Introduction to Stochastic Actor-Based Models for Network Dynamics *

Tom A.B. Snijders[†]
Gerhard G. van de Bunt[‡]
Christian E.G. Steglich[§]

Abstract

Stochastic actor-based models are models for network dynamics that can represent a wide variety of influences on network change, and allow to estimate parameters expressing such influences, and test corresponding hypotheses. The nodes in the network represent social actors, and the collection of ties represents a social relation. The assumptions posit that the network evolves as a stochastic process 'driven by the actors', i.e., the model lends itself especially for representing theories about how actors change their outgoing ties. The probabilities of tie changes are in part endogenously determined, i.e., as a function of the current network structure itself, and in part exogenously, as a function of characteristics of the nodes ('actor covariates') and of characteristics of pairs of nodes ('dyadic covariates'). In an extended form, stochastic actor-based models can be used to analyze longitudinal data on social networks jointly with changing attributes of the actors: dynamics of networks and behavior.

This paper gives an introduction to stochastic actor-based models for dynamics of directed networks, using only a minimum of mathematics. The focus is on understanding the basic principles of the model, understanding the results, and on sensible rules for model selection.

Keywords: statistical modeling, longitudinal, Markov chain, agent-based model, peer selection, peer influence.

*Draft article for special issue of Social Networks on *Dynamics of Social Networks*. We are grateful to Andrea Knecht who collected the data used in the example, under the guidance of Chris Baerveldt. We also are grateful to Matthew Checkley and two reviewers for their very helpful remarks on earlier drafts.

[†]University of Oxford and University of Groningen
[‡]Free University, Amsterdam
[§]University of Groningen

1. Introduction

Social networks are dynamic by nature. Ties are established, they may flourish and perhaps evolve into close relationships, and they can also dissolve quietly, or suddenly turn sour and go with a bang. These relational changes may be considered the result of the structural positions of the actors within the network – e.g., when friends of friends become friends–, characteristics of the actors ('actor covariates'), characteristics of pairs of actors ('dyadic covariates'), and residual random influences representing unexplained influences. Social network research has in recent years paid increasing attention to network dynamics, as is shown, e.g., by the three special issues devoted to this topic in *Journal of Mathematical Sociology* edited by Patrick Doreian and Frans Stokman (1996, 2001, and 2003; also see Doreian and Stokman, 1997). The three issues shed light on the underlying theoretical micro mechanisms that induce the evolution of social network structures on the macro level. Network dynamics is important for domains ranging from friendship networks (e.g., Pearson and Michell, 2000; Burk, Steglich, and Snijders, 2007) to, for example, organizational networks (see the review articles Borgatti and Foster, 2003; Brass, Galaskiewicz, Greve, and Tsai, 2004).

In this article we give a tutorial introduction to what we call here *stochastic actor-based models for network dynamics*, which are a type of models that have the purpose to represent network dynamics on the basis of observed longitudinal data, and evaluate these according to the paradigm of statistical inference. This means that the models should be able to represent network dynamics as being driven by many different tendencies, such as the micro mechanisms alluded to above, which could have been theoretically derived and/or empirically established in earlier research, and which may well operate simultaneously. Some examples of such tendencies are reciprocity, transitivity, homophily, and assortative matching, as will be elaborated below. In this way, the models should be able to give a good representation of the stochastic dependence between the creation, and possibly termination, of different network ties. These stochastic actor-based models allow to test hypotheses about these tendencies, and to estimate parameters expressing their strengths, while controlling for other tendencies (which in statistical terminology might be called 'confounders').

The literature on network dynamics has generated a large variety of mathematical models. To describe the place in the literature of stochastic actor-based models (Snijders 1996, 2001), these models may be contrasted with other dynamic network models.

Most network dynamics models in the literature pay attention to a very specific set of micro mechanisms –allowing detailed analyses of the properties of these models–, but lack an explicit estimation theory. Examples are models pro-

posed by Bala and Goyal (2000), Hummon (2000), Skyrms and Pemantle (2000), and Marsili, Vega-Redondo, and Slanina (2004), all being actor-based simulation models that focus on the expression of a single social theory as reflected, e.g., by a simple utility function; those proposed by Jin, Girvan, and Newman (2001) which represent a larger but still quite restricted number of tendencies; and models such as those proposed by Price (1976), Barabási and Albert (1999), and Jackson and Rogers (2007), which are actor-based, represent one or a restricted set of tendencies, and assume that nodes are added sequentially while existing ties cannot be deleted, which is a severe limitation to the type of longitudinal data that may be faithfully represented. Since such models do not allow to control for other influences on the network dynamics, and how to estimate and test parameters is not clear for them, they cannot be used for purposes of theory testing in a statistical model.

The earlier literature does contain some statistical dynamic network models, mainly those developed by Wasserman (1979) and Wasserman and Iacobucci (1988), but these do not allow complicated dependencies between ties such as are generated by transitive closure. Further there are papers that present an empirical analysis of network dynamics which are based on intricate and illuminating descriptions such as Holme, Edling, and Liljeros (2004) and Kossinets and Watts (2006), but which are not based on an explicit stochastic model for the network dynamics and therefore do not allow to control one tendency for other ('confounding') tendencies.

Distinguishing characteristics of stochastic actor-based models are flexibility, allowing to incorporate a wide variety of actor-driven micro mechanisms influencing tie formation; and the availability of procedures for estimating and testing parameters that also allow to assess the effect of a given mechanism while controlling for the possible simultaneous operation of other mechanisms or tendencies.

We assume here that the empirical data consist of two, but preferably more, repeated observations of a social network on a given set of actors; one could call this *network panel data*. Ties are supposed to be the dyadic constituents of relations such as friendship, trust, or cooperation, directed from one actor to another. In our examples social actors are individuals, but they could also be firms, countries, etc. The ties are supposed to be, in principle, under control of the sending actor (although this will be subject to constraints), which will exclude most types of relations where negotiations are required for a tie to come into existence. Actor covariates may be constant like sex or ethnicity, or subject to change like opinions, attitudes, or lifestyle behaviors. Actor covariates often are among the determinants of actor similarity (e.g., same sex or ethnicity) or spatial proximity between actors (e.g., same neighborhood) which influence the existence of ties.

Dyadic covariates likewise may be constant, such as determined by kinship or formal status in an organization, or changing over time, like friendship between parents of children or task dependencies within organizations.

This paper is organized as follows. In the next section, we present the assumptions of the actor-based model. The heart of the model is the so-called objective function, which determines probabilistically the tie changes made by the actors. One could say that it captures all theoretically relevant information the actors need to 'evaluate' their collection of ties. Some of the potential components of this function are structure-based (endogenous effects), such as the tendency to form reciprocal relations, others are attribute-based (exogenous effects), such as the preference for similar others. In Section 3 we discuss several statistical issues, such as data requirements and how to test and select the appropriate model. Following this we present an example about friendship dynamics, focusing on the interpretation of the parameters. Section 4 proposes some more elaborate models. In Section 5, models for the coevolution of networks and behavior are introduced and illustrated by an example. Section 6 discusses the difference between equilibrium and out-of-equilibrium situations, and how these longitudinal models relate to cross-sectional statistical modeling of social networks. Finally, in Section 7, a brief discussion is given, the *Siena* software is mentioned which implements these methods, and some further developments are presented.

2. Model assumptions

A dynamic network consists of ties between actors that change over time. A foundational assumption of the models discussed in this paper is that the network ties are not brief *events*, but can be regarded as *states* with a tendency to endure over time. Many relations commonly studied in network analysis naturally satisfy this requirement of gradual change, such as friendship, trust, and cooperation. Other networks more strongly resemble 'event data', e.g., the set of all telephone calls among a group of actors at any given time point, or the set of all e-mails being sent at any given time point. While it is meaningful to interpret these networks as indicators of communication, it is not plausible to treat their ties as enduring states, although it often is possible to aggregate event intensity over a certain period and then view these aggregates as indicators of states.

Given that the network ties under study denote states, it is further assumed, as an approximation, that the changing network can be interpreted as the outcome of a Markov process, i.e., that for any point in time, the current state of the network determines probabilistically its further evolution, and there are no additional effects of the earlier past. All relevant information is therefore assumed to be included in the current state. This assumption often can be made more plausible by

choosing meaningful independent variables that incorporate relevant information from the past.

This paper gives a non-technical introduction into actor-based models for network dynamics. More precise explanations can be found in Snijders (2001, 2005) and Snijders, Steglich, and Schweinberger (2007). However, a modicum of mathematical notation cannot be avoided. The tie variables are binary variables, denoted by x_{ij}. A tie from actor i to actor j, denoted $i \to j$, is either present or absent (x_{ij} then having values 1 and 0, respectively). Although this is in line with traditional network analysis, an extension to valued networks would often be theoretically sound, and could make the Markov assumption more plausible. This is one of the topics of current research. The tie variables constitute the network, represented by its $n \times n$ adjacency matrix $x = \left(x_{ij}\right)$ (self-ties are excluded), where n is the total number of actors. The changes in these tie variables are the dependent variables in the analysis.

2.1. Basic assumptions

The model is about directed relations, where each tie $i \to j$ has a sender i, who will be referred to as *ego*, and a receiver j, referred to as *alter*. The following assumptions are made.

1. The underlying time parameter t is continuous, i.e., the process unfolds in time steps of varying length, which could be arbitrarily small. The parameter estimation procedure, however, assumes that the network is observed only at two or more discrete points in time. The observations can also be referred to as 'network panel waves', analogous to panel surveys in non-network studies.

   This assumption was proposed already by Holland and Leinhardt (1977), and elaborated by Wasserman (1979 and other publications) and Leenders (1995 and other publications) – but their models represented only reciprocity, and no other structural dependencies between network ties. The continuous-time assumption allows to represent dependencies between network ties as the result of processes where one tie is formed as a reaction to the existence of other ties. If, for example, three actors who at the first observation are mutually disconnected form at the second observation a closed triangle, where each is connected to both of the others, then a discrete-time model that has the observations as its time steps would have to account for the fact that this closed triangle structure is formed 'out of nothing', in one time step. In our model such a closed triangle can be formed tie by tie, as a consequence of reciprocation and transitive closure. Thus, the appearance

of closed triangles may be explained based on reciprocity and transitive processes, without requiring a special process specifically for closed triangles.

Since many small changes can add up to large differences between consecutively observed networks, this does not preclude that what is observed shows large changes from one observation to the next.

2. The changing network is the outcome of a Markov process. This was explained above. Thus, the total network structure is the social context that influences the probabilities of its own change.

