such as triangles or star-shaped structures—network ties that do not share local configurations are assumed to be conditionally independent. Patterns of dependence in an ERGM are commonly interpreted as evidence for various local processes that operate simultaneously (Lusher and Robins 2013; Robins and Lusher 2013). The formulation of the ERGM is interpretable as the conditional probability of a tie, given the rest of the network (Wasserman and Pattison 1996).

The SAOM was introduced in 1996 (Snijders 2001) using elements of econometric discrete choice models (McFadden 1974) and dynamic Markov models for social networks (Holland and Leinhardt 1977) as an attempt to integrate theoretical and statistical models for the evolution of social networks. The SAOM is a probability model for network changes in continuous

time, and the estimation theory was elaborated for network data available at discrete time points. Network ties are assumed to change one tie at a time (Holland and Leinhardt 1977), with actors choosing which of the outgoing ties to change based on a multinomial choice probability model (McFadden 1974). Dependence between ties is modeled by allowing the actors' choices concerning tie changes to be influenced by the embedding of the potential ties in local configurations. These are akin to the local configurations in the ERGM but take the perspective of a focal actor, while there is no focal actor in the ERGM. Due to its actor-oriented formulation, the SAOM offers a direct interpretation of parameters as reflecting differential probabilities of actors' choices with regard to their outgoing ties.

What should be regarded as the unit of analysis is a moot point in these models. One might say that for the ERGM, this is the tie variable, whereas it is the actor for the SAOM. However, this would be ambiguous, or incomplete, because in both models there is potentially—depending on the specification—a pervasive dependence between all network ties. It is clear that for the SAOM, the actors have a special role in the model: In the ERGM, the ties are considered in the context of the rest of the network, while in the SAOM, the ties are considered in the context of the other outgoing ties of the sending actor, which in turn are considered in the context of the rest of the network.

While both models have a similar mathematical basis, the detailed differences and similarities between the models have not been studied in depth. The consequences of the different model assumptions on the meso- and macro-level structures that can be represented in principle by ERGMs and SAOMs, and the arguments that should determine model choice to a given empirical problem, require a deeper treatment than available in the current literature.

There are some general introductions to network modeling (such as Broekel et al. 2014) treating both models, but these do not present a detailed comparison of the consequences of the different model assumptions. To our knowledge, there are two studies that explicitly compare the two models (or a variant of the ERGM, the tERGM [temporal ERGM], to the SAOM), conducted by Desmarais and Cranmer (2012) and Lerner et al. (2013). However, both articles do not explain the fundamental differences between the principles underlying the models but instead engage in empirical, fit-based model comparison. Furthermore, they come to differing conclusions. It may be noted that Snijders (2001, section 9) gives a specification of the SAOM under which its stationary distribution is an ERGM. However, this requires an artificial specification of the rate function which is never used in practice. This result therefore is of theoretical importance only.

The present article aims at a deeper understanding of the differences between the two models on a conceptual level. We limit our comparison to directed networks, as the undirected case is not unambiguously defined in the SAOM. A basic difference is that the ERGM is defined as a probability distribution for graphs, or for tie variables given the rest of the graph, while the SAOM is defined as a probability distribution for changes in tie variables given the rest of the graph. The ERGM has a firm basis in conditional independence assumptions, thanks to the Hammersley–Clifford theorem (Frank and Strauss 1986; Lusher et al. 2013). This is lacking in the SAOM, but on the other hand, this gives to the SAOM the freedom of a more versatile specification. Second, although similar statistics are used, in the ERGM, they are defined on graph level while they are nested in actors in the SAOM. This gives an extra source of dependence between ties. Parameters in the ERGM are interpreted as tendencies in the graph or of pairs of actors conditional on the rest of the graph, whereas parameters in the SAOM are interpreted as tendencies of actors. The focus of our article is to establish analytically how these differences translate into distinct meso- and macro-level features of networks that can be modeled appropriately by either approach.

The remainder of this article is structured as follows. First, we outline the different levels on which the models can be compared and justify why we focus our attention on the comparison between the tie- versus actor-oriented nature of the two models. Subsequently, we discuss how the model assumptions imply a different extent of dependency between ties and how the models allow for symmetric or asymmetric transition dependence between ties, accompanied by the implications for nonmodeled network statistics. This is followed by showing how the aforementioned differences can lead to different conclusions in an empirical example. Finally, the results and their implications as well as recommendation for researchers choosing between either models are discussed.

Levels of Comparison

In the literature, the SAOM and the directed ERGM are generally distinguished on two dimensions. First, the ERGM traditionally applies to cross-sectional data, while the SAOM was created to analyze longitudinal network data. This distinction now is starting to be outdated. Recent developments have proposed numerous extensions to the ERGM that allow the analysis of longitudinal data, such as the longitudinal ERGM (lERGM; Snijders and Koskinen 2013), the separable temporal ERGM (StERGM ; Krivitsky and Handcock 2014), and different variants of the tERGM (Desmarais and Cranmer 2012; Hanneke, Fu, and Xing 2010). Also, researchers are starting to use the SAOM

to model cross-sectional data (Snijders and Steglich 2015), even though published applications still are rare. The cross-sectional model implied by the SAOM is its stationary distribution, entirely comparable to the way of defining the ERGM as a stationary distribution implied by dependent tie processes.

Second, the ERGM is tie oriented, while the SAOM is actor oriented. These are admittedly vague terms. The original introduction of actor-oriented network models (Snijders 1996, 2001) was in terms of actors myopically optimizing their evaluation of the network, but in more recent publications (e.g., Snijders et al. 2010), this is presented as just "one possible interpretation." ERGMs may also be interpreted as actors, or pairs of actors, myopically optimizing their ties one by one. Such stochastic optimization interpretations result from the same interpretations of binary and multinomial logistic regression models (McFadden 1974) and do not essentially distinguish between these two classes of models. The difference between tie orientation and actor orientation can be understood in the sense that a natural definition of the ERGM is in terms of conditional probabilities for tie variables given the rest of the network, while the SAOM is defined in terms of choices made by actors for their entire collection of outgoing ties, given the network. To understand how these differences in approach lead to different properties of the models, we compare the classical ERGM to an SAOM applied to cross-sectional data. We compare the two models both from a *micro-perspective* focusing on how single tie changes are modeled and from a *macro-perspective* focusing on emergent differences of the stationary distributions of either model at the (cross-sectional) network level.

The Model Definitions

Both models are in the tradition of generalized linear modeling and have a linear predictor in their heart. For the ERGM, the linear predictor is defined globally for the entire graph, and we denote it by:

$$\sum_k \theta_k z_k(x),$$

where $z_k(x)$ are statistics of the network x and θ_k are statistical parameters. For the SAOM, the linear predictor is defined with reference to the actors, and for actor i, we denote it as:

$$\sum_k \beta_k s_{k,i}(x),$$

where $s_{k,i}(x)$ are statistics of the network x, as "seen" from the perspective of node i, and β_k are the parameters.

In order to analytically compare the two models, in this section, we define a dynamic micro-level as well as a macro-level formulation of both models. A dynamic micro-level formulation describes the probabilities to transition from one network state to an adjacent network state, that is, how likely is the change in each tie variable given a current network. For both models, these transition probabilities define a Markov chain on the space of all possible network states. A macro-level definition of the models is a formulation of a probability distribution on the space of all networks and should be consistent in the sense that it is the stationary distribution of the micro-model.

There are different ways to define the ERGM. A macro-level definition is the probability distribution for random graphs X given by:

$$P_{\text{ERGM}}(X = x; \theta) = \frac{\exp\left(\sum_k \theta_k z_k(x)\right)}{\kappa}, \tag{1}$$

where θ is a vector of statistical parameters with K elements indexed by k, $z_k(x)$ are statistics describing the network x, and κ is a normalizing constant. This normalizing constant cannot be easily calculated; this has severe consequences for the analysis but is immaterial for our present discussion.

A micro-level definition of the ERGM can be formulated at the tie level. This is provided by the Metropolis algorithm explained in Koskinen and Snijders (2013), which is one of the algorithms that can be used to simulate draws from the ERGM, because it has the ERGM as its limiting stationary distribution. Many scholars probably will agree that this is the "natural" micro-level specification of the ERGM, but it is not unique, as there are many graph processes that converge to the same stationary distribution. In this algorithm, given that the current state of the network is x, the equation that specifies the transition probability from a network state x to an adjacent network state $x^{\pm ij}$ that differs only by the state of the tie from i to j is:

$$\begin{aligned} p_{\text{ERGM}}(x \to x^{\pm ij}; \theta) &= \frac{1}{N(N-1)} \cdot \frac{\exp\left(\sum_k \theta_k z_k(x^{\pm ij})\right)}{\exp\left(\sum_k \theta_k z_k(x)\right) + \exp\left(\sum_k \theta_k z_k(x^{\pm ij})\right)} \\ &= \underbrace{\frac{1}{N(N-1)}}_{\text{select tie } x_{ij}} \cdot \underbrace{\frac{\exp\left(\sum_k \theta_k \Delta z_k(x, x^{\pm ij})\right)}{1 + \exp\left(\sum_k \theta_k \Delta z_k(x, x^{\pm ij})\right)}}_{\text{change tie } x_{ij}}, \end{aligned} \tag{2}$$

where θ, k, and $z_k(x)$ are as above, N is the number of nodes in the network, and $\Delta z_k(x, x^{\pm ij}) = z_k(x^{\pm ij}) - z_k(x)$ is the difference in statistics, called the *change statistic* (Wasserman and Pattison 1996), between network x and $x^{\pm ij}$. The first factor describes the probability to select randomly the specific ordered pair (i,j) that distinguishes network state x from $x^{\pm ij}$, while the second factor (a logit model) specifies the probability to change the tie, as opposed to not changing and thus remaining in the current state x. The second factor follows from the conditional distribution of the tie variable from i to j, given the rest of the graph.

The two definitions are equivalent because the probability to reach a specific network x in a random walk over the space of all possible networks with transition probabilities as defined in equation (2) (the stationary distribution) is given by equation (1).