   The assumption of a Markov process has been made in practically all models for social network dynamics, starting by Katz and Proctor's (1959) discrete Markov chain model. This is an assumption that will usually not be realistic, but it is difficult to come up with manageable models that do not make it. We could say that this assumption is a lens through which we look at the data – it should help but it also may distort. If there are only two panel waves, then the data have virtually no information to test this assumption. For more panel waves, there is in principle the possibility to test this assumption and propose models making less restrictive assumption about the time dependence, but this will require quite complicated models.

3. The actors control their outgoing ties. This means not that actors can change their outgoing ties at will, but that changes in ties are made by the actors who send the tie, on the basis of their and others' attributes, their position in the network, and their perceptions about the rest of the network. This assumption is the reason for using the term 'actor-based model'. This approach to modeling is in line with the methodological approach of structural individualism (Udehn, 2002; Hedström, 2005), where actors are assumed to be purposeful and to behave subject to structural constraints. The assumption of purposeful actors is not required, however, but a question of model interpretation (see below).
   It is assumed formally that actors have full information about the network and the other actors. In practice, as can be concluded from the specifications given below, the actors only need more limited information, because the probabilities of network changes by an actor depend on the personal network (including actors' attributes) that would result from making any given change, or possibly the personal network including those to whom the actor has ties through one intermediary (i.e., at geodesic distance two).

4. At a given moment one probabilistically selected actor – 'ego' – may get the opportunity to change one outgoing tie. No more than one tie can change at

any moment – a principle first proposed by Holland and Leinhardt (1977). This principle decomposes the change process into its smallest possible components, thereby allowing for relatively simple modelling. This implies that tie changes are not coordinated, and depend on each other only sequentially, via the changing configuration of the whole network. For example, two actors cannot decide jointly to form a reciprocal tie; if two actors are not tied at one observation and mutually tied at the next, then one of the must have taken the initiative and extended a one-sided tie, after which, at some later moment, the other actor reciprocated and formed a reciprocal tie. This assumption excludes relational dynamics where some kind of coordination or negotiation is essential for the creation of a tie, or networks created by groups participating in some activity, such as joint authorship networks.

For directed networks, this usually is a reasonable simplifying assumption. In most cases, panel data of directed networks have many tie differences between successive observations and do not provide information about the order in which ties were created or terminated, so that this is an assumption about which the available data contain hardly any empirical evidence.

Summarizing the status of these four basic assumptions: the first (continuous-time model) makes sense intuitively; the second (Markov process) is an as-if approximation and it would be interesting in future to construct models going beyond this assumption; the third (actor-based) is often a helpful theoretical heuristic; and the fourth (ties change one by one) is an assumption which limits the applicability to a wide class of panel data of directed networks for which this assumption seems relatively harmless.

The actor-based network change process is decomposed into two sub-processes, both of which are stochastic.

5. The change opportunity process, modeling the frequency of tie changes by actors. The change rates may depend on the network positions of the actors (e.g., centrality) and on actor covariates (e.g., age and sex).

6. The change determination process, modeling the precise tie changes made when an actor has the opportunity to make a change. The probabilities of tie changes may depend on the network positions, as well as covariates, of ego and the other actors ('alters') in the network. This is explained below.

The actor-based model can be regarded as an agent-based simulation model (Macy and Willer 2002). It does not deviate in principle from other agent-based models, only in ways deriving from the fact that the model is to be used for statistical inference, which leads to requirements of flexibility (enough parameters that can be estimated from the data to achieve a good fit between model and data) and

parsimony (not more fine detail in the model than what can be estimated from the data). The word 'actor' rather than 'agent' is used, in line with other sociological literature (e.g., Hedström 2005), to underline that actors are not regarded as subservient to others' interests in any way.

The actor-based model, when elaborated for a practical application, contains parameters that have to be estimated from observed data by a statistical procedure. Since the proposed stochastic models are too complex for the straightforward application of classical estimation methods such as maximum likelihood, Snijders (1996, 2001) proposed a procedure using the method of moments implemented by computer simulation of the network change process. This procedure uses the basic principle that the first observed network is itself not modeled but used only as the starting point of the simulations. In statistical terminology: the estimation procedure conditions on the first observation. This implies that it is the change between two observed periods time points that is being modeled, and the analysis does not have the aim to make inferences about the determinants of the network structure at the first time point.

2.2. Change determination model

The first step in the model is the choice of the focal actor (ego) who gets the opportunity to make a change. This choice can be made with equal probabilities or with probabilities depending on attributes or network position, as elaborated in Section 4.1. This selected focal actor then may change one outgoing tie (i.e., either initiate or withdraw a tie), or do nothing (i.e., keep the present status quo). This means that the set of admissible actions contains n elements: $n - 1$ changes and one non-change. The probabilities for a choice depend on the so-called *objective function*. This is a function of the network, as perceived by the focal actor. Informally, the objective function expresses how likely it is for the actor to change her/his network in a particular way. On average, each actor 'tries to' move into a direction of higher values of her/his objective function, subject to the constraints of the current network structure and the changes made by the other actors in the network; and subject to random influences. The objective function will depend in practice on the personal network of the actor, as defined by the network between the focal actor plus those to whom there is a direct tie (or, depending on the specification, the focal actor plus those to whom there is a direct or indirect – i.e., distance-two – tie), including the covariates for all actors in this personal network. Thus, the probabilities of changes are assumed to depend on the personal networks that would result from the changes that possibly could be made, and their composition in terms of covariates, via the objective function values of those networks.

The precise interpretation is given by equation (4) in the appendix. This is the

core of model and it must represent the research questions and relevant theoretical and field-related knowledge. The objective function is explained in more detail in the next section.

The name 'objective function' was chosen because one possible interpretation is that it represents the short-term objectives (net result of preferences and structural as well as cognitive constraints) of the actor. Which action to choose out of the set of admissible actions, given that ego has the opportunity to act (i.e., change a network tie), follows the logic of discrete choice models (McFadden, 1973; Maddala, 1983) which have been developed for modeling situations where the dependent variable is a choice made from a finite set of actions.

2.3. Specification of the objective function

The objective function determines the probabilities of change in the network, given that an actor has the opportunity to make a change. One could say it represents the 'rules for network behavior' of the actor. This function is defined on the set of possible states of the network, as perceived from the point of view of the focal actor, where 'state of the network' refers not only to the ties but also to the covariates. When the actor has the possibility of moving to one out of a set of network states, the probability of any given move is higher accordingly as the objective function for that state is higher.

Like in generalized linear statistical models, the objective function is assumed to be a linear combination of a set of components called *effects*,

$$f_i(\beta, x) = \sum_k \beta_k \, s_{ki}(x) \; . \tag{1}$$

In this section and elsewhere, the symbol i and the term 'ego' are ways of referring to the focal actor. Here $f_i(\beta, x)$ is the value of the objective function for actor i depending on the state x of the network; the functions $s_{ki}(x)$ are the effects, functions of the network that are chosen based on theory and subject-matter knowledge, and correspond to the 'tendencies' mentioned in the introductory section; and the weights β_k are the statistical parameters. The effects represent aspects of the network as 'viewed' from the point of view of actor i. As examples, one can think of the number of reciprocated ties of actor i, representing tendencies toward reciprocity, or the number of ties from i toward actors of the same gender, representing tendencies toward gender homophily. Many more examples are presented below. The effects $s_{ki}(x)$ depend on the network x but may also depend on actor attributes (actor covariates), on variables depending on pairs of actors (dyadic covariates), etc. If β_k equals 0, the corresponding effect plays no role in the network dynamics; if β_k is positive, then there will be a higher probability of

moving into directions where the corresponding effect is higher, and the converse if β_k is negative.

For the model selection, an essential part is the theory-guided choice of effects included in the objective function in order to test the formulated hypotheses. A good approach may be to progressively build up the model according to the method of decreasing abstraction (Lindenberg, 1992). An additional consideration here is, however, that the complexity of network processes, and the limitations of our current knowledge concerning network dynamics, imply that model construction may require data-driven elements to select the most appropriate precise specification of the endogenous network effects. For example, in the investigation of friendship networks one might be interested in effects of lifestyle variables and background characteristics on friendship, while recognizing the necessity to control for tendencies toward reciprocation and transitive closure. As discussed below in the section on triadic effects, multiple mathematical specifications are available (as 'effects' $s_{ki}(x)$ to be included in equation (1)) expressing the concept of transitive closure. Usually there are no prior theoretical or empirical reasons for choosing among these specifications. It may then be best to use theoretical considerations for deciding to include lifestyle-related and background variables as well as tendencies toward reciprocation and transitive closure in the model, and to choose the best specification for transitive closure, by one or several specific effects, in a data-driven way.

In the following we give a number of effects that may be considered for inclusion in the objective function. They are described here only conceptually, with some brief pointers to empirical results or theories that might support them; the formulae are given in the appendix. Effects depending only on the network are called *structural* or endogenous effects, while effects depending only on externally given attributes are called *covariate* or exogenous effects. The complexity of networks is such that an exhaustive list cannot meaningfully be given. To simplify formulations, the presentation shall assume that the relation under study is friendship, so the existence of a tie $i \to j$ will be described as i calling j a friend. Higher values of the objective function, leading to higher tendencies to form ties, will sometimes be interpreted in shorthand as preferences.

Basic effects

The most basic effect is defined by the *outdegree* of actor i, and this will be included in all models. It represents the basic tendency to have ties at all, and in a decision-theoretic approach its parameter could be regarded as the balance of benefits and costs of an arbitrary tie. Most networks are sparse (i.e., they have a density well below 0.5) which can be represented by saying that for a tie to an arbitrary other actor – arbitrary meaning here that the other actor has no characteristics or tie pattern making him/her especially attractive to i –, the costs will

usually outweigh the benefits. Indeed, in most cases a negative parameter is obtained for the outdegree effect.

Another quite basic effect is the tendency toward *reciprocity*, represented by the number of reciprocated ties of actor i. This is a basic feature of most social networks (cf. Wasserman and Faust, 1994, Chapter 13) and usually we obtain quite high values for its parameter, e.g., between 1 and 2.

Transitivity and other triadic effects

Next to reciprocity, an essential feature in most social networks is the tendency toward transitivity, or transitive closure (sometimes called clustering): friends of friends become friends, or in graph-theoretic terminology: two-paths tend to be, or to become, closed (e.g., Davis 1970, Holland and Leinhardt 1971). In Figure 1.a, the two-path $i \rightarrow j \rightarrow h$ is closed by the tie $i \rightarrow h$.

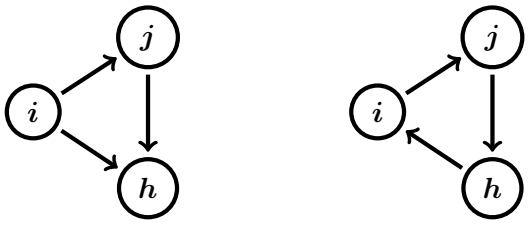

Figure 1. a. Transitive triplet (i, j, h) b. Three-cycle

The *transitive triplets* effect measures transitivity for an actor i by counting the number of pairs j, h such that there is the transitive triplet structure of Figure 1.a. However, this is just one way of measuring transitivity. Another one is the *transitive ties* effect, which measures transitivity for actor i by counting the number of other actors h for which there is at least one intermediary j forming a transitive triplet of this kind. The transitive triplets effect postulates that more intermediaries will add proportionately to the tendency to transitive closure, whereas the transitive ties effect expects that given that one intermediary exists, extra intermediaries will not further contribute to the tendency to forming the tie $i \rightarrow h$.

An effect closely related to transitivity is *balance* (cf. Newcomb, 1962), which in our implementation is the same as *structural equivalence with respect to outties* (cf. Burt, 1982), which is the tendency to have and create ties to other actors who make the same choices as ego. The extent to which two actors make the same choices can be expressed simply as the number of outgoing choices and non-choices that they have in common.

Transitivity can be represented by still more effects: e.g., negatively, by the number of others to whom i is indirectly tied but not directly (geodesic distance equal to 2). The choice between these representations of transitivity may depend

both on the degree to which the representation is theoretically convincing, and on what gives the best fit.

A different triadic effect is the number of three-cycles that actor i is involved in (Figure 1.b). Davis (1970) found that in many social network data sets, there is a tendency to have relatively few three-cycles, which can be represented here by a negative parameter β_k for the three-cycle effect. The transitive triplets and the three-cycle effects both represent closed structures, but whereas the former is in line with a hierarchical ordering, the latter goes against such an ordering. If the network has a strong hierarchical tendency, one expects a positive parameter for transitivity and a negative for three-cycles. Note that a positive three-cycle effect can also be interpreted, depending on the context of application, as a tendency toward generalized exchange (Bearman, 1997).

Degree-related effects

In- and outdegrees are primary characteristics of nodal position and can be important driving factors in the network dynamics.

One pair of effects is *degree-related popularity* based on *indegree* or *outdegree*. If these effects are positive, nodes with higher indegree, or higher outdegree, are more attractive for others to send a tie to. This can be measured by the sum of indegrees of the targets of i's outgoing ties, and the sum of their outdegrees, respectively. A positive indegree-related popularity effect implies that high indegrees reinforce themselves, which will lead to a relatively high dispersion of the indegrees (a Matthew effect in popularity as measured by indegrees, cf. Merton, 1968 and Price, 1976). A positive outdegree-related popularity effect will increase the association between indegrees and outdegrees, or keep this association relatively high if it is high already.

Another pair of effects is *degree-related activity* for *indegree* or *outdegree*: when these effects are positive, nodes with higher indegree, or higher outdegree respectively, will have an extra propensity to form ties to others. These effects can be measured by the indegree of i times i's outdegree; and, respectively, the outdegree of i times i's outdegree, that is, the square of the outdegree.[1] The outdegree-related activity effect again is a self-reinforcing effect: when it has a positive parameter, the dispersion of outdegrees will tend to increase over time, or to be sustained if it already is high. The indegree-related activity effect has the same consequence as the outdegree-related popularity effect: positive parameters lead to a relatively high association between indegrees and outdegrees. Therefore these two effects will be difficult, or impossible, to distinguish empirically,

[1] Experience has shown that for the degree-related effects, often the 'driving force' is measured better by the square roots of the degrees than by raw degrees. In some cases this may be supported by arguments about diminishing returns of increasingly high degrees. See the formulae in the appendix.

and the choice between them will have to be made on theoretical grounds. These four degree-related effects can be regarded as the analogues in the case of directed relations of what was called *cumulative advantage* by Price (1976) and *preferential attachment* by Barabási and Albert (1999) in their models for dynamics of non-directed networks: a self-reinforcing process of degree differentiation.

These degree-related effects can represent hierarchy between nodes in the network, but in a different way than the triadic effects of transitivity and 3-cycles. The degree-related effects represent global hierarchy while the triadic effects represent local hierarchy. In a perfect hierarchy, ties go from the bottom to the top, so that the bottom nodes have high outdegrees and low indegrees and the top nodes have low outdegrees and high indegrees. This will be reflected by positive indegree popularity and negative outdegree popularity, and by positive outdegree activity and negative indegree activity. Therefore, to differentiate between local and global hierarchical processes, it can be interesting to estimate models with triadic and degree-related effects, and assess which of these have the better fit by testing the triadic parameters while controlling for the degree-related parameters, and vice versa.

Other degree-related effects are *assortativity-related*: actors might have preferences for other actors based on their own and the other's degrees (Morris and Kretzschmar 1995; Newman 2002). In settings where degrees reflect status of the actors, such preferences may be argued theoretically based on status-specific preferences, constraints, or matching processes. This gives four possibilities, depending on in- and outdegree of the focal actor and the potential friend.

Together, this list offers eight degree-related effects. The outdegree-related popularity and indegree-related activity effects are nearly collinear, and it was already mentioned that theory, not empirical fit, will have to decide which one is a more meaningful representation. Some of the other effects also may be confounded, but this depends on the data set. The four effects described as degree-related popularity and activity are more basic than the assortativity effects (cf. the relation between main effects and interactions in linear regression). Because of this, when testing any assortativity effects, one usually should control for three of the degree-related popularity and activity effects.

Covariates: exogenous effects

For an actor variable V, there are three basic effects: the *ego* effect, measuring whether actors with higher V values tend to nominate more friends and hence have a higher outdegree (which also can be called covariate-related activity effect or sender effect); the *alter* effect, measuring whether actors with higher V values will tend to be nominated by more others and hence have higher indegrees (covariate-related popularity effect, receiver effect); and the *similarity effect*, measuring whether ties tend to occur more often between actors with similar values on

V (homophily effect). Tendencies to homophily constitute a fundamental characteristic of many social relations, see McPherson, Smith-Lovin, and Cook (2001). When the ego and alter effects are included, instead of the similarity effect one could use the *ego-alter interaction effect*, which expresses that actors with higher V values have a greater preference for other actors who likewise have higher V values.

For categorical actor variables, the *same* V effect measures the tendency to have ties between actors with exactly the same value of V.

For a dyadic covariate, i.e., a variable defined for pairs of actors, there is one basic effect, expressing the extent to which a tie between two actors is more likely when the dyadic covariate is larger.

Interactions

Like in other statistical models, interactions can be important to express theoretically interesting hypotheses. The diversity of functions that could be used as effects makes it difficult to give general expressions for interactions. The ego-alter interaction effect for an actor covariate, mentioned above, is one example.

Another example is given by de Federico (2004) as an interaction of a covariate with reciprocity. In her analysis of a friendship network between exchange students, she found a negative interaction between reciprocity and having the same nationality. Having the same nationality has a positive main effect, reflecting that it is easier to become friends with those coming from the same country. The negative interaction effect was unexpected, but can be explained by regarding reciprocation as a response to an initially unreciprocated tie, the latter being a unilateral invitation to friendship. Since contacts between those with the same nationality are easier than between individuals from different nationalities, extending a unilateral invitation to friendship is more remarkable (and perhaps more costly) between individuals of different nationalities than between those of the same nationality. Therefore it will be noticed and appreciated, and hence reciprocated, with a higher probability. Thus, the rarity of cross-national friendships leads to a stronger tendency to reciprocation in cross-national than same-nationality friendships.

As a further class of examples, note that in the actor-based framework it may be natural to hypothesize that the strength of certain effects depends on attributes of the focal actor. For example, girls might have a greater tendency toward transitive closure than boys. This can be modeled by the interaction of the ego effect of the attribute and the transitive triplets, or transitive ties effect.

Other interactions (and still other effects) are discussed in Snijders et al. (2008). As the selection presented here already illustrates, the portfolio of possible effects in this modeling approach is very extensive, naturally reflecting the multitude of possibilities by which networks can evolve over time. Therefore, the selection

of meaningful effects for the analysis of any given data set is vital. This will be discussed now.

3. Issues arising in statistical modeling

When employing these models, important practical issues are the question how to specify the model – boiling down mainly to the choice of the terms in the objective function – and how to interpret the results. This is treated in the current section.

3.1. Data requirements

To apply this model, the assumptions should be plausible in an approximate sense, and the data should contain enough information. Although rules of thumb always must be taken with many grains of salt, we first give some numbers to indicate the sizes of data sets which might be well treated by this model. These rules of thumb are based on practical experience.

The amount of information depends on the number of actors, the number of observation moments ('panel waves'), and the total number of changes between consecutive observations. The number of observation moments should be at least 2, and is usually much less than 10. There are no objections in principle against analyzing a larger number of time points, but then one should check the assumption that the parameters in the objective function are constant over time, or that the trends in these parameters are well represented by their interactions with time variables (see point 10 below).

If one has more than two observation points, then in practice one may wish to start by analyzing the transitions between each consecutive pair of observations (provided these provide enough information for good estimation – see below). For each parameter one then can present the trend in estimated parameter values, and depending on this one can make an analysis of a larger stretch of observations if the parameters appear approximately constant, or do the same while including for some of the parameters an interaction with a time variable.

The number of actors will usually be larger than 20 – but if the data contain many waves, a smaller number of actors could be acceptable. The number of actors will usually not be more than a few hundred, because the implicit assumption that each actor is a potential network partner for any other actor might be implausible for networks with so many actors that not all actors are aware of each others' existence.

The total number of changes between consecutive observations should be large enough, because these changes provide the information for estimating the parameters. A total of 40 changes (cumulated over all successive panel waves) is on

the low side. More changes will give more information and, thereby, allow more complicated models to be fitted. Between any pair of consecutive waves, the number of changes should not be too high, because this would call into question the assumption that the waves are consecutive observations of a gradually changing network; or, if they were, the consecutive observations would be too far apart.

This implies that, when designing the study, the researcher has to have a reasonable estimate of how much change to expect. For instance, if one is interested in the development of the friendship network of a group of initially mutual strangers (e.g., university freshmen), it may be good to plan the observation moments to be separated by only a few weeks, and to enlarge the period between observations after a couple of months. On the other hand, if one studies inter-firm network dynamics, given the time delays involved for firms in the planning and executing of their ties to other firms, it may be enough to collect data once every year, or even less frequently.