The fact that the stationary distribution of equation (2) is the ERGM is helpful for the interpretation of parameters, as it associates a micro-model to the macro-model given by equation (1). If the researcher is willing to assume that the network is in equilibrium, the parameters then may be interpreted in terms, for example, of motivations of actors.

Contrary to this, the SAOM has one natural definition on the microlevel of network transitions. It is defined by the probability of transitions between adjacent network states (i.e., states differing by at most one tie) that depend on local statistics defined from the perspective of an actor. The probability of transitioning from a network x to a network $x^{\pm ij}$ is defined as the probability of actor i to change the relation to actor j:

$$p_{\mathrm{SAOM}}(x \to x^{\pm ij}; \beta) = \rho_i \times \frac{\exp\left(\sum_k \beta_k s_{k,i}(x^{\pm ij})\right)}{\sum_h \exp\left(\sum_{jk} \beta_k s_{k,i}(x^{\pm ih})\right)}$$

$$= \underbrace{\rho_i}_{\text{select actor } i} \times \underbrace{\frac{\exp\left(\sum_k \beta_k \Delta s_{k,i}(x, x^{\pm ij})\right)}{1 + \sum_{h'} \exp\left(\sum_{jk} \beta_k \Delta s_{k,i}(x, x^{\pm ih'})\right)}}_{\text{actor } i \text{ changing tie } \pm ij}, \tag{3}$$

where β is a statistical parameter with K elements, $s_{k,i}(x^{\pm ij})$ are statistics that describe the network from the perspective of actor i if it would change the tie to actor j, and ρ_i is the rate function for actor i. Variable h runs over all members of the network including i, h' all members without i. Function $\Delta s_{k,i}(x, x^{\pm ij})$ is the difference in statistics between networks x and $x^{\pm ij}$ from

actor i's perspective, called the *change statistic* just as for the ERGM. The first term of this equation specifies the probability that actor i rather than any other actor is allowed to make a tie change. In many applications, these probabilities are uniform, $\rho_i = 1/N$, and we limit our analysis to this case. The second term defines the probability that actor i executes the tie change to actor j, as opposed to any actor h, and is defined like in the multinomial logit model. The probabilities do not sum to unity when j assumes all values from 1 to n except i, as actor i can make the choice not to change the network at all.

The macro-level definition of the cross-sectional SAOM is, in principle, the stationary distribution of a Markov chain defined by the transition probabilities in equation (3). For this distribution, however, the probability of observing a graph x cannot be expressed in a closed form as in the case of the ERGM. However, the definition of the stationary distribution corresponding to equation (3) allows us to express these probabilities in a recursive form:

$$P_{\mathrm{SAOM}}(x;\beta) = \sum_{(h,l)} P_{\mathrm{SAOM}}(x^{\pm hl};\beta) \cdot p_{\mathrm{SAOM}}(x^{\pm hl} \to x;\beta), \qquad (4)$$

with (h,l) running over all pairs of actors in the network and the additional variables as explained below equation (3). This formula is not specific to the SAOM, but a general formulation for the stationary distribution of a Markov chain. In principle, the corresponding system of linear equations could be solved, but the dimension of this system ($2^{N(N-1)}$) is too high for practical computations. In practice, the stationary distribution of an SAOM is not tractable for even moderately sized networks ($N > 5$).

With equations (2) to (4), we can populate all entries in Table 1. Using the insights from these equations, we can now analytically compare how the micro-level transition probabilities from one network state to another differ between the SAOM and the ERGM. In a second step, we can extrapolate how the differences in transition probabilities lead to differences in the structure of the stationary distributions of seemingly equivalent models in the ERGM and SAOM. This, in turn, shall allow us to draw conclusions why we can find distinct, or even contradictory results in analyses employing these two models, and give help to decide which model should be used for which empirical questions. The detailed comparison procedure is outlined in the following section.

Analytical Approach

To have a basis for comparison, we study specifications of the ERGM and the SAOM that are equivalent with regard to the expected value of explicitly

Table 1. Mathematical Definitions of the ERGM and SAOM.

	Defined on	
	Microlevel	Macrolevel
Tie oriented	Equation (2)	Equation (1)
Actor oriented	Equation (3)	Equation (4)

Note: ERGM = exponential random graph model; SAOM = stochastic actor-oriented model.

modeled graph statistics, that is, the statistics in the functions $z_k(x)$. In other words, we compare the stationary distribution of ERGMs and SAOMs of models that appear identical at first glance. In order to *understand* how these seemingly identical models bring about different unmodeled meso- and macro-level properties, we compare how the micro-level transition probabilities between adjacent network states (equations 2 and 3) differ. By contrasting the transition probabilities, we can (i) show conceptually how they differ, and how these differences extrapolate to differing network-level outcomes, and (ii) show mathematically that there are no common micro-level model specifications under which the transition probabilities for the two models are identical.

The detailed approach we take in the subsequent sections is as follows. Using simple configurations of three nodes, we show how a set of transition probabilities from one configuration to a set of other configurations that are identical in the ERGM can differ in the SAOM by virtue of its actor-oriented nature. Starting from these micro-level differences, we formulate a set of expectations how this results in macro-level differences for a range of network-level properties.

To assess these expectations, we compare the stationary distributions of an ERGM and an SAOM that include the "same effects" and that have the same expected macro-level statistics related to these effects. The "same" effects refer to the same macro-level statistics that they model. For example, in Dependence in a Density Model subsection, we compare the stationary distribution of an ERGM and an SAOM that only include a density parameter, with an identical expected number of ties. The unmodeled network-level feature that is compared in the section is the implied outdegree distribution (OD).

As calculating the exact stationary distribution for larger networks is computationally not feasible, we use networks of five nodes for our comparison. A network of five nodes already has $2^{5\times4} = 1{,}048{,}576$ possible states, of which 9,036 configurations are nonisomorphic,[3] which is a manageable

number for our purposes. At the same time, a network of five nodes is large enough to show the consequences of the model choices we are interested in. In the following two sections, we briefly outline how we obtain the statistical parameters θ and β for the micro-level ERGMs and SAOMs that generate networks with the same values for the macro-level network statistics (called target statistics).

Procedure for ERGMs

For the ERGM, the procedure to find a parameter vector $\hat{\theta}$ that defines a stationary distribution on the space of all five-node networks with specific expected macro-level target statistics $\bar{z}$ is relatively straightforward. Finding a $\hat{\theta}$ that, for example, generates a distribution of networks with an expected density statistic of 0.3 is a convex optimization problem.[4] We can apply a numerical algorithm to find a $\hat{\theta}_{\text{density}}$ that generates a stationary distribution with average target statistics $\bar{z}_{density} = 0.3$. Finding the $\hat{\theta}$ for which the expected statistics under the model are equal to the empirically observed statistic is the solution to the so-called Moment equation:

$$E_{\theta}(z(x)) = z_{\text{obs}}(x). \tag{5}$$

For exponential family models, such as the ERGM, the solution to the Moment equation is also the maximum likelihood estimate.

The general procedure is as follows: An initial parameter θ_0 is defined and the expected target statistics are obtained by weighting the statistics of all 9,036 networks with equation (1) and their relative frequency (the number of isomorphic graphs with this specific structure). Based on the deviation of the average statistics from $\bar{z}$, θ_0 is updated and the new average statistics are calculated. This procedure is repeated until a $\theta_n = \hat{\theta}$ is found that generates a stationary distribution with average target statistics $\bar{z}$. We use a simplex algorithm based on Nelder and Mead (1965) implemented in the statistical software R (R Core Team 2015), which is relatively slow but very robust, and proved most useful for our purpose. Alternatively, gradient-based or Newton-based methods could be employed.

Procedure for SAOMs

The procedure to obtain a stationary distribution with macro-level target statistics $\bar{s}$ in the SAOM is more involved, because the macro-level probability distribution on the space of all 9,036 nonisomorphic network states (the stationary distribution) cannot be expressed in a closed form; it is defined

implicitly as the solution of equation (4) for a given β and given graph statistics $s(x)$. However, in case of a five-node network, we can calculate it as follows. The probability of all potential transitions between the 9,036 network states for a given β and $s(x)$ is calculated to obtain a complete $9{,}036 \times 9{,}036$ transition matrix, taking the frequencies of isomorphic transitions into account. The stationary distribution is equal to the first eigenvector of this transition matrix (the eigenvector that is associated with the eigenvalue 1). Employing this more complex way of calculating the stationary distribution of an SAOM, from this point onward the same procedure as for the ERGMs can be applied. An initial parameter β_0 is iteratively updated until a $\hat{\beta}$ is found that generates a stationary distribution with average target statistics $\bar{s}$.

Dependence Between Ties

In the literature on ERGMs, various types of dependence between network ties are distinguished. A partially ordered dependence hierarchy for ERGMs is presented in Pattison and Snijders (2013). We mention here only the types of dependence that are most frequently encountered in the practical modeling literature. (1) The Bernoulli model of *independence* between ties. In an Erdös–Rényi random graph, that is, an ERGM including only a density parameter, all ties are assumed to be independent. (2) The first departure from tie independence assumes *dyadic independence*. Here, a tie from *i* to *j* is dependent on a tie from *j* to *i*, but independent of all other ties in the network. An ERGM with a density and a reciprocity parameter reflects dyadic independence. (3) The so-called *Markov Dependence* proposed by Frank and Strauss (1986) is more complex. Two ties are assumed to be independent, conditionally on the rest of the graph, unless they share a node. Transitive triplet and in-star or out-star effects in ERGMs satisfy Markov dependence. (4) The level of dependence that, since the publication of Snijders et al. (2006), is widely used in empirical analyses using ERGMs is *social circuit dependence*, introduced by Pattison and Robins (2002) and explained in Lusher et al. (2013, section 6.5.4). Two ties x_{ij} and x_{hl} are assumed to be conditionally independent, given the rest of the graph, if no tie between the dyad *ij* and the dyad *hl* exists. Four cycles and geometrically-weighed edgewise shared partner (GWESP) effects are examples of terms in ERGM specifications expressing circuit dependence.