To express quantitatively whether the data collection points are not too far apart, one may use the Jaccard (1900) index (also see Batagelj and Bren, 1995), applied to tie variables. This measures the amount of change between two waves by

$$\frac{N_{11}}{N_{11} + N_{01} + N_{10}}\,, \tag{2}$$

where N_{11} is the number of ties present at both waves, N_{01} is the number of ties newly created, and N_{10} is the number of ties terminated. Experience has shown that Jaccard values between consecutive waves should preferably be higher than .3, and – unless the first wave has a much lower density than the second – values less than .2 would lead to doubts about the assumption that the change process is gradual, compared to the observation frequency. If the network is in a period of growth and the second network has many more ties than the first, one may look instead at the proportion, among the ties present at a given observation, of ties that have remained in existence at the next observation ($N_{11}/(N_{10} + N_{11})$ in the preceding notation). Proportions higher than .6 are preferable, between .3 and .6 would be low but may still be acceptable. If the data collection was such that values of ties (ranging from weak to strong) were collected, then these numbers may be used as rough rules of thumb and give some guidance for the decision where to dichotomize the tie values – although, of course, substantive concerns related to the interpretation of the results have primacy for such decisions.

The methods require in principle that network data are complete. However, it is allowed that some of the actors enter the network after the start, or leave before the end of the panel waves (Huisman and Snijders, 2003), and a limited amount of missing data can be accommodated (Huisman and Steglich, 2008). Another option to represent that some actors are not yet, or no more, present in the net-

work, is to specify that certain ties cannot exist ('structural zeros') or that some ties are prescribed ('structural ones'), see Snijders et al. (2008). The use of structural zeros allows, e.g., to combine several small networks into one structure (with structural zeros forbidding ties between different networks), allowing to analyze multiple independent networks that on themselves would not yield enough information for estimating parameters, under the extra assumption that all network follow dynamics with the same parameter values in the objective function.

3.2. Testing and model selection

It turns out (supported by computer simulations) that the distributions of the estimates of the parameters β_k in the objective function (1), representing the importance of the various terms mentioned in Section 2.3, are approximately normally distributed. Therefore these parameters can be tested by referring the t-ratio, defined as parameter estimate divided by standard error, to a standard normal distribution.

For actor-based models for network dynamics, information-theoretic model selection criteria have not yet generally been developed, although Koskinen (2004) presents some first steps for such an approach. Currently the best possibility is to use ad hoc stepwise procedures, combining forward steps (where effects are added to the model) with backward steps (where effects are deleted). The steps can be based on significance test for the various effects that may be included in the model. Guidelines for such procedures are the following. We prefer not to give a recipe, but rather a list of considerations that a researcher might have in mind when constructing a strategy for model selection.

1. Like in all statistical models, exclusion of one effect may mask the existence of another effect, so that pure forward selection may lead to overlooking some effects, and it is advisable to start with a model including all effects that are expected to be strong.

2. Fitting complicated models may be time-consuming and lead to instability of the algorithm, and a resulting failure to obtain good estimates. Therefore, forward selection is technically easier than backward selection, which is unfortunately at variance with the preceding remark.

3. The estimation algorithm (Snijders, 2001) is iterative, and the initial value can determine whether or not the algorithm converges. For relatively simple models, a simple standard initial value usually works fine. For complicated models, however, the algorithm may converge more easily if started from an initial value obtained as the estimate for a somewhat simpler model. Estimates obtained from a more complicated model by simply omitting the

deleted effects sometimes do not provide good starting values. Therefore, forward selection steps often work better from the algorithmic point of view than backward steps. This implies that, to improve the performance of the algorithm, it is advisable to retain copies of the parameter values obtained from good model fits, for use as possible initial values later on.

4. Network statistics can be highly correlated just because of their definition. This also implies that parameter estimates can be rather strongly correlated, and high parameter correlations do not necessarily imply that some of the effects should be dropped. For example, the parameter for the outdegree effect often is highly correlated with various other structural parameters. This correlation tells us that there is a trade-off between these parameters and will lead to increased standard errors of the parameter for the outdegree effect, but it is not a reason for dropping this effect from the model.

5. Parameters can be tested by a so-called score-type test without estimating them, as explained in Schweinberger (2008). Since estimating many parameters can make the algorithm instable, and in forward selection steps it may be necessary to have tests available for several effects to choose the most important one to include, the score-type tests can be very helpful in model selection. In this procedure, a model (null hypothesis) including significant and/or theoretically relevant parameters is tested against a model (alternative hypothesis) extended by one or several parameters one is also interested in. Under the null hypothesis, those parameters are zero. The procedure yields a test statistic with a chi-squared null distribution, along with standard normal test statistics for each separate parameter. The parameters for which a significant test result was found, then may be added to the model for a next estimation round.

6. It is important to let the model selection be guided by theory, subject-matter knowledge, and common sense. Often, however, theory and prior knowledge are stronger with respect to effects of covariates – e.g., homophily effects – than with respect to structure. Since a satisfactory fit is important for obtaining generalizable results, the structural side of model selection will of necessity often be more of an inductive nature than the selection of covariate effects. The newness of this method implies that we still need to accumulate more experience as to what is a 'satisfactory' fit, and how complicated models should be in practice.

7. Among the structural effects, the outdegree and reciprocity effect should be included by default. In almost all longitudinal social network data sets, there also is an important tendency toward transitivity (Davis 1970). This should

be modeled by one, or several, of the transitivity-related effects described above.

8. Often, there are some covariate effects which are predicted by theory; these may be control effects or effects occurring in hypotheses. It is good practice to include control effects from the start. Non-significant control effects might be dropped provisionally and then tested again as a check in the presumed final model; but one might also retain control effects independent of their statistical significance. We do not think there are unequivocal rules whether or not to include from the start the effects representing the main hypotheses in a given study.

9. The degree-based effects (popularity, activity, assortativity) can be important structural alternatives for actor covariate effects, and can be important network-level alternatives for the triad-level effects. It is advisable, at some moment during the model selection process, to check these effects; note that the square-root specification usually works best.

10. If the data have three or more waves and the model does not include time-changing variables, then the assumption is made that the time dynamics is homogeneous, which will lead to smooth trajectories of the main statistics from wave to wave. It is good as a first general check to consider how average degree develops over the waves, and if this development does not follow a rather smooth curve (allowing for random disturbances), to include time-varying variables that can represent this development. Another possibility is to analyze consecutive pairs of waves first, which will show the extent of inhomogeneity in the process (cf. the example in Snijders, 2005).

11. The model assumes that the 'rules for network change' are the same for all actors, except for differences implied by covariates or network position. This leaves only moderate room for outlying actors, such as are indicated by relatively very large outdegrees or indegrees. Very high or very low outdegrees or indegrees should be 'explainable' from the model specification; if they are only explainable from earlier observations ('path dependence'), they will have a tendency to regress toward the mean. The model specification may be able to explain outliers by covariates identifying exceptional actors, but also by degree-related endogenous effects such as the self-reinforcing Matthew effect mentioned above. It is good as a first general check to inspect the indegrees and outdegrees for outliers. If there are strong outliers then it is advisable to seek for actor covariates which can help to explain the outlying values, or to investigate the possibility that degree-related effects, explainable or at least interpretable from a theoretical point

of view, may be able to represent these outlying degrees. If these cannot be found, then one solution is to use dummy variables for the actors concerned, to represent their outlying behavior. Such an ad hoc model adaptation, which may improve the fit dramatically, is better than the alternative of working with a model with a large unmodeled heterogeneity. In case there is a theoretical argument to expect certain outliers, these can be pointed out by including a dummy variable, but of a different kind. In contrast to the former type of outliers, the latter one is expected, so should be captured in advance by a covariate. If one is not capable to make a difference between the two types, one has to rely on the ad hoc model adaptation.

3.3. Example: friendship dynamics

By way of example, we analyze the evolution of a friendship network in a Dutch school class. The data were collected between September 2003 and June 2004 as part of a study reported in Knecht (2008). The 26 students were followed over their first year at secondary school during which friendship networks as well as other data were assessed at four time points at intervals of three months. There were 17 girls and 9 boys in the class, aged 11-13 at the beginning of the school year. Network data were assessed by asking students to indicate up to twelve classmates which they considered good friends. The average number of nominated classmates ranged between 3.6 and 5.7 over the four waves, showing a moderate increase over time. Jaccard coefficients for similarity of ties between waves are between .4 and .5, which is somewhat low (reflecting fairly high turnover) but not too low.

Some data were missing due to absence of pupils at the moment of data collection. This was treated by ad-hoc model-based imputation using the procedure explained in Huisman and Steglich (2008). One pupil left the classroom. Such changes in network composition can also be treated by the methods of Huisman and Snijders (2003), but this simple case was treated here by using *structural zeros*: starting with the first observation moment where this pupil was not a member of the classroom any more, all incoming and outgoing tie variables of this pupil were fixed to zero and not allowed to change in the simulations.

Considering point 1 above, effects known to play a role in friendship dynamics, such as basic structural effects and effects of basic covariates, are included in the baseline model. From earlier research, it is known that friendship formation tends to be reciprocal, shows tendencies towards network closure, and in this age group is strongly segregated according to the sexes. The model includes, for each of these tendencies, effects corresponding to these expectations. Structural effects included are reciprocity; transitive triplets and transitive ties, measuring transitive closure that is compatible with an informal local hierarchy in the friendship net-

work; and the three-cycles effect measuring anti-hierarchical closure. Homophily based on the sexes is included as the same sex effect. All variables are centered. For example, the dummy variable for sex (boys = 1, girls = 0) has mean 0.346 (9 boys and 17 girls), which leads to the centered values v_i = –0.346 for girls and v_i = 0.654 for boys.

As exogenous control variables, we include sender and receiver effects of sex, and a dyadic covariate indicating friendship in primary school reflecting relationship history. In addition, several degree-related endogenous effects are included as control effects: in- and outdegree-related popularity, and outdegree-related activity, explained above. Estimates for this model are given in Table 1 as Model 0. All calculations were done using Siena version 3.2 (Snijders et al., 2008).

The parameters reported for the rate function in periods 1-3 are defined in the simulation model as the expected frequencies, between successive waves, with which actors get the opportunity to change a network tie. For these parameters no p-values are given in the tables, as testing that they are zero is meaningless (if these would be zero there would be no change at all). These estimated rate parameters will be higher than the observed numbers of changes per actor, however, because in the model an actor may get the opportunity to change a tie but choose not to make any change, and because actors may add a tie during the simulations, and withdraw the same tie before the next observation moment.

Table 1 about here

This analysis confirms, for this data set, several of the known properties of friendship networks: there is a high degree of reciprocity, as seen in the significant *reciprocity* parameter; there is segregation according to the sexes, as seen in the significant *same sex* parameter; there is an almost equally strong effect of having been friends at primary school already, and there is evidence for transitive closure, as seen in the significant effects of *transitive triplets* and *transitive ties*. A direct comparison of the size of parameter estimates is possible, given that they occur in the same linear combination in the objective function, but it should be kept in mind that these are unstandardized coefficients. Other significant effects are the negative *3-cycles* parameter, which indicates that the tendencies toward closure are not completely egalitarian (as one might have thought based on the reciprocity parameter), but do show some evidence for local hierarchization in the network. This also is suggested by the marginally significant negative effect of the outdegree-related popularity which indicates that active pupils, i.e., those who nominate particularly many friends, are less likely to be chosen as friends – this could be a status effect negatively associated with nomination activity. Also significant is the sender effect of sex (*sex (M) ego*), which in our coding of the variable means that the boys tend to be more active in the classroom friendship network than the girls.

Rate parameters, finally, suggest that the amount of friendship change seems to peak in the second period (perhaps due to a higher friendship turnover after the Christmas break) and slow down towards the end of the school year. These differences are small, however. The same descriptive conclusion can be drawn also by inspecting the observed amounts of change, without needing to refer to a statistical model.

In a subsequent model (Model 1 in Table 1), more parsimony is obtained by eliminating the non-significant effects in a backward selection procedure. The sex alter effect was retained in spite of its non-significance, because the three sex-related effects belong together as a representation of sex-related friendship preferences. One by one, the least significant of the insignificant effects were dropped from the model. While doing so, score-type tests were made for the earlier omitted parameters (now constrained to zero) to check whether the parameter does not become significant upon dropping other effects from the model. This is possible in models with correlated effects like ours, but it did not occur for our data set. Estimates of Model 1 give the same qualitative results as those of Model 0. The parameters dropped due to insignificance were the outdegree-related activity effect (suggesting that the value for ego of an individual friendship does not depend on how many other friends the friend currently has) and the indegree-related popularity effect (suggesting that receiving many friendship nominations is not a self-reinforcing process).

3.4. Parameter interpretation

For the general understanding of the numerical values of the parameters, it may be kept in mind that the parameters β_k in the objective function are unstandardized coefficients of the statistics of which the mathematical formulae are given in the appendix.

The parameters in the objective function can be interpreted in two ways. In the first place, by interpreting this function as the "attractiveness" of the network for a given actor. For getting a feeling of what are small and large values, it may be noted (see the interpretation in terms of myopic optimization in Snijders, 2001) that the objective functions are used to compare how attractive various different tie changes are, and for this purpose random disturbances are added to the values of the objective function with standard deviations equal[2] to 1.28.

The objective function is a weighted sum of effects $s_{ik}(x)$; their mathematical definitions are given in the appendix. In most cases the contribution of a single tie variable x_{ij} is just a simple component of this formula.

For example, consider the actor variable sex, denoted as V, and originally

[2]More exactly, the value is $\sqrt{\pi^2/6}$, the standard deviation of the Gumbel distribution.

with values 1 for girls and 2 for boys. All variables are centered. The global mean of this variable is 1.346 (9 boys and 17 girls), which leads to the centered values $v_i = -0.346$ for girls and $v_i = 0.654$ for boys. For this variable the model includes the 'ego' effect, the 'alter' effect, and the 'same' effect. Let us denote the parameters by $\beta_{\rm e}$, $\beta_{\rm a}$, and $\beta_{\rm s}$. Then, using the formulae in the appendix, the joint contribution of these V-related effects to the objective function is

$$\beta_{\rm e} \sum_j x_{ij}\, v_i \;+\; \beta_{\rm a} \sum_j x_{ij}\, v_j \;+\; \beta_{\rm s} \sum_j x_{ij}\, I\{v_i = v_j\}$$

where $I\{v_i = v_j\} = 1$ if $v_i = v_j$, and 0 otherwise. This means that the contribution of the single tie x_{ij} to the objective function, considering only the sex-related effects, is given by

$$\beta_{\rm e}\, v_i \;+\; \beta_{\rm a}\, v_j \;+\; \beta_{\rm s}\, I\{v_i = v_j\} \;=\; 0.35\, v_i \;+\; 0.10 v_j \;+\; 0.49\, I\{v_i = v_j\}$$

Substituting the values -0.346 for females and 0.654 for males yields the following table.

	alter	
ego	F	M
F	0.33	–0.06
M	0.19	0.78

This table shows that girls as well as boys prefer friendships with same-sex alters, but for boys the difference is more pronounced than for girls.

A second interpretation is that when actor i has the opportunity to make a change in her outgoing ties (where no change also is an option), and x_a and x_b are two possible results of this change, then $f_i(x_b, \beta) - f_i(x_a, \beta)$ is the log odds ratio for choosing between these two alternatives – so that the ratio of the probability of x_b and x_a as next states is

$$\exp\big(f_i(x_b, \beta) - f_i(x_a, \beta)\big) \;=\; \frac{\exp\big(f_i(x_b, \beta)\big)}{\exp\big(f_i(x_a, \beta)\big)}\;.$$

Note that, when the current state is x, the possibilities for x_a and x_b are x itself (no change), or x with one extra outgoing tie from i, or x with one less outgoing tie from i. Explanations about log odds ratios can be found in texts about logistic regression and loglinear models, e.g., Agresti (2002). A further elaborated example of this is given below in Section 4.2.

4. More complicated models

This section treats two generalizations of the model sketched above.

4.1. Differential rates of change: the rate function

Depending on actor attributes or on positional characteristics such as indegree or outdegree, actors might change their ties at differential frequencies. This can be the case, e.g., in networks between organizations with clear differences in degrees, where the outdegrees reflect the importance to the organizations of the network under study, and the resources they devote to positioning themselves in it. The average frequency at which actors get the opportunity to change their outgoing ties then is called the *rate function*, depending on attributes and network position of the actors.

Model 2 in Table 1 gives an example of such an analysis. It extends Model 1 by adding an effect of sex on the rate function. The estimated negative effect indicates that in the data set under study, boys change their network ties less frequently than girls, but the difference is not significant ($p = 0.13$).

To interpret the parameter values, one should know that a so-called exponential link function is used (Snijders, 2001; Snijders et al., 2008), which means that the variables have an effect on the rate function after an exponential transformation, with a multiplicative effect. For example, the parameter estimate of –0.42 for the effect of sex on the rate function implies that the estimated rate function is the base rate multiplied by $\exp(-0.42\, v_i)$. Recall that the values of the variable 'sex' are, centered, $v_i = -0.346$ for females and $v_i = 0.654$ for males. Thus, for period 1, for girls the expected number of opportunities for change is $9.69 \times \exp(-0.42 \times (-0.346)) = 11.2$, and for boys it is $9.69 \times \exp(-0.42 \times (0.654)) = 7.4$. The difference seems rather large but is not significant in view of the small sample size.

4.2. Differences between creating and terminating ties: the endowment function

In the treatment given above, terminating a tie is just the opposite of creating one. This is not always a good representation of reality. It is conceivable, for example, that the loss when terminating a reciprocal tie is greater than the gain in creating one; or that transitive closure works especially for the creation of new ties, but hardly guards against termination of existing ties. This can be modeled by having two components of the objective function: the *evaluation function*, which considers only the network that will be the case as a result of the change to be made; and the *endowment function*, which is a component that operates only for the termination of ties and not for their creation. Everything discussed above about

the objective function concerned the evaluation function – in other words, in those discussions and the example, the endowment function was nil. The endowment function gives contributions to the objective function that do not play a role when creating ties, but that are lost when dissolving ties.

Model 3 in Table 1 gives the results of an analysis that includes, in addition to the effects of Model 1, also an endowment effect related to *reciprocity*. It was estimated as significant and positive, while the corresponding evaluation function effect of reciprocity dropped in size and significance. To interpret this result, jointly consider the reciprocity evaluation effect with parameter 0.71 and the reciprocity endowment effect with parameter 1.42. The contribution of a tie being reciprocated then is 0.71 for the creation of the tie and 0.71 + 1.42 = 2.13 against the termination of the tie. Thus, reciprocity here is more important against terminating friendships – that is, for maintaining friendships – than for creating friendships.

To elaborate this example, consider how the friendship choices of a girl towards other girls depend on reciprocity. Suppose that actor i can change one of her ties, while there are two girls j_1 and j_2, both of them choosing i as a friend, and two others j_3 and j_4 not choosing i. In addition, suppose that currently i chooses j_1 and j_3 as friends, but not the other two. Assume finally (artificially, for the sake of explanation) that these four girls do not choose each other and further also are isolated from i's network so that other structural effects besides reciprocity do not matter. Since the actor variable 'sex' has centered value $v_i = -0.346$ for girls, the parameter estimates for Model 3 give as the total contribution of the three sex-related effects for girl-girl ties $0.41\, v_i + 0.16\, v_j + 0.56\, I\{v_i = v_j\} = (0.41 + 0.16) \times (-0.346) + 0.56 = 0.36$. With the outdegree effect of –1.59, this yields $-1.59 + 0.36 = -1.23$ as the basic contribution of a tie to the evaluation function.

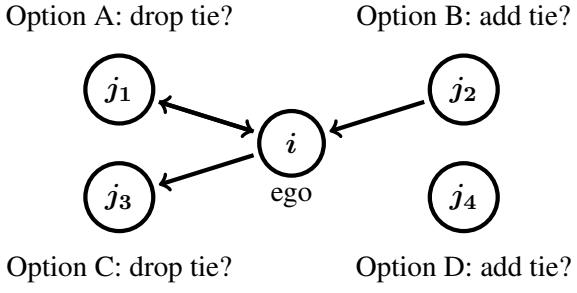

Figure 2. Four options for actor i.

When girl i can change a tie variable, using this value of –1.23 for the combined effect of outdegree and the three sex-related effects for a girl-girl tie, five of the options for i are the following:

(A): drop reciprocated friendship tie to j_1 : –(–1.23) – 2.13 = – 0.90;
(B): reciprocate friendship tie from j_2 : –1.23 + 0.71 = –0.52;
(C): drop non-reciprocated friendship tie to j_3 : –(–1.23) = 1.23;
(D): initiate friendship tie to j_4 : –1.23;
(E): do nothing: 0.0.