Dependence assumptions have been very central in the literature about ERGMs since Frank and Strauss (1986), Wasserman and Pattison (1996), and in the later writings of Pattison and Robins. The central message of Snijders et al. (2006) was that the assumption of Markov dependence is too restrictive

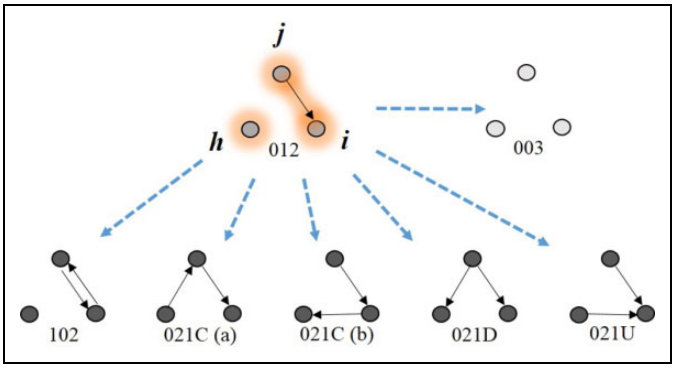

Figure 1. Possible transitions from triad 012.

in practice, and for modeling empirical networks, this should be relaxed at least to social circuit dependence.

Dependence assumptions have hardly been considered in the literature about SAOMs, however, with the exception of two earlier publications (Snijders and van Duijn 1997; Snijders 2005), in which independence and dyadic models were discussed. In this section, we show that the SAOM implies stronger dependence between ties for common model specifications. These dependence assumptions implicit in the SAOM are not yet properly defined and discussed in the literature. Tie dependence in the SAOM is discussed in detail on the example of a density and a reciprocity term. Toward the end of the section, we broaden our discussion to more complex effects.

Dependence in a Density Model

To discuss the dependence assumption in a model including only a density term, consider a network with only three nodes and the transitions depicted in Figure 1. Note that the probability of a transition in the ERGM is calculated as the product of the probabilities of (a) choosing the pair of actors at either side of the tie and (b) changing the tie variable; whereas for the SAOM, it is the product of the probabilities of (a) choosing the sending actor and (b) changing the tie variable. Given that the network, using the MAN notation[5] of the triad census (TC; Holland and Leinhardt 1976), is in state 012, we can calculate the transition probability to any adjacent graph depicted in Figure 1 according to the ERGM for a given θ and the SAOM for a given β.

For example, the transition probability from state 012 to state 102 in the ERGM is

$$p_{\mathrm{ERGM}}(012 \rightarrow 102) = \frac{1}{6} \times \frac{\exp(\theta_{\mathrm{density}})}{1 + \exp(\theta_{\mathrm{density}})}, \tag{6}$$

where the first fraction is the probability to consider the tie from i to j and the second fraction is the probability to create this tie, as opposed to leaving it absent (see equation [2]). At the same time,

$$\begin{aligned} p_{\mathrm{ERGM}}(012 \rightarrow 102) &= \frac{1}{2} \times p_{\mathrm{ERGM}}(012 \rightarrow 021C) \\ &= p_{\mathrm{ERGM}}(012 \rightarrow 021D) \\ &= p_{\mathrm{ERGM}}(012 \rightarrow 021U). \end{aligned} \tag{7}$$

The transition probability to state 021C is twice as high as the other three transitions, as the figure shows that there are two possible tie changes that result in this state.

In comparison, the transition probability from state 012 to state 102 for the SAOM that only includes a density parameter is given by:

$$p_{\mathrm{SAOM}}(012 \rightarrow 102) = \frac{1}{3} \times \frac{\exp(\beta_{\mathrm{density}})}{\exp(0) + \exp(\beta_{\mathrm{density}}) + \exp(\beta_{\mathrm{density}})}, \tag{8}$$

where the first fraction is the probability to choose actor i for the next tie change and the second fraction is the probability to form the tie to actor j (the numerator) rather than any other possible tie change, including no change (the denominator, see equation [3]). Further,

$$\begin{aligned} p_{\mathrm{SAOM}}(012 \rightarrow 102) &= \frac{1}{2} \times p_{\mathrm{SAOM}}(012 \rightarrow 021C) \\ &= p_{\mathrm{SAOM}}(012 \rightarrow 021U). \end{aligned} \tag{9}$$

However, the transition probability from state 012 to state 021D differs:

$$p_{\mathrm{SAOM}}(012 \rightarrow 021D) = \frac{1}{3} \times \frac{\exp(2 \times \beta_{\mathrm{density}})}{\exp(0) + \exp(\beta_{\mathrm{density}}) + \exp(2 \times \beta_{\mathrm{density}})}, \tag{10}$$

as the alternative to this transition for actor j is to *drop* a tie. This changes the denominator of the equation. Therefore, if $\beta_{\mathrm{density}} < 0$,

$$p_{\mathrm{SAOM}}(012 \rightarrow 102) > p_{\mathrm{SAOM}}(012 \rightarrow 021D). \tag{11}$$

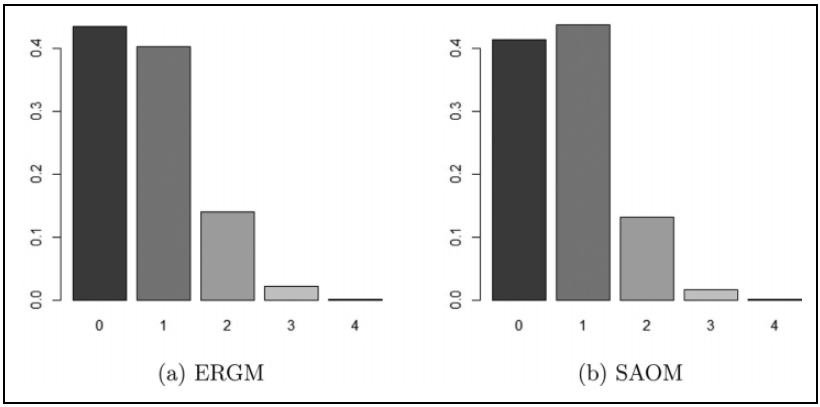

Figure 2. Outdegree distribution of a tie-only model for (a) exponential random graph model and (b) stochastic actor-oriented model with a density of 0.1879 and $n = 5$.

The previous equations show that in the SAOM, a negative density parameter makes the transition to state 021D less likely than a transition to any other state with two ties, because actor j has, when chosen to make a tie change, a "more attractive"[6] alternative for a tie change, which is dropping the already existing tie to actor i. Consequently, the creation of a tie from actor j to actor h is *dependent* on the existence of a tie from actor j to actor i in Figure 1. Thus, an SAOM that includes only a density parameter already assumes dependence between all ties that share the sender. We can refer to this type of dependence as *sender dependence.*[7]

In a second step, we relate these differences in transition probabilities to differences in the meso-level structures of a network consistent with either model. To this end, we calculate the stationary distribution of an ERGM and an SAOM with five nodes that only include a density parameter and with an expected density of $p = .1879$. This relates to the value $\beta_{\text{density}} = -1$ and an average degree of ≈ 0.75. For the ERGM, the parameter is $\theta_{\text{density}} = \log(p/(1-p)) = -1.464$. The corresponding ODs per actor can be calculated for either model. The OD of the ERGM follows a binomial distribution, as the presence or absence of each tie is determined independently with the same probability. For the SAOM, the approach of Procedure for SAOMs subsection can be used. The results are shown in Figure 2. Comparing the two distributions, it becomes apparent that for the SAOM, the likelihood to observe actors with an extreme value (0, 3, and 4) is lower than for the ERGM. The reason is that the more ties an actor has, the less

attractive it is in the SAOM framework to form additional ties, as the alternative options of dropping a tie becomes more attractive. In compensation, to achieve the same number of expected ties in the stationary distribution as the ERGM, the probability to form a tie when having an outdegree of 0 is increased in the SAOM. None of this holds for the ERGM, in which ties of an actor are independent.

Analytical solution for the density model. While calculating the stationary distribution for an SAOM as a function of only graph statistics is in practice not tractable for almost all commonly used model specifications (see The Model Definitions subsection), the density-only model provides, due to its simplicity, an exception. Thus, we can analytically compare the ERGM and SAOM for the density model to confirm and theoretically prove the results obtained by simulation.

The simplest ERGM and SAOM specification has $K = 1$ (so that we can drop the index k) and includes only the degree, or outdegree term: $z(x) = \sum_{ij} x_{ij}$ for the ERGM and $s_i(x) = \sum_j x_{ij}$ for the SAOM. The total number of ties in the ERGM then has the binomial distribution with parameters $N(N-1)$ and $p = e^{\beta}/(1+e^{\beta})$. For the SAOM, the explicit calculation of the stationary distribution is still possible. The definition of the model implies that the rows of the adjacency matrix must be independent (i.e., *sender dependence*), so we can focus on the distribution of a single row, that is, all outgoing ties of a single actor. This model is invariant under permuting the actors, so that the distribution of this row must depend only on the number of ties. Denote this sufficient statistic by T. Equation (3) implies for T the process with transition probabilities:

$$\begin{aligned} p_{\mathrm{SAOM}}(t \to t+1; \beta) &= \frac{(N-t-1)e^{\beta}}{1+(N-t-1)e^{\beta}+te^{-\beta}} \\ p_{\mathrm{SAOM}}(t \to t-1; \beta) &= \frac{te^{-\beta}}{1+(N-t-1)e^{\beta}+te^{-\beta}}. \end{aligned} \tag{12}$$

Denote the stationary probabilities by $\pi(t) = P\{T = t\}$. A sufficient condition for determining the values of $\pi(t)$ is the so-called detailed balance equation (Norris 1997)

$$\pi(t)p_{\mathrm{SAOM}}(t \to t+1; \beta) = \pi(t+1)p_{\mathrm{SAOM}}(t+1 \to t; \beta) \quad \text{for all } t = 0, \ldots, N-2,$$

and it is clear that this equation will indeed always have a solution because it can be used to recursively calculate $\pi(t)/\pi(0)$ for $t = 1, \ldots, N-2$, and

finally determine $\pi(0)$ by the requirement that the probabilities sum to 1. With some further calculations, this shows that

$$\frac{\pi(t+1)}{\pi(t)} = e^{2\beta}\frac{N-t-1}{t+1}\left\{1+\frac{e^{-\beta}-e^{\beta}}{1+(N-t-1)e^{\beta}+te^{-\beta}}\right\}. \quad (13)$$

This now allows direct calculation of the expected distribution of number of ties for each actor for any given β. For $\beta = -1$, this is indeed the OD shown in Figure 2 on the right-hand side.

The preceding discussion also allows us to understand why the likelihood to observe a specific realization of a network x under an SAOM cannot be expressed merely as a function of graph-level statistics, as mentioned in the second section. From equation (13), it follows that the likelihood to observe the 021D configuration is different from, for example, the 021U configuration in an SAOM with only a density parameter. Thus, two graphs that have the same global graph statistics (number of ties) have a different likelihood of being observed. The probability to observe a specific graph is, therefore, a function not only of graph statistics (as it is in the ERGM) but also of how these configurations are nested within actors.

Consequences for empirical research. The seemingly trivial difference in the ODs for the two models studied above has implications for empirical research regarding estimated outdegree centralization parameters of empirical data. In simple cases, the OD can be modeled by the out-star or outdegree activity effects in the ERGM or SAOM, respectively. For the example presented above, additional analysis has shown that the expected OD of the SAOM including only a density parameter in Figure 2 can be closely approximated by an ERGM that includes a density and a negative out-star parameter. In turn, this means when simulating a network with an ERGM that includes only a density parameter (a Bernoulli graph), estimating this network with an SAOM and a model specification that includes an outdegree activity effect will tend to lead to finding a positive parameter for the latter effect (whether it will be large or significant is a different matter).

Therefore, when analyzing the same empirical network with an ERGM and an SAOM under a more complicated model, we can expect similar differences. It is possible that one of the two analyses finds evidence for (or against) outdegree centralization, while the other one does not. This is a consequence of the difference between the two null models (the density-only model). However, this does not imply any problems of identifiability of different effects in one or the other model. Also in the SAOM, there is no problem in

distinguishing outdegree centralization effects from the null model, even though some dependence between outgoing ties is generally present in the SAOM. Rather, the point in observed networks where the parameter is zero is different between the two models. We should note, however, that these differences are quite small, and vanishingly small for large N.

Dependence in a Reciprocity Model

After discussing the consequences arising from a tie-based or actor-based perspective for a pure density model, we turn to a very basic structural network mechanism: reciprocity. There are few social networks that do not exhibit a tendency toward reciprocation; it is often the strongest predictor of the presence of a tie. In this section, we show that including a reciprocity parameter in the SAOM leads to a higher level of tie dependence than in the ERGM.

To this end, consider Figure 1 once again. We now look at the transition probabilities from state 012 to the states 021C(a) and 021C(b). Note that for neither transition, a reciprocated tie is formed. For the ERGM including a density and a reciprocity parameter, the transition probabilities are

$$\begin{aligned} p_{\mathrm{ERGM}}(012 \to 021C(a)) &= p_{\mathrm{ERGM}}(012 \to 021C(b)) \\ &= \frac{1}{6} \times \frac{\exp(1 \times \theta_{\mathrm{density}} + 0 \times \theta_{\mathrm{reciprocity}})}{1 + \exp(1 \times \theta_{\mathrm{density}} + 0 \times \theta_{\mathrm{reciprocity}})} \\ &= \frac{1}{6} \times \frac{\exp(\theta_{\mathrm{density}})}{1 + \exp(\theta_{\mathrm{density}})}. \end{aligned} \tag{14}$$

Thus, not only are the transition probabilities identical, but they are also independent of $\theta_{\mathrm{reciprocity}}$. This is intuitively plausible, as neither newly created tie is reciprocated. For the SAOM, on the other hand, the transition probabilities are

$$p_{\mathrm{SAOM}}(012 \to 021C(a)) = \frac{1}{3} \times \frac{\exp(\beta_{\mathrm{den}})}{\exp(0) + \exp(\beta_{\mathrm{den}}) + \exp(\beta_{\mathrm{den}})} \tag{15}$$

and

$$p_{\mathrm{SAOM}}(012 \to 021C(b)) = \frac{1}{3} \times \frac{\exp(\beta_{\mathrm{den}})}{\exp(0) + \exp(\beta_{\mathrm{den}}) + \exp(\beta_{\mathrm{den}} + \beta_{\mathrm{rec}})}. \tag{16}$$

Note that the transition probability to state 021C(b) includes a reciprocity parameter in the denominator. Consequently, if $\beta_{\text{reciprocity}} > 0$,

$$p_{\text{SAOM}}(012 \rightarrow 021C(a)) > p_{\text{SAOM}}(012 \rightarrow 021C(b)). \quad (17)$$

Intuitively, this means that actor i is less likely to form a two path by connecting to actor h, because there is a more attractive alternative, namely, to reciprocate the tie to actor j. Actor h, on the other hand, does not have a more attractive alternative, thus is relatively more likely to create a two path by connecting to actor j. From this simple example, we can conclude that using a reciprocity effect in an SAOM implies additional dependence compared to the ERGM: The transition probabilities for all outgoing ties of an actor are dependent on all incoming ties. We can refer to this type of dependence as *Markovian transition dependence*.

These differences in tie dependence between the ERGM and SAOM allow us to extrapolate expectations about the resulting macro-level consequences. We expect that in the SAOM, the number of two paths in the network depends on the strength of the reciprocity parameter. For a negative parameter, every incoming tie increases the probability of an actor to form a nonreciprocated tie (as the alternative options become less attractive). In turn, there are more two paths in this network compared to a network without a tendency against reciprocity. For a positive reciprocity parameter, we expect the opposite, that is, a decrease in number of two paths. For the ERGM, there is no reason to expect a difference in the number of two paths, as the likelihood of forming a nonreciprocated tie is not affected by the number of incoming ties an actor has—after all, the model is tie oriented, not actor oriented.

We show this numerically by calculating the stationary distribution of a series of ERGMs and SAOMs that include a density and a reciprocity parameter for a network with five nodes. Now we have to rely exclusively on the simulation approach for the SAOM. The expected number of ties in each stationary distribution is eight, while the expected number of reciprocated ties was varied between 1 and 6.5. In total, we calculated the stationary distribution of 12 ERGMs and 12 SAOMs. For each of the stationary distributions, the expected number of two paths was extracted. The results are shown in Figure 3. For the ERGM, the number of two paths is independent of the number of reciprocated ties, following dyadic independence (blue line and triangles). For the SAOM, on the other hand, a negative reciprocity parameter leads to an increase in two paths, while a positive reciprocity parameter is associated with a decrease in the number of two paths (red line

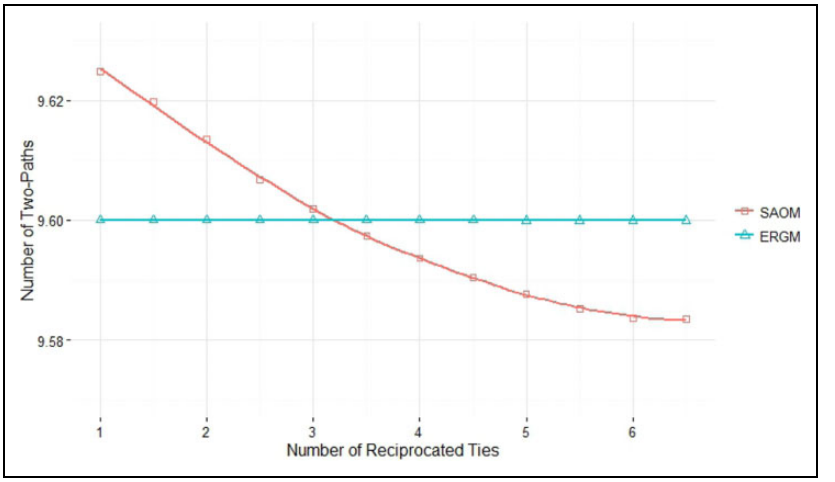

Figure 3. Expected number of two paths in a model with fixed density (0.4) by expected number of reciprocated ties for $n = 5$.

and squares). This is in line with our previous reasoning, although the differences are very small in size.

Consequences for empirical research. The consequences for researchers that analyze empirical networks are similar to the ones outlined in the subsection on empirical consequences for the density model. The difference in observed two paths between the two models that include a density and a reciprocity parameter can be compensated by including a two-path parameter in the ERGM or an indegree activity or outdegree popularity parameter in the SAOM. However, for empirical researchers, this means that when analyzing the same network with an ERGM and an SAOM, it might be the case that only one of the two finds evidence for or against a tendency to form two paths.

Higher-Level Dependence

The pattern that we observed for the two simplest endogenous network mechanisms, density and reciprocity, can be summarized as follows. In the ERGM, the conditional probability of presence or absence of a tie, given the rest of the graph, is only influenced by how this changes the modeled statistics of the entire graph. In the SAOM, the conditional probability of a tie is also determined by the other available options that are possible for the actor sending the tie, and how these are reflected in differences in the graph

statistics. The ERGM is based on choices with binary option sets, the SAOM on choices with option sets having N elements. This induces in the SAOM additional dependencies between ties originating from the same sender that do not exist in the ERGM. The outlined logic can be applied to all endogenous network effects, and the dependence assumption in the SOAM can be specified accordingly. While we will not engage in a deep discussion and mathematical proofs for the dependence of the following effects, the intuition should be clear from the preceding sections.

We first look at transitivity, one of the most common effects that accounts for the tendency to form cohesive groups within the network. In the ERGM, Frank and Strauss (1986) proposed to model transitivity using Markov dependence, as other ties of the sender and recipient influence the likelihood of the tie in question. Snijders et al. (2006) proposed that, rather, social circuit dependence be used to model transitivity.[8] We use the Markovian formulation of transitivity for our discussion; however, an analogous argument can be made for transitivity using social circuit dependence.