Since these are contributions to logarithms of probabilities, the proportionality factors between the probabilities of these events must be calculated as the exponential transformations of these values, which are, respectively, $e^{-0.90} = 0.41$, 0.59, 3.42, 0.29, and 1.

These are the relative probabilities of changes toward any given other girl. One should note, however, that there may be different numbers of the four cases (A-D) for a given girl, and the probability of severing *any* reciprocated tie, or of creating *any* non-reciprocated tie, depends also on these numbers for the 'ego' girl under consideration. Since the friendship network is sparse, with average degrees between 3.6 and 5.7, the cases of type (D) will be most numerous. Consider, for instance, a girl with 3 mutual girlfriends (A), who has 2 non-reciprocated friendships to girls (C), 1 other girl who mentions her as a friend without reciprocation (B), and 10 girls without a friendship either way (D). Suppose that in addition she has no friendships with any of the 9 boys, and denote the option of establishing a friendship to a boy by (F). Retain the unrealistic simplifying assumption that all her network members are mutually unrelated, also after adding one hypothetical new friend, so that transitivity does not influence probabilities of change. The baseline value of a tie from a girl to a boy is $-1.59 + 0.41 \times (-0.346) + 0.16 \times 0.654 = -1.63$, with exponential transform 0.20. Taking into account the fact that the number of opportunities for options (A) to (F) are 3, 1, 2, 10, 1, and 9, the six proportionality factors have to be divided by the denominator $(3 \times 0.41) + (1 \times 0.59) + (2 \times 3.42) + (10 \times 0.29) + (1 \times 1) + (9 \times 0.20) = 14.36$. Thus, for this girl, the probabilities are:

(A): of dropping any of the three reciprocated friendship ties: $(3 \times 0.41)/14.36 = 0.09$;
(B): of reciprocating the incoming friendship tie: $(1 \times 0.59)/14.36 = 0.04$;
(C): of dropping one of the non-reciprocated friendship ties: $(2 \times 3.42)/14.36 = 0.48$;
(D): of initiating some new friendship tie to a girl: $(10 \times 0.29)/14.36 = 0.20$;
(E): of doing nothing: 0.07;
(F): and of extending a new friendship tie to a boy: $(9 \times 0.20)/14.36 = 0.13$.

Thus, in line with theories about reciprocation such as balance theory, the probability is slightly larger than 0.5 that the proportion of reciprocity in friendships will be increased. There are many random influences, however, that would decrease reciprocity – but most of these are proposals of new ties which could be

seen by the other party as an invitation toward future reciprocation.

5. Dynamics of networks and behavior

Social networks are so important also because they are relevant for behavior and other actor-level outcomes: related actors may influence one another (e.g., Friedkin, 1998), and ties will be selected in part based on the similarity between ego and potential relational partners (homophily, see McPherson, Smith-Lovin, and Cook, 2001). This means that not only is the network changing as a function of itself and of the actor variables, but likewise the actor variables are changing as a function of themselves and of the network. We use the term *behavior* as shorthand for endogenously changing actor variables, although these could also refer to attitudes, performance, etc.; there could be one or more of such variables. It is assumed here that the behavior variables are ordinal discrete variables, with values 1, 2, etc., up to some maximum value, for instance, several levels of delinquency, several levels of smoking, etc. The dependence of the network dynamics on the total network-behavior configuration will be also called the *social selection* process, while the dependence of the behavior dynamics on the total network-behavior configuration will be called the *social influence* process. Both social influence and social selection can lead to similarity between tied actors, which is often observed. A fundamental question then is whether this similarity is caused mainly by influence or mainly by selection, as discussed by Ennett and Bauman (1994) for smoking behavior and Haynie (2001) for delinquent behavior.

This combination of selection and influence can be modeled by an extension of the actor-based model to a structure where the dependent variables consist not only of the tie variables but also of the actors' behavior variables, as specified in Snijders, Steglich and Schweinberger (2007) and Steglich, Snijders and Pearson (2009). Of course there usually will be, in addition, also exogenous actor and/or dyadic variables in the role of independent variables.

The assumptions for the actor-based model for the dynamics of networks and behavior are extensions of the assumptions for network dynamics. The extended formulations are as follows, given without the background explanations which were given above and which apply also for this case.

1. As above, the underlying time parameter is continuous.

2. The changing system consisting of network and behavior is the outcome of a Markov process. Thus, the probabilities of change of the network as well as those of the behavioral variables depend, at each moment, on the current combination of network structure and behavior variables for all actors.

3. At a given moment either one probabilistically selected actor may change a tie, or one actor may change his/her behavior by going one unit up or down (recall that the behavior variables are assumed to be integer-valued). This excludes coordination between changes in the network and in the behavior.

   The fact that changes in behavior are assumed to be by one unit in a single time point imply that a 'natural' application of the model requires that the total number of ordinal scale values is not too large; in practice applications mostly have had two to five, and sometimes up to ten, scale values.

4. The actors control their outgoing ties as well as their own behavior. This is meant not in the sense of conscious control, but in the sense that the explanation of the actor's outgoing ties and behavior is based in the actor and the structural and other limitations provided by the actor and his/her social context.

5. The moments were actors get the opportunity for a tie change or a behavior change are modeled as distinct processes, so these are governed by a priori unrelated parameters.

6. There are distinct processes also for tie changes and behavior changes, conditional on the possibility to make the respective type of change, so these are governed by a priori unrelated parameters.

The changes in behavior depend on an objective function similar to the objective function for network changes. However, this function will be different because it needs to represent primarily the actor's behavior rather than his/her network position, and because choices of behavior changes may be framed differently from choices of tie changes, depending on different goals and restrictions.

The model assumptions imply that the dependent behavior variable will change endogenously during the simulations, representing the endogenous social influence process. Since the network and the behavior variables both influence the dynamics of the network ties and of the actors' behavior, the sequence of changes in the network and in the behavior, reacting on each other, generates a mutual dependence between the network dynamics and the behavior dynamics.

5.1. The objective function for behavior

We only consider models where increasing the behavior variable has just the opposite effect of decreasing it, and the objective function for behavior is the same as the evaluation function (a separate endowment function is not considered). The

objective, or evaluation, function can be represented, analogously to (1), as

$$f_i^Z(\beta, x, z) = \sum_k \beta_k^Z \, s_{ki}^Z(x, z) \; , \tag{3}$$

where $s_{ki}^Z(x, z)$ are functions depending on the behavior of the focal actor i, but also on the behavior of his network partners, his network position, etc. The strength of the effects of these functions on behavior choices are represented by the parameters β_k^Z. The superscript Z is used to distinguish the effects and parameters for behavior change from those for network change (which could be given the superscript X). The main possible terms of the evaluation function are as follows.

Basic shape effects

We first discuss basic tendencies determining behavior change that are independent of actor attributes and network position. A baseline definition for the evaluation function will be a curve, depending on the actor's own behavior z_i, that can be loosely interpreted as the relative preference for the specific value z_i of the behavior. The term 'prefer' should be taken with much reservation, as a shorthand 'as-if' term – we could just as well see this, e.g., as a matter of constraints. When the behavior variable is dichotomous, then a linear function suffices, as each function of two values can be represented by a linear function; but for three or more possible values a unimodal 'preference' function will often be reasonable, so that a specification will be required that allows the function to be curvilinear. Thus, in Figure 3, a simple evaluation function is drawn for a behavior variable with range 1–4, which is maximal at the value $z = 2$, indicating that when actors have a possibility for change they will be drawn toward the value $z = 2$: if their current value for Z is higher than 2 then the probability is higher that they will decrease their value, if the current value is lower than 2 then the probability is higher that they will increase their value of Z. To represent this mathematically, a quadratic function can be used. The linear and quadratic coefficients in this function are called the *linear shape effect* and the *quadratic shape effect*. Note that the latter is superfluous for a dichotomous behavior variable.

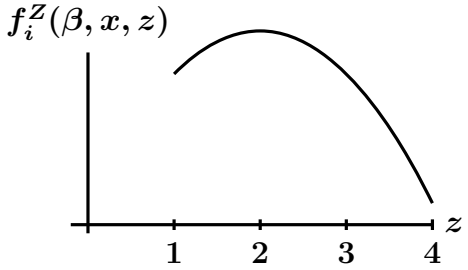

Figure 3. Basic shape of the evaluation function for behavior; in this case, with the maximum at $z = 2$.

The quadratic shape effect can also be called the effect of Z on itself, and is a kind of feedback effect. When this parameter is negative there is negative feedback, or a self-correcting mechanism: when the value of the actor's behavior increases, the further push toward *higher* values of the behavior will become smaller and when it decreases, the further push toward *lower* values of the behavior will become smaller. Conversely, when the coefficient of the quadratic term is positive, the feedback will be positive, so that changes in the behavior are self-reinforcing. This can be an indication of addictive behavior. Negative values, or values not significantly different from 0, are more often seen than positive values.

When the coefficient for the shape effect is denoted β_1^Z and the coefficient for the quadratic shape effect is β_2^Z, then the total contribution of these two effects is $\beta_1^Z z_i + \beta_2^Z z_i^2$. With a negative coefficient β_2^Z, this is a unimodal preference function. The maximum then is attained for $z_i = -2\,\beta_1^Z / \beta_2^Z$, or more precisely, the integer value within the prescribed range that is closest to this number. (Of course additional effects will lead to a different picture; but if the additional effects are linear as a function of z_i in the permitted range – which can be checked from the formulae in the appendix, and which is not the case for all similarity effects as defined below! –, this will change the location of the maximum but not the unimodal shape of the function.)

Influence and position-dependent effects

The tendencies expressed in the shape parameters affect every actor in the same way, irrespective of his/her characteristics or network position. To capture social network effects, additional terms in the behavior evaluation function are needed, differentiating between actors on the basis of their network position and the behavior of the others to whom they are tied.

The actor-based model can represent social influence, i.e., influence from alters' behavior on ego's behavior, in various ways, because there are several different ways to measure and aggregate the influences from different alters. Three different representations are as follows.

1. The *average similarity effect*, expressing the preference of actors to be similar in behavior to their alters, in such a way that the total influence of the alters is the same regardless of the number of alters (i.e., ego's outdegree).

2. The *total similarity effect*, expressing the preference of actors to be similar in behavior to their alters, in such a way that the total influence of the alters is proportional to the number of alters.

3. The *average alter effect*, expressing that actors whose alters have a higher average value of the behavior, also have themselves a stronger tendency toward high values on the behavior.

The choice between these three will be made on theoretical grounds and/or on the basis of statistical significance.

In addition to this type of social influence, network position itself could also have an effect on the dynamics of the behavior. Indeed, the actor's outdegree or indegree may be terms in the objective function; in the case of positive parameters, this expresses that those who are more active (higher outdegree) or more popular (higher indegree) have a stronger tendency to display higher values of the behavior.

Effects of other actor variables

For each actor-dependent covariate as well as for each of the other dependent behavior variables (if any), a main effect on the behavior can be included, representing the influence of this actor variable on changes in the behavior. In addition, it is possible that such a variable will moderate the influence effect, leading to an interaction between the variable and the influence effect.

5.2. Specification of models for dynamics of networks and behavior

There is a natural advantage of the network part of the model over the behavior part in terms of the amount of information available on the two dimensions. For n actors, there are n measurements of the behavior variable, while there are as many as $n(n-1)$ measurements of network tie variables. In statistical terms, there will be less power to detect determinants of behavioral evolution than there is to detect determinants of network evolution. Therefore, backward model selection is not a good route to follow for specifying a model of network-behavior co-evolution: weak and unsystematic effects on behavioral change, when estimated, subtract from the ability to identify the stronger and more systematically occurring ones. Instead of starting with an extensive model, it is better to start with a small one, and proceed by way of forward model selection to arrive at a good model.

Since it is more difficult to estimate a model of network-behavior co-evolution than to estimate a model of only network evolution, it makes sense first to estimate a good model for the dynamics of only the network according to the approach of the preceding section, and to use this as a baseline for the network part of the co-evolution model.

It has become clear above that there are several specifications of the social selection part, such as Z-similarity and Z - ego $\times$ alter; like for any actor variable, also for behavior variable Z the ego and alter effects may be relevant. The similarity effect in the network dynamics part is directly interpretable, whereas the Z - ego $\times$ alter interaction effect needs also the ego and alter effects to be well interpretable. For the social influence part likewise there are several possible specifications, such as average similarity, total similarity, and average alter.

A difficulty is that we often have no clear theoretical clue as to which of these three specifications is better in a particular case; the alternative specifications are defined by effects that may be highly mutually correlated and therefore are not readily estimated jointly in the same model; and the power of detecting selection and influence will depend on the specification chosen. If we do have prior information as to the best specification, then it is preferable to work with this specification. If we do not have such information, and we wish to avoid the chance capitalization inherent in using the 'most significant' effect without taking into account that it was chosen exactly because it was the most significant, then we could proceed along something like the following lines. An example of this approach is given in the next subsection. This procedure uses score-type tests (Schweinberger, 2008) for several parameters simultaneously, which are chi-squared tests with number of degrees of freedom equal to the number of tested effects, and which have the advantage that parameters can be tested without estimating them.

1. Specify and estimate a baseline model '*BM*' for network-behavior co-evolution in which the network and behavior dynamics are independent; that is, the model for network dynamics contains no effects dependent on the behavior variable, and vice versa.

2. Choose a number of candidate social selection effects '*SEL*' and a number of candidate social influence effects '*INF*' on theoretical grounds, without considering the data.

3. Test the effects in the sets *SEL* and *INF* by score-type tests in the baseline model *BM*. This gives the statistical evidence about the existence of influence and selection, not controlling each effect for the other.

4. Select the effects that are individually most significant in either set, and denote these effects by *SEL1* and *INF1*.

(In this formulation, the effects are selected based on their significance when hypothesized to be added to the baseline model. Another possibility is to select the effects based on the following two models.)

5. To test influence effects while controlling for selection, estimate the model *BM + SEL1* (i.e., the baseline model extended with the most significant selection effect) and within this model test all of the effects *INF* jointly by a score-type test.

6. To test selection effects while controlling for influence, estimate the model *BM + INF1* and test all of the effects *SEL* jointly by a score-type test.

7. Finally, to estimate a model with influence and selection, estimate the model *BM + SEL1 + INF1*.

8. It often will be sensible to conduct some further checks to guard against the danger of overlooking important effects. This can be done again with score-type tests. Candidate effects to be checked include the indegree and outdegree effects on the behavior variable.

5.3. Example: dynamics of friendship and delinquency

Substantively, what we address is again the dynamics of the friendship network in the school class of 11-13 year old pupils investigated above, now co-evolving with their delinquency (Knecht, 2008). This variable is defined as a rounded average over four types of minor delinquency (stealing, vandalism, graffiti, and fighting), measured in each of the four waves of data collection. The five-point scale ranged from 'never' to 'more than 10 times', and the distribution is highly skewed, most students reporting no delinquency. In a range of 1–5, the mode was 1 at all four waves, the average rose over time from 1.4 to 2.0, and the value 5 was never observed.

The question addressed is, whether the data provide evidence for network influence processes playing a role in the spread of delinquency through the group defined by the classroom. Analyses were carried out by Siena version 3.17 (Snijders et al. 2008). Reported results all were taken from runs in which all 't-ratios for convergence' (see the Siena manual) were less than 0.1 in absolute value, indicating good convergence of the algorithm.

For the model selection we follow the steps laid out in the preceding section. The baseline model is the model which for the friendship dynamics is model 3 in Table 1, and for the delinquency dynamics includes the linear and quadratic shape effects and the effect of sex (as a control variable). The effects potentially modelling social selection based on delinquency (the set *SEL* in the preceding

section) are delinquency ego, delinquency alter, delinquency similarity, and delinquency ego $\times$ alter. Delinquency ego and delinquency alter are included here as control variables for delinquency ego $\times$ delinquency alter. The effects potentially modelling social influence with respect to delinquency (the set *INF*) are average similarity, total similarity, and average delinquency alter. Score tests for the selection and influence effects tested with the baseline model as the null hypothesis yielded the following p-values. For both parts of the model, first the results of the overall score test (of *SEL* and *INFL*, respectively) are given, and then the results for the separate degrees of freedom of which this test is composed.

Effect	p
Friendship dynamics	
overall test for social selection (4 d.f.)	< 0.001
delinquency ego	< 0.001
delinquency alter	0.32
delinquency similarity	< 0.001
delinquency ego $\times$ delinquency alter	0.02
Delinquency dynamics	
overall test for social influence (3 d.f.)	0.04
average similarity	0.03
total similarity	0.12
average delinquency alter	0.48

The overall tests show evidence for social selection ($p < .001$) as well as for social influence ($p < .05$), when these are not being controlled for each other. The separate tests suggest that strongest effects are delinquency similarity for the network dynamics and average similarity for the behavior dynamics. These are the effects denoted above as *SEL1* and *INF1*, respectively.

To test selection while controlling for influence, the baseline model was extended with influence operationalized as average similarity, and the four selection effects (delinquency ego, delinquency alter, delinquency similarity, and delinquency ego $\times$ delinquency alter) were jointly tested by a score-type test. The test result was highly significant, $p < 0.001$.

To test influence while controlling for selection, the baseline model was extended with selection operationalized as delinquency similarity, and the three influence effects (average similarity, total similarity, and average delinquency alter) were jointly tested by a score-type test. This led to $p = 0.04$. Thus, for this classroom there is clear evidence ($p < 0.001$) for delinquency-based friendship selection, and evidence ($p = 0.04$) for influence from pupils on the delinquent behavior of their friends.

To estimate a model incorporating selection and influence, the baseline model was extended with delinquency similarity for the network dynamics and average similarity for the behavior dynamics. This model is estimated and presented in Table 2. The delinquency ego and delinquency alter effects were also tested for inclusion in the network dynamics model, and the indegree and outdegree effects were tested for inclusion in the behavior dynamics model, but none of these were significant.

Table 2 about here

It can be concluded for this data set that there is evidence for delinquency-based friendship selection, expressed most clearly by the delinquency similarity measure; and for influence from pupils on the delinquent behavior of their friends, expressed best by the average similarity measure. The delinquent behavior does not seem to be influenced by sex. The model for network dynamics yields estimates that are quite similar to those for the model without the simultaneous delinquency dynamics, except that the reciprocity effect has shifted more strongly towards the endowment effect.

6. Cross-sectional and longitudinal modeling

For a further understanding of this actor-based model, it may be helpful to reflect about equilibrium and out-of-equilibrium social systems. Equilibrium is understood here not as a fixed state but as dynamic equilibrium, where changes continue but may be regarded as stochastic fluctuations without a systematic trend. This can be combined with discussing the relation between cross-sectional and longitudinal statistical modeling of social networks.

For cross-sectional modeling the exponential random graph model ('ERGM'), or p^* model, is similar to the model of this paper in its statistical approach to network structure (cf. Wasserman and Pattison 1996, Robins et al., 2007, and the references cited there). A telling characteristic of the ERGM is that the only feasible way to obtain random draws from such a probability distribution is to simulate a process longitudinally until it may be assumed to have reached dynamic equilibrium, and then take samples from the process (Snijders, 2002). Thus, the ERGM can be best understood as a model of a process in equilibrium. If we would have longitudinal data from a process in dynamic equilibrium, then modeling them by the approach of this paper would give roughly the same results as modeling its cross-sections by an ERGM. It would not be exactly the same because the ERGM is not actor-based; a tie-based version of our longitudinal model (Snijders, 2006) is possible, which does correspond exactly to the ERGM. On the contrary, if we would have cross-sectional data which may be assumed to have been observed

far from an equilibrium situation, then it is difficult to see what would be the precise meaning of results of an ERGM analysis, based as this is on an implicit equilibrium assumption.

If one has observed a longitudinal network data set of which the consecutive cross-sections have similar descriptive properties – no discernible trends or important fluctuations in average degree, in proportion of reciprocated ties, in proportion of transitive closure among all two-paths, etc. –, then it would be a mistake to infer that the development is not subject to structural network tendencies just because the descriptive network indices are stationary. For example, if the network shows a persisting high extent of transitive closure, in a process which is dynamic in the sense that quite some ties are dissolved while other new ties appear, then it must be concluded that the dynamics of the network contains an aspect which sustains the observed extent of transitive closure against the random influences which, without this aspect, would make the transitive closure tend to attenuate and eventually to disappear.