In the SAOM, actor i considers every potential tie change it could make, given the rest of the graph. Consequently, the presence of the two-path $i \rightarrow h \rightarrow j$ makes the option to form a tie to j more likely, while at the same time, the alternative option to form a tie to, say, actor k becomes less likely. Thus, the likelihood of $i \rightarrow k$ depends on the presence of the tie $h \rightarrow j$, given that the tie $i \rightarrow h$ exists. Consequently, an SAOM including a transitivity effect assumes dependence on the configuration of outgoing ties of the focal actor at distance two. This is a transition version of the D_1 class in the dependence hierarchy of Pattison and Snijders (2013): at least one of the nodes i, k is at distance 1 of at least one of the nodes h, j. We may call this *distance-two transition dependence*, or *local transition dependence*, referring to the distance-two neighborhood as the local neighborhood. Schematically, this is sketched in Figure 4. The red bold tie on the left is dependent on the green dotted tie on the very right for the mentioned effect in the SAOM.

As a last example, consider the indegree popularity effect in the SAOM, which is the equivalent to the in-star effect in the ERGM. This models preferential attachment. In the ERGM, the in-star effect still satisfies Markov dependence. In the SAOM, there is a completely new level of dependence. With a positive indegree popularity parameter, the formation of a tie from i to j becomes more likely with every incoming tie to actor j, for example, $h \rightarrow j$. As a tie to actor j becomes more and more attractive to actor i, the probability that i forms a tie to an alternative alter, for example, k decreases. Therefore, the tie $i \rightarrow k$ is dependent on the presence of the tie $h \rightarrow j$, *even if* there is no

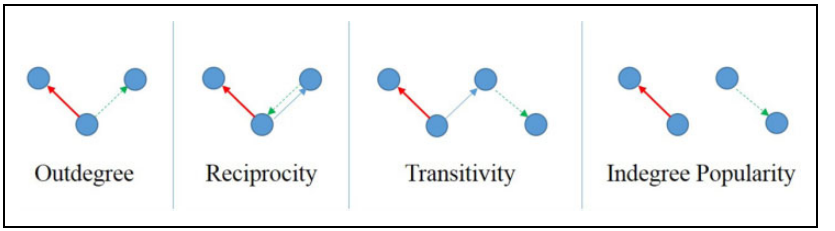

Figure 4. Dependence of the red tie on the green tie for different stochastic actor-oriented model effects.

tie $i \rightarrow h$. Thus, dependence is not local any more, but *nonlocal transition dependence* is assumed, as depicted in Figure 4.

Consequences for empirical research. While it was relatively simple to understand how the dependence induced by the density and reciprocity effects results in straightforward differences at the meso- and macro-level of the network, it is difficult to formulate the meso- and macro-level outcomes of the last two dependence assumptions. The dependence in the transitivity effect might lead to changes in average path lengths in the network. We abstain from further speculation.

Nevertheless, this section has outlined how the different dependence assumptions in the ERGM and the SAOM can lead to different findings when analyzing empirical networks. The choice of method to analyze a network should thus not be a matter of taste for a model or software, but resulting from a careful consideration whether the tie formation process is rather in line with a tie-oriented or actor-oriented framework. A basic contrast is that in the ERGM, potential ties are considered by themselves whereas in the SAOM, outgoing potential ties from any given actor are in the same basket, with a trade-off between them; in both cases, given their further embeddedness in the network.

Asymmetric Transition Dependence and Elementary Effects

While the previous section was concerned with the different dependence assumptions of the ERGM and the SAOM, the present section treats the phenomenon of *asymmetric transition dependence*. Asymmetric transition dependence describes a situation in which the toggle (change) of the tie $i \rightarrow j$ is, by the nature of the effect formulation, dependent on a tie $h \rightarrow l$, but the

toggle of $h \rightarrow l$ is *not* dependent on $i \rightarrow j$. Effects that model asymmetric transition dependence are not only possible in the SAOM but also frequently used. For the ERGM, on the other hand, asymmetric transition dependence is impossible, by virtue of our formulation of the dynamic model where toggles are made according to the conditional distribution of the tie given the rest of the network. The situation for ERGM and SAOM are different because transitions in the SAOM take the position of the focal actor i into account, while in the ERGM, there is no focal actor. The resulting meso- and macro-level differences are considerable, and for empirical researchers, they may be more severe than the differences discussed in the section on dependence between ties. We will discuss asymmetric transition dependence using the example of transitivity.

Regardless of whether the micro- or macro-level formulation (according to equations [2] or [1], respectively) of the ERGM is used, the likelihood of a graph only depends on the count of structures $z(x)$, such as transitive triplets. The transition probability between two adjacent graphs depends on the change of the substructure count $\Delta z(x)$, the *change statistic*. This means that the position of the changing tie in the substructure is irrelevant, as long as it contributes to the count. For the transitive triplet consisting of the three ties $i \rightarrow h$, $i \rightarrow j$, and $h \rightarrow j$, the formation of each tie contributes to the change statistic $\Delta z(x)$, given the other two ties exist.

In contrast, the SAOM is actor-based and defined by transitions following equation (3), that is, the probability of transitions between two adjacent graphs depends on the structures $s_i(x)$ *from the perspective of actor i*. The change statistics $\Delta s_{k,i}(x)$ take the position of the focal actor into account—only the ties formed by the focal actor that create the substructure from the perspective of the focal actor contribute to the change statistic. For the transitive triplet with ties $i \rightarrow h$, $i \rightarrow j$, and $h \rightarrow j$, only the formation or termination of the ties $i \rightarrow h$ and $i \rightarrow j$ contribute to the change statistics, *not* the tie $h \rightarrow j$. In an SAOM with a positive transitivity parameter, forming a transitive triplet by closing an out-two-star does not contribute to a higher score in the numerator of equation (3).

This difference is illustrated in Figure 5. The green arrows show the transitions that contribute to the change statistic for the transitivity effect in the ERGM, while the red ones show the transitions that contribute in the SAOM. On this example, the notion of asymmetric transition dependence in the SAOM should become clear: The toggle of the tie from the lower left to the lower right actor is dependent on the tie from the upper to the lower right actor, but not vice versa. On the one hand, allowing for asymmetric transition dependence gives the SAOM a degree of flexibility that enables the

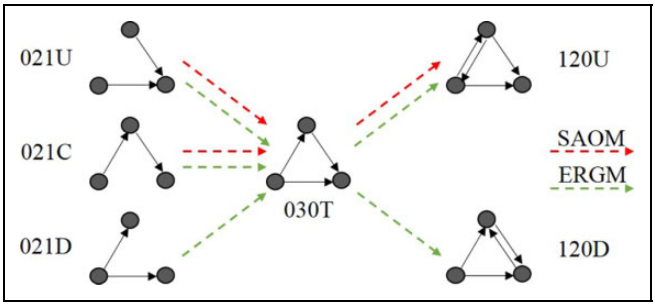

Figure 5. Transitions that contribute to the change statistic for the transitive triplets effect.

formulation of effects that are very close to social theory. We might assume that the lower left actor *i* in the transitivity example has a strong desire to be in a situation where a friend of a friend is its friend. For the upper actor *h*, however, whether himself and a friend are both liked by the same person might not matter. On the other hand, the flexibility to specify asymmetric effects comes at a cost of losing parsimony of the statistical model, as exemplified by comparing equations (1) and (4).

The possibility to specify effects that model a defined network evolution mechanism of theoretical interest can be made even more precise in the SAOM than outlined in the example above, by so-called *elementary effects*, which in 2014 were defined in the RSiena manual (Ripley et al. 2016). For elementary effects, only one "focal" tie in a substructure contributes to the change statistic. For example, the closing of a two path to form a transitive triplet can be specified as an elementary effect, that is, only the transition from 021C to 030T in Figure 5 contributes to the change statistic. This can be advantageous, if it is theoretically assumed that the tendency to close a two path or the tendency to close an out-two-star might differ and reflect distinct underlying motivations of actors. The RSiena software contains three effects that refer specifically each to the creation of one of the ties in the transitive triplet. Two of these are elementary effects.

Differences in Meso-Level Structures

Asymmetric transition dependence poses the question how differences in effect specification impact the meso- and macro-level characteristics of a network. We use the transitive triplets effect as an example once again. A fruitful approach to these differences starts at comparing the structures that precede the transitive triplet. For the ERGM micro-model, transitive triplets are evenly created from

the triads 021D, 021C, and 021U as antecedents. In the SAOM, 030T triads that add to the change statistics are created only from the 021U and 021C triads. Consequently, if we look at the stationary distribution of an ERGM and an SAOM with a positive transitivity parameter and the same global number of transitive triplets, we expect to observe more 021D triads in the SAOM than in the ERGM, because in the ERGM, it tends to close to the 030T triad. At the same time, we expect more 021U triads in the ERGM than in the SAOM, as in the SAOM there is a higher likelihood of closing this triad to counterbalance the smaller number of preceding structures for the 030T triad. The same argument can be made for the transition from the 030T triad (the transitive triplet) to the 120D and 120U triad. Only the transition to the 120U triad contributes to the change statistic of the SAOM, while both contribute to the change statistic of the ERGM. Thus, we expect to observe more 120D triads in the stationary distribution of the ERGM and more 120U triads in the SAOM.

To test these expectations, we performed a similar analysis as in the subsection on dependence in a reciprocity model. We calculate the stationary distribution for a set of ERGMs and SAOMs with five nodes and nine expected ties, while the expected number of transitive triplets varies between 2.5 and 11, as outlined in Procedure for ERGMs and Procedure for SAOMs subsections. In total, we calculated 18 stationary distributions for either model type. We retrieved the expected number of triads 021D, 021U, 120D, and 120U from the calculated stationary distribution and analyzed how they change dependent on the number of transitive triplets in the network. The results are shown in Figure 6.

Figure 6a shows that for a higher number of transitive triplets associated with a positive transitivity parameter, in the SAOM, indeed more expected 021D triads in the stationary distribution are observed than in the ERGM. For lower numbers of transitive triplets, associated with a negative transitivity parameter, the pattern is reversed. The further expectations, outlined in the previous paragraph, are equally met. For networks with a positive transitivity parameter, we find a higher expected number of 021U and 120D triads in the ERGM, while the expected number of 120U triads is higher in the stationary distribution of the SAOM. Additional analysis shows that the different effect specifications lead to differences in the relative expected frequency for most of the 16 nonisomorphic triads between the two models, where some differences can be explained more intuitively than others.

Consequences for Empirical Research

As the previous sections have shown, the transitivity effect models distinct processes in the ERGM compared to the SAOM. Thus, if two seemingly

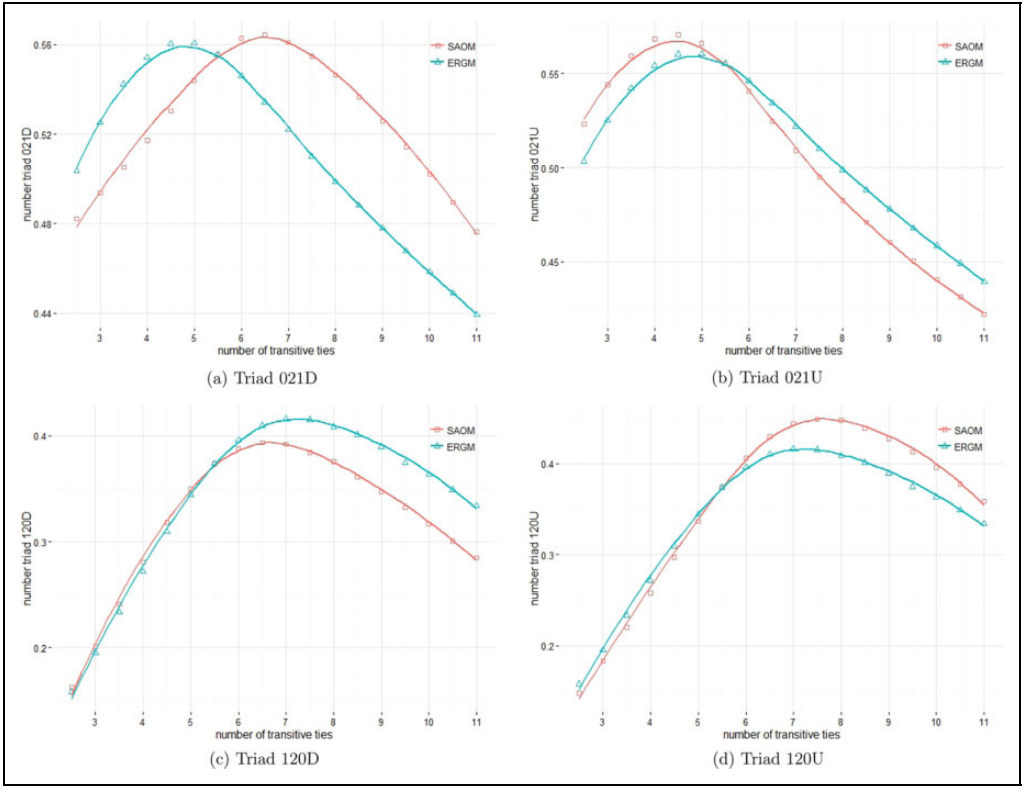

Figure 6. Various triads for exponential random graph model and stochastic actor-oriented model with $n = 5$ and an expected density of 0.45 for different number of transitive triplets.

equivalent models do not model the same process, as the underlying generative mechanisms differ, they imply a different interpretation of the estimated parameters. In other words, a positive transitivity parameter means something else in the ERGM compared to the SAOM. Using the ERGM, it can be inferred that ties in transitive structures are more likely to exist, while the SAOM suggests that ties that close otherwise open in-two-stars or two paths are more likely to exist.

This seemingly obvious difference has important implications. First, indegree and outdegree centralization that is naturally represented by a model including a transitivity parameter differs. This is clear from the previous discussion and Figure 6a and b.[9] Similar to the consequences discussed in Dependence in a Density Model and Dependence in a Reciprocity Model subsections, varying indegree and outdegree centralization can lead to diverging findings for the respective model parameters in the ERGM, the in-star and out-star parameter, compared to the SAOM, the indegree popularity and outdegree activity parameter.

The second consequence of the option of asymmetric transition dependence in the SAOM relates to model selection. Parameters included in a model are often, at least partly, based on the fit of model to data, as outlined in Hunter, Goodreau, and Handcock (2008) and Lospinoso (2012), in particular the fit of the triad census (TC) (Block 2015). As Differences in Meso-level Structures subsection has shown, the expected counts in the TC can vary in an analysis employing an ERGM and an SAOM with "equivalent" model specifications. Consequently, the fit of the model to the data regarding the TC in the ERGM might be good, while the fit for the same model is poor when using an SAOM. In this case, the model specification for the SAOM, but not the ERGM, will be changed.

Employing different model specifications often means different control effects, included in addition to effects of theoretical interest. Still, the chosen control effects will contribute to the explanations. Thus, if two different model specifications in an analysis with an ERGM and an SAOM are employed, the difference in control variables might further impact the substantive conclusions.

A further consequence of different model specifications being preferable in the ERGM compared to the SAOM implies that impartial model comparison in terms of model *performance* becomes difficult. A common strategy in model comparison is to estimate two equivalently specified models and see which one performs better with respect to some criterion. As we have seen, an ERGM and an SAOM that include effects that try to capture the same concept do not model exactly the same process: There are no truly equivalent

model specifications. Consequently, the premise that we can compare the two models by including the same effects falls apart, and comparing the empirical performance of an ERGM to an SAOM with identical model specification has restricted validity.

Asymmetric transition dependence is relevant beyond the discussed example of transitivity. Other, frequently used effects in the SAOM model asymmetric transition dependence, such as the indegree activity, the outdegree popularity, and most other triadic effects (e.g., the transitive reciprocated triplets effect). For all of these effects, the modeled network structures have an equivalent effect in the ERGM, but one ERGM effect will correspond to several SAOM effects. For example, the indegree activity and outdegree popularity effects in the SAOM both correspond to the two-path effect in the ERGM. This is because different dynamic microprocesses can lead to the same cross-sectional consequences. The theoretical relevance of asymmetric transition dependence also points to the importance of dynamic theoretical approaches and longitudinal designs.

Identifiability of Asymmetric Transition Dependence in Cross-Sectional Data

A question that arises naturally from the discussed asymmetric transition dependence is whether it is identifiable in cross-sectional data. In longitudinal data, the additional time dimension allows identifying asymmetric transition dependence. However, in RSiena, this is not allowed by the default estimation (Method of Moments), but rather maximum likelihood estimation is necessary for this purpose.

The discussion in the preceding subsections shows that the different parameters available by asymmetric transition dependence lead to differences in triad frequencies. Therefore, it must be possible to empirically identify these parameters in the SAOM. In the ERGM, although asymmetric transition dependence is impossible, nevertheless parameters for the different triads can be included and estimated. One basic definition of the ERGM is based on the elaboration of conditional independence assumptions, utilizing the Hammersley–Clifford theorem (Besag 1974), which states that in the ERGM for each clique in the dependence graph (e.g., corresponding to a transitive triplet), there is only one parameter.[10] This is discussed in Lusher et al. (2013), and figure 7.8 in their section 7.2 shows the 15 two- and three-node sufficient subgraphs that can be used to parameterize triadic dependencies in directed ERGMs. Since there are 16 triads in a directed graph (Holland and Leinhardt 1976), this shows that—although it is not usual—the

ERGM methodology permits in principle to identify all differential triad frequencies in directed graphs. However, identifiability both for the ERGM and for the SAOM is only possible in principle and to date not implemented in any software. Furthermore, often scientific interest is in processes and these can convincingly be studied only with a longitudinal design, because in many practical cases, the assumption of a stationary distribution will be debatable. More research on this is needed.

Model Fitting Exercise

All analyses in the previous sections illustrating the conceptual and theoretical arguments were conducted on a five-node network, and each effect was studied in isolation. What is of practical relevance to an empirical researcher, however, is whether and how these differences actually impact the conclusions drawn when applied to real data.

The complex nature of the different effects and their interdependencies makes it difficult to predict, let alone understand how they influence the results of an empirical analysis of a real-world network, when included all at once. Thus, we intend to show in this section not analytically, but by the fitting a model that seems reasonable from previous research to the same empirical network data using an ERGM and an SAOM to illustrate how the previously discussed differences impact conclusions drawn from either model.

Data and Implementation

The data that are used for this comparison is a friendship network between adolescents, which was collected within the context of the *Teenage Friends and Lifestyle Study*[11] (Pearson and Michell 2000). The data were gathered in Glasgow between 1995 and 1997. All 160 participants of the study were members of a school cohort (aged $\sim$13 in the first wave) and were followed over a period of three years. Complete data are available for 129 adolescents, who participated in all three waves. Demographic and lifestyle characteristics of the adolescents were collected; furthermore, participants could nominate up to six persons in their year group as close friends at each time point. For the empirical analysis of this study, we use the friendship network in the third wave and analyze this using an ERGM and a cross-sectional SAOM.

Analysis of the network with an SAOM was conducted using RSiena (Ripley et al. 2016). The identical network was specified as the first and second time point in the analysis, and the basic rate parameter was fixed at a

very high value to approximate a stationary distribution.[12] Overall, the procedure outlined in Snijders and Steglich (2015) was followed for the analysis of cross-sectional data with RSiena. The ERGM analysis of the empirical network was conducted using MPNet (Wang et al. 2014).

Model Selection Strategy

For the empirical comparison, we select a specification that is reasonable on theoretical grounds, known from prior research, as well as provides a satisfying fit for either model. To determine model fit, we use the procedure outlined in Hunter et al. (2008) and Lospinoso (2012).

The final model that fulfills both criteria includes the basic density and reciprocity parameter, three triadic parameters, three degree-related parameters and one covariate-based parameter. The three triadic parameters are the geometrically weighted version of transitivity over multiple two paths (GWESP FF) and transitivity over multiple in-stars (GWESP FB), as discussed in, for example, Ripley et al. (2016) and Wang et al. (2014), as well as a dense triad parameter, discussed in Block (2015). The degree-related parameters are the outdegree activity effect in the SAOM, or the respective out two-star effect in the ERGM, an out-isolate parameter, modeling the tendency toward or against having isolates in the network, and the outdegree popularity effect in the SAOM, or the respective two-path effect in the ERGM. The often included parameter modeling indegree centralization (indegree popularity or in two-star) was omitted, as it was not significant in either model and led to large collinearity. All aforementioned effects are discussed in standard text about ERGMs and SAOMs. Finally, a parameter modeling gender homophily was included, with a centered gender variable.

Results

The results of the estimated model can be found in Table 2. The fit of both models is satisfactory with regard to the OD, the indegree distribution, and the TC.[13] Some important conclusions do not differ regardless of which method is used. Both estimations find a tendency toward reciprocation and sex homophily, but against the formation of dense triads and two paths. However, and more importantly, other conclusions differ. First, both models find a tendency toward clustering, but through different paths. In the ERGM, clustering is represented best by the GWESP FB configuration, while in the SAOM it is modeled by the GWESP FF effect. A naive interpretation of this difference in clustering effects would be that in the ERGM, the important

Table 2. Estimates and s.e.'s for the Glasgow Data Wave 3 SAOM and ERGM.

SAOM			ERGM		
	Estimates	s.e.	Estimates	s.e.	
Outdegree	−1.21	0.35*	−3.87	0.20*	Density
Reciprocity	2.80	0.35*	3.22	0.25*	Reciprocity
GWESP FF	2.27	0.47*	0.39	0.27	GWESP FF
GWESP FB	−0.26	0.46	1.21	0.25*	GWESP FB
Dense triads	−0.31	0.08*	−0.80	0.24*	Dense triads
Outdegree activity	−0.17	0.04*	−0.04	0.06	Out-two-star
Out isolate	−2.23	1.12*	−0.70	0.58	Out isolate
Outdegree popularity	−0.22	0.04*	−0.23	0.03*	Two-path
Sex Ego × Alter	0.77	0.13*	1.25	0.17*	Sex Ego × Alter
GoF OD	0.250		0.074		GoF OD
GoF ID	0.636		0.655		GoF ID
GoF TC	0.081		0.683		GoF TC

Note: Values for GoF test represent *p* values testing whether the model does not fit the data with regards to the tested statistic. ERGM = exponential random graph model; SAOM = stochastic actor-oriented model; GoF = goodness of fit; OD = outdegree distribution; ID = indegree distribution; TC = triad census; s.e. = standard error.
*$p < .05$.

closure mechanisms is the closure of in-two-stars, while the important closure mechanism in the SAOM is the closure of two paths. The further differences regard the effects that model the outdegree of nodes in the network. In the SAOM, there is a tendency against outdegree centralization and a tendency against out isolates. Both are not found in the ERGM.

What we learn from this model fitting exercise is that the differences discussed in the previous sections indeed change conclusions that can be drawn. We should keep in mind, however, that this is only one case and more research is needed to establish potential patterns of differences between estimated parameters.

Discussion and Conclusion

Currently, there are two dominant statistical approaches that model the micro-mechanisms underlying the formation of social networks, the ERGM and the SAOM. While both models are widely applied in empirical research, their main similarities and differences are not well understood to date. We have tried to highlight the differences in fundamental principles underlying the tie-oriented and actor-oriented approach, relating to the different meso-

and macro-level network features that can be represented by either model. The discussed differences hold irrespective of whether cross-sectional or longitudinal data is analyzed, but we have positioned our presentation in the framework of cross-sectional modeling. For the SAOM, this is rather unusual, and we have considered the stationary distribution for the SAOM (Snijders and Steglich 2015) as its cross-sectional version.

We identify two main distinguishing features that relate to the form that dependence between ties can take in either model: First, there are intrinsic model differences in tie dependence that arise from using a multinomial (actor oriented), rather than binary (tie oriented) choice function in the SAOM compared to the ERGM. Second, the option to specify asymmetric transition dependence between ties exists in the SAOM, but not the ERGM. This stems from the former being defined on the level of transitions between adjacent graphs, while the latter is defined globally, using counts of substructures in the network. We illustrated both differences analytically as well as showing an example where these intrinsic model differences have an effect on conclusions when analyzing empirical data.

Advice for Model Selection

Given that employing different models can lead to differing conclusions about the micro-mechanisms that bring about empirically observed networks, the question arises which model should be used for which type of research question and data. The choice of method for an empirical analysis can be based on at least two considerations: Which method provides a better fit of model to data, and which model assumptions reflect the assumed social processes that give rise to the observed network more accurately.

Generally, the goal of network modeling is the *explanation* of the micro-mechanisms that bring about a social network. In this case, the statistical model should be a reasonable representation of the processes that are assumed to take place in the real world, that is, the integration of statistical and theoretical model is desirable. Hence, for most social science applications that try to explain observed networks, the method should be preferred whose assumptions represent most accurately the real-world process that has brought the network about. A basic criterion is whether the process leading to the existence of ties involves in some sense a comparison between the various ties that potentially could be "sent" by a given node. For networks where the nodes are social actors who decide in some way about their collection of outgoing ties, this relates to the agency of the nodes sending the ties, and the considerations that play a role in the choices made by these nodes or actors.

First, assuming that ties are chosen by actors in a way that can be expressed using rewards and costs, is it reasonable to assume that actors compare the potential ties with one another with respect to costs and rewards and have a greater likelihood of choosing the relatively more rewarding ties, or do actors consider the costs and rewards of each tie separately, without a mutual comparison? The former case is in accordance with the multinomial formulation of the SAOM, while the latter is in line with the binary ERGM. In case the formation of ties is costly due to constraints in resources, the choice for one tie implies the choice against another tie. This holds, for example, for friendship ties. Friends require investments such as time, of which people only have a limited stock; therefore, having an unlimited number of friends is not possible. Every choice to be friends with one person consequently lowers the probability to nominate a different person as a friend, reflecting the dependence between all outgoing ties present in the SAOM. This suggests the SAOM as the more appropriate model for this kind of tie.

For costless ties, the opposite is the case. If an individual can have, in principle, an unlimited amount of ties, ties will not be evaluated in competition to one another, but the existence of each tie can be considered by itself. This might be the case, for example, for ties of positive affection, or "liking ties." There is no constraint to how many people an individual can like (as opposed to how many one can be friends with). Here the binary formulation of the ERGM may be more appropriate. Thus, a first rule of thumb for model selection between the ERGM and SAOM could be whether the nature of the tie is that actors cannot have them in unlimited amounts.

Second, in modeled substructures such as triplets, is there one particular sending node that has a crucial theoretical importance in this structure and can be considered the actor who (in first approximation) has "control" over the tie? Many social science theories are specified from an individual's position. For example, a friend of a friend will become my friend; this verbal claim specifies the "ego" who will form which tie to whom within a given structure. This can be translated into one specific transition that leads from a two path (the 012C triad) to a transitive triplet (the 030T triad). In case the network model is employed to test such specific, actor-based theories, an actor-oriented model that can take these asymmetric transition dependencies between ties into account is closer to the assumed processes. In other words, actor-based hypotheses and theories are more directly tested and modeled with the SAOM. The prime example in which asymmetric transition dependence is common in theoretical reasoning is the friendship network.

As a contrasting example, when analyzing a network of flight connections between airports, the airlines and the airports do have agency, but there is not

a priori special position for the origin airport as compared to the destination, and the ERGM might better reflect the assumed process of network formation. Thus, a second rule for model selection is whether the theories to be tested are formulated on the actor or tie level, for which asymmetric transition dependence is a good indicator.

Third, we note the crucial role played by conditional independence assumptions for defining the ERGM. If the theories investigated imply clear consequences, or research questions, about conditional dependence and independence, a modeling approach based on the ERGM may give a better link between theory and empirical investigation.

The previous paragraphs indicate that a good knowledge of the context, data, and the assumed processes leading to the formation and dissolution of ties will be helpful for deciding which model to use. We think that in many cases, it will be possible to make an a priori model selection based on the outlined considerations. If this is not the case, a fit-based model selection might be able to show which model can more naturally, and more parsimoniously, represent observed features of a network.

Finally, from the perspective of statistical theory, it should be noted that some properties of the ERGM as a statistical model may be better understood than properties of the SAOM. This is because there is a vast literature on exponential family models (of which the ERGM is a member), while less statistical theory is available for the SAOM. Nevertheless, we believe that theoretical and substantive considerations should have the primacy over statistical fit in model selection.

Outlook

We discussed the principal differences between the ERGM and the SAOM. However, in the empirical comparisons, this included only one case study so far, with a comparison based on an admittedly arbitrary model. We do not know how large the differences will generally be between ERGM and SAOM analyses using apparently the same model specifications. A more encompassing picture of the consequences for empirical data analysis would be desirable, resulting from comparing multiple cases from different disciplines and finding common differences in parameter estimates.

A series of questions that remained unanswered in this article are the differences between the variants and methods for longitudinal network analysis. Especially for the ERGM, multiple extensions have been proposed that enable the analysis of longitudinal data. How they compare to one another and the differences to the SAOM of either ERGM modification is not yet

studied. The main differences that we foresee are the differences between continuous time models, such as the lERGM (Snijders and Koskinen 2013) and the SAOM, on the one hand, and discrete time, or auto-regressive models, such as the tERGM (Hanneke et al. 2010) and the StERGM (Krivitsky and Handcock 2014), on the other. A further difference is that the SAOM allows specifying a rate function that models speed differences in tie changes between actors. Analyzing these differences will enable a better model selection for the analysis of longitudinal data, additional to the differences discussed in the article at hand.

We hope our study is a helpful step toward understanding the differences between the two most widely used statistical models for social network modeling, the ERGM and the SAOM (or "Siena model") at a time when the statistical analysis of social networks takes an increasingly important role in the social sciences. Using analytical model comparison as well as an empirical study, we have argued that a theoretically informed model choice is of crucial importance for empirical researchers.

Acknowledgment

The authors would like to thank members of the Tuscany group at Nuffield College, the Siena group at Groningen University, and the SoNAR-C at Lugano University for their valuable comments and feedback to this work.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. A brief history of the development of exponential random graph models (ERGMs) is summarized, for example, in Robins and Lusher (2013). Among the most important contributions in the development of the statistical model are Erdős and Rényi (1959), Holland and Leinhardt (1981), Frank and Strauss (1986), Wasserman and Pattison (1996), and Snijders et al. (2006).
2. "Far away" in the social space refers to a large sociometric distance.
3. Isomorphic networks are the same after a permutation of the nodes.
4. It should be noted that the mean value and the canonical or natural parameterizations are equivalent, with a 1:1 correspondence.

5. Three numbers of which the first is the number of Mutual dyads, the second the number of Asymmetric dyads, and the third the number of Null dyads, with an extra letter if necessary to resolve remaining ambiguities.
6. We use the term "attractive" to refer to tie changes that would increase the linear predictor more strongly, and therefore are more probable.
7. Sender dependence is a special case of the stochastic actor-oriented model (SAOM), which is discussed in Snijders (2005). It is similar to dyadic dependence in the ERGM, as it allows to compartmentalize the network into independent sections. For models with sender dependence, the rows of the adjacency matrix representing the network are independent.
8. The concept of social circuit dependence was proposed already in Pattison and Robins (2002), but without this name, the term was coined later and was used, for example, in Robins, Snijders, et al. (2007)
9. The contribution of closing the 021U triad to the change statistics in the ERGM led to an increased indegree centralization. At the same time, the relatively more prominent closure of the 021C and 021D triads in the SAOM lead to higher outdegree centralization. Additional analyses have confirmed this pattern, even though differences are small in size.
10. An introduction to dependence graphs and how they underlie ERGMs is beyond the scope of this article. Interested readers are referred to Besag (1974), Frank and Strauss (1986), and Whittaker (1990)
11. The chief scientist office of the Scottish Home and Health Department funded the study under their Smoking Initiative (grant K/OPR/17/8). Data collection was executed by Lynn Michell and Patrick West of the Medical Research Council and Medical Sociology Unit, University of Glasgow.
12. Very high is 100 in our analyses. A rate parameter of 100 consistently resulted in a turnover of more than 90 percent of ties by the end of the simulation, which is satisfying for our purposes.
13. A large p value indicates good model fit, as the associated null hypothesis is that the model fits the data.

References

Besag, J. 1974. "Spatial Interaction and the Statistical Analysis of Lattice Systems." *Journal of the Royal Statistical Society. Series B (Methodological)* 36:192-236.

Block, P. 2015. "Reciprocity, Transitivity, and the Mysterious Three-cycle." *Social Networks* 40:163-73.

Broekel, T., P.-A. Balland, M. Burger, and F. van Oort. 2014. "Modeling Knowledge Networks in Economic Geography: A Discussion of Four Methods." *The Annals of Regional Science* 53:423-52.

Desmarais, B. A. and S. J. Cranmer. 2012. "Micro-level Interpretation of Exponential Random Graph Models with Application to Estuary Networks." *Policy Studies Journal* 40:402-34.

Erdős, P. and A. Rényi. 1959. "On Random Graphs." *Publicationes Mathematicae Debrecen* 6:290-97.

Frank, O. and D. Strauss. 1986. "Markov Graphs." *Journal of the American Statistical Association* 81:832-42.

Hanneke, S., W. Fu, and E. P. Xing. 2010. "Discrete Temporal Models of Social Networks." *Electronic Journal of Statistics* 4:585-605.

Holland, P. W. and S. Leinhardt. 1976. "Local Structures in Social Networks." *Sociological Methodology* 7:1-45.

Holland, P. W. and S. Leinhardt. 1977. "A Dynamic Model for Social Networks." *Journal of Mathematical Sociology* 5:5-20.

Holland, P. W. and S. Leinhardt. 1981. "An Exponential Family of Probability Distributions for Directed Graphs." *Journal of the American Statistical Association* 76:33-50.

Hunter, D. R., S. M. Goodreau, and M. S. Handcock. 2008. "Goodness of Fit of Social Network Models." *Journal of the American Statistical Association* 103: 248-58.

Koskinen, J. H. and T. A. B. Snijders. 2013. "Simulation, Estimation, and Goodness of Fit." Pp. 141-66 in *Exponential Random Graph Models for Social Networks: Theories, Methods and Applications*, edited by D. Lusher, J. H. Koskinen, and G. L. Robins, Chap. 12. Cambridge, UK: Cambridge University Press.

Krivitsky, P. N. and M. S. Handcock. 2014. "A Separable Model for Dynamic Networks." *Journal of the Royal Statistical Society. Series B, Statistical Methodology* 76:29-46.

Lerner, J., N. Indlekofer, B. Nick, and U. Brandes. 2013. "Conditional Independence in Dynamic Networks." *Journal of Mathematical Psychology* 57:275-83.

Lospinoso, J. 2012. *Statistical Models for Social Network Dynamics*. DPhil thesis. Oxford, UK: Department of Statistics, University of Oxford.

Lusher, D., J. Koskinen, and G. Robins. 2013. *Exponential Random Graph Models for Social Networks*. Cambridge, UK: Cambridge University Press.

Lusher, D. and G. L. Robins. 2013. "Formation of Social Network Structure." Pp. 16-28 in *Exponential Random Graph Models for Social Networks: Theories, Methods and Applications*, edited by D. Lusher, J. H. Koskinen, and G. L. Robins, Chap. 3. Cambridge, UK: Cambridge University Press.

McFadden, D. 1974. "Condition Logit Analysis of Qualitative Choice Behavior." *Frontiers in Econometrics* 1:105-43.

Nelder, J. A. and R. Mead. 1965. "A Simplex Method for Function Minimization." *The Computer Journal* 7:308-13.

Norris, J. R. (1997). *Markov Chains*. Cambridge, UK: Cambridge University Press.

Pattison, P. E. and G. L. Robins. 2002. "Neighbourhood based Models for Social Networks." *Sociological Methodology* 32:301-37.

Pattison, P. E. and T. A. B. Snijders. 2013. "Modeling Social Networks: Next Steps." Pp. 287-301 in *Exponential Random Graph Models for Social Networks: Theories, Methods and Applications*, edited by D. Lusher, J. H. Koskinen, and G. L. Robins, Chap. 22. Cambridge, UK: Cambridge University Press.

Pearson, M. and L. Michell. 2000. "Smoke Rings: Social Network Analysis of Friendship Groups, Smoking and Drug-taking." *Drugs-Education Prevention and Policy* 7:21-37.

R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ripley, R. M., T. A. B. Snijders, Z. Boda, A. Vörös, and P. Preciado. 2016. *Manual for RSiena, version 1.1-294*. Oxford, UK: University of Oxford, Nuffield College.

Rivera, M. T., S. B. Soderstrom, and B. Uzzi. 2010. "Dynamics of Dyads in Social Networks: Assortative, Relational, and Proximity Mechanisms." *Annual Review of Sociology* 36:91-115.

Robins, G., P. Pattison, Y. Kalish, and D. Lusher. (2007). "An Introduction to Exponential Random Graph (p*) Models for Social Networks." *Social Networks* 29:173-91.

Robins, G. L. and D. Lusher. 2013. "What are Exponential Random Graph Models?" Pp. 9-15 in *Exponential Random Graph Models for Social Networks: Theories, Methods and Applications*, edited by D. Lusher, J. H. Koskinen, and G. L. Robins. Cambridge, UK: Cambridge University Press.

Robins, G. L., T. A. B. Snijders, P. Wang, M. S. Handcock, and P. E. Pattison. 2007. "Recent Developments in Exponential Random Graph (p^*) Models for Social Networks." *Social Networks* 29:192-215.

Snijders, T. A. B. 1996. "Stochastic Actor-oriented Dynamic Network Analysis." *Journal of Mathematical Sociology* 21:149-72.

Snijders, T. A. B. 2001. "The Statistical Evaluation of Social Network Dynamics." *Sociological Methodology* 31:1-33.

Snijders, T. A. B. 2005. "Models for Longitudinal Network Data." Pp. 215-47 in *Models and Methods in Social Network Analysis*, edited by P. Carrington, J. Scott, and S. Wasserman. New York, UK: Cambridge University Press.

Snijders, T. A. B. and J. Koskinen. 2013. "Longitudinal Models." in *Exponential Random Graph Models for Social Networks: Theories, Methods and Applications*, edited by D. Lusher, J. H. Koskinen, and G. L. Robins, Chap. 11. Cambridge, UK: Cambridge University Press.

Snijders, T. A. B., P. E. Pattison, G. L. Robins, and M. S. Handcock. 2006. "New Specifications for Exponential Random Graph Models." *Sociological Methodology* 36:99-153.

Snijders, T. A. B. and C. E. G. Steglich. 2015. "Representing Micro-macro Linkages by Actor-based Dynamic Network Models." *Sociological Methods & Research* 44:222-71.

Snijders, T. A. B., G. G. van de Bunt, and C. E. G. Steglich. 2010. "Introduction to Stochastic Actor-based Models for Network Dynamics." *Social Networks* 32: 44-60.

Snijders, T. A. B. and M. van Duijn. 1997. "Simulation for Statistical Inference in Dynamic Network Models." Pp. 493-512 in *Simulating Social Phenomena*, edited by R. Conte, R. Hegselmann, and P. Terna. Berlin: Springer.

Wang, P., G. L. Robins, P. E. Pattison, and J. H. Koskinen. 2014. *MPNet*. Melbourne: Melbourne School of Psychological Sciences.

Wasserman, S. and P. Pattison. 1996. "Logit Models and Logistic Regressions for Social Networks: I. An Introduction to Markov Graphs and p*." *Psychometrika* 61:401-25.

Whittaker, J. 1990. *Graphical Models in Applied Multivariate Statistics*. Chichester, UK: Wiley Publishing.

Author Biographies

Per Block is an ETH Fellow at the chair of Social Networks at ETH Zürich, Switzerland. His research focuses on the evolution of social networks and dependence between ties, actors and social positions.

Christoph Stadtfeld is an assistant professor of Social Networks at ETH Zürich, Switzerland. His research focuses on the development and application of theories and methods for social network dynamics.

Tom A. B. Snijders is a professor of Methodology and Statistics in the Department of Sociology at the University of Groningen, and an emeritus fellow of Nuffield College (University of Oxford). His research focuses on methods of network analysis and multilevel analysis.