The longitudinal actor-based model is more general than the ERGM in that it does not require that the observed process be in equilibrium. Given a sequence of consecutively observed networks, if one were to make an analysis of the first one by an ERGM and of the further development by an actor-based model, then in theory it is possible to obtain opposite results for these two analyses, and this would point toward a non-equilibrium situation. For example, it would be possible that the first observed network shows no transitive closure at all, but the dynamics does show a transitivity effect; then over time the extent of transitive closure would increase, perhaps to reach some dynamic equilibrium later on. Conversely, it is possible that there is a strong transitivity effect at the first observation but no transitivity in the longitudinal model, which means that the observed extent of transitive closure in repeated cross-sectional analyses would eventually peter out to nil.

The advantage of longitudinal over cross-sectional modeling is that the parameter estimates provide a model for the rules governing the dynamic change in the network, which often are better reflections of social rules and regularities than what can be derived from a single cross-sectional observation. This also is an argument for not modeling the first observation but using it only as an initial condition for the network dynamics (as mentioned at the end of Section 2.1). In many situations the first observation cannot be regarded as coming from a process in equilibrium, and then it is unclear what the first observation by itself can tell us about social rules and regularities.

7. Discussion

This paper has given a tutorial introduction in the use of actor-based models for analyzing the dynamics of directed networks – expressed by the usual format of a directed graph – and of the joint interdependent dynamics of networks and behavior – where 'behavior' is an actor variable which may refer to behavior, attitudes, performance, etc., measured as an ordinal discrete variable. The purpose of these models is to be used to test hypotheses concerning network dynamics and represent the strength of various tendencies driving the dynamics by estimated parameters. To be useful in this way for statistical inference, the models must be able to give a good representation of the dependencies between network ties, and between network positions and behavior of the actors. The models also have to contain parameters that can express theoretical considerations about tendencies driving network change. Further, the models must be flexible enough to represent several different explanations of change – which may be competing but also potentially complementary – and test these against each other, or controlling for each other. These goals are accomplished by the stochastic actor-driven model in which the central object of modeling is the objective function (1), (3), analogous to the linear predictor in generalized linear modeling (e.g. Long, 1997).

These models can be estimated by software called *Siena* ('Simulation Investigation for Empirical Network Analysis'), of which the manual is Snijders et al. (2008). Program and documentation can be downloaded free from the Siena web-page, http://www.stats.ox.ac.uk/siena/. Some examples of applications of this model are Van de Bunt, Van Duijn, and Snijders (1999) and Burk, Steglich, and Snijders (2007). Further applications of these models are presented in some of the papers in this special issue.

These models are relatively new, and more complicated than many other statistical models to which social scientists are used. Applications are starting to be published, but more experience is needed to get a better understanding of their applicability and the interpretation of the results. Especially important will be the further development of ways to assess the goodness of fit of these models and to diagnose what in the data-model combination may be mainly responsible for a possible lack of fit. Another issue is the assessment of the robustness of results with respect to misspecified models, and the development of other models for network dynamics which may serve as potential alternatives for cases where the models presented here do not fit to a satisfactory degree. A case in point would be the development of models permitting a more elaborate temporal dependence than Markovian dependence. All this should lead to better knowledge about how to fit longitudinal models to network data, and hence to more reliable results. In this tutorial we have tried to represent the current knowledge about the specification of longitudinal network models in a concise but more or less complete way, but

we hope that this knowledge will expand rapidly.

Various extensions of the model are the topic of recent and current research. Models for the dynamics of non-directed networks, e.g., alliance networks, have been developed and were applied in Checkley and Steglich (2007) and van de Bunt and Groenewegen (2007). Extensions to valued ties are in preparation. Other estimation procedures have been proposed: Bayesian inference by Koskinen and Snijders (2007) and Schweinberger (2007), Maximum Likelihood estimation by Snijders, Koskinen and Schweinberger (2009). All these developments will be tracked at the Siena website mentioned above.

Appendix

This appendix contains some formulae to support the understanding of the verbal descriptions in the paper.

Objective function

When actor i has the opportunity to make a change, he/she can choose between some set $\mathcal{C}$ of possible new states of the network. Normally this will be set consisting of the current network and all other networks where one outgoing tie variable of i is changed. The probability of going to some new state x in this set is given by

$$\frac{\exp\big(f_i(\beta, x)\big)}{\sum_{x' \in \mathcal{C}} \exp\big(f_i(\beta, x')\big)} \,. \tag{4}$$

In words: the probability that an actor makes a specific change is proportional to the exponential transformation of the objective function of the new network, that would be obtained as the consequence of making this change. Similarly, when actor i can make a change in the behavior variable z and the current value is z_0, then the possible new states are $z_0 - 1, z_0$, and $z_0 + 1$ (unless the first or last of these three falls outside the range of the behavior variable). Denoting this allowed set also by $\mathcal{C}$, the probability of going to some new state z in this set is given by

$$\frac{\exp\big(f_i^Z(\beta^Z, x, z)\big)}{\sum_{z' \in \mathcal{C}} \exp\big(f_i^Z(\beta, x, z')\big)} \,. \tag{5}$$

These are the same formulae as used in multinomial logistic regression.

Effects

Some formulae for effects $s_{ik}(x)$ are as follows. Replacing an index by a $+$ sign denotes summation over this index. Exogenous actor covariates are denoted by v_i and dyadic covariates by w_{ij}.

reciprocity	$\sum_j x_{ij}\, x_{ji}$	(6)
transitive triplets	$\sum_{j,h} x_{ih}\, x_{ij}\, x_{jh}$	(7)
transitive ties	$\sum_h x_{ih}\, \max_j(x_{ij}\, x_{jh})$	(8)
three-cycles	$\sum_{j,h} x_{ij}\, x_{jh}\, x_{hi}$	(9)
balance	$\frac{1}{n-2} \sum_{j=1}^{n} x_{ij} \sum_{\substack{h=1 \\ h\neq i,j}}^{n} (b_0 - \mid x_{ih} - x_{jh} \mid)$,	(10)

where b_0 is the mean of $\mid x_{ih} - x_{jh} \mid$;

indegree popularity (sqrt)	$\sum_j x_{ij} \sqrt{x_{+j}}$	(11)
outdegree popularity (sqrt)	$\sum_j x_{ij} \sqrt{x_{j+}}$	(12)
indegree activity (sqrt)	$\sqrt{x_{+i}}\, x_{i+}$	(13)
outdegree activity (sqrt)	$x_{i+}^{1.5}$	(14)

out-outdegree assortativity (sqrt)	$\sum_j x_{ij} \sqrt{x_{i+}\, x_{j+}}$	(15)
(other assortativity effects similar)		
V - ego	$\sum_j x_{ij}\, v_i$	(16)
V - alter	$\sum_j x_{ij}\, v_j$	(17)
V - similarity	$\sum_j x_{ij}\, (\mathrm{sim}_{ij} - \overline{\mathrm{sim}})$,	(18)

where $\mathrm{sim}_{ij} = \big(1 - |v_i - v_j|/\Delta\big)$, with $\Delta = \max_{ij} |v_i - v_j|$;

same V $$\sum_j x_{ij}\, I\{v_i = v_j\}\ , \tag{19}$$

where $I\{v_i = v_j\} = 1$ if $v_i = v_j$, and 0 otherwise;

V - ego $\times$ alter $$\sum_j x_{ij}\, v_i\, v_j \tag{20}$$

dyadic covariate W $$\sum_j x_{ij}\, w_{ij} \tag{21}$$

dyadic cov. W $\times$ reciprocity $$\sum_j x_{ij}\, x_{ji}\, w_{ij} \tag{22}$$

actor cov. V $\times$ transitive triplets $$v_i \sum_{j,h} x_{ij}\, x_{jh}\, x_{ih}\ . \tag{23}$$

Some formulae for behavior effects $s_{ik}^{Z}(x, z)$ are the following.

linear shape $$z_i \tag{24}$$

quadratic shape $$z_i^2 \tag{25}$$

outdegree $$z_i\, x_{i+} \tag{26}$$

indegree $$z_i\, x_{+i} \tag{27}$$

average similarity $$x_{i+}^{-1} \sum_j x_{ij}(\mathrm{sim}_{ij}^{z} - \overline{\mathrm{sim}^{z}})\ , \tag{28}$$

where $\mathrm{sim}_{ij}^{z} = \big(1 - |z_i - z_j|/\Delta_Z\big)$ with $\Delta_Z = \max_{ij} |z_i - z_j|$

total similarity $$\sum_j x_{ij}(\mathrm{sim}_{ij}^{z} - \overline{\mathrm{sim}^{z}}) \tag{29}$$

average alter $$z_i\big(\sum_j x_{ij}\, z_j\big)/\big(\sum_j x_{ij}\big) \tag{30}$$

main effect covariate V $$z_i\, v_i\ . \tag{31}$$

References

Agresti, A. 2002. Categorical Data Analysis. Wiley, New Jersey.

Barabási, A.L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.

Bala, V., Goyal, S., 2000. A noncooperative model of network formation. Econometrica 68, 1181–1229.

Batagelj, V., Bren, M., 1995. Comparing resemblance measures. Journal of Classification 12, 73–90.

Bearman, P.S., 1997. Generalized exchange. American Journal of Sociology 102, 1383–1415.

Borgatti, S.P., Foster, P.C., 2003. The network paradigm in organizational research: A review and typology. Journal of Management 29, 991–1013.

Brass, D.J., Galaskiewicz, J., Greve, H.R., Tsai, W., 2004. Taking stock of networks and organizations: a multilevel perspective. Academy of Management Journal 47, 795–817.

Burk, W.J., Steglich, C.E.G., Snijders, T.A.B., 2007. Beyond dyadic interdependence: Actor-oriented models for co-evolving social networks and individual behaviors. International Journal of Behavioral Development 31, 397–404.

Burt, R.S., 1982. Toward a Structural Theory of Action. Academic Press, New York.

Checkley, M., Steglich, C.E.G., 2007. Partners in Power: Job Mobility and Dynamic Deal-Making. European Management Review 4, 161–171.

Davis, J.A., 1970. Clustering and hierarchy in interpersonal relations: testing two graph theoretical models on 742 sociomatrices. American Sociological Review 35, 843–852.

de Federico de la Rúa, A., 2004. L'Analyse longitudinale de réseaux sociaux totaux avec SIENA – Méthode, discussion et application. BMS, Bulletin de Méthodologie Sociologique 84, 5–39.

Doreian, P., Stokman, F.N. (Eds.), 1997. Evolution of Social Networks. Gordon and Breach Publishers, Amsterdam.

Ennett, S.T., Bauman, K.E., 1994. The contribution of influence and selection to adolescent peer group homogeneity: The case of adolescent cigarette smoking. Journal of Personality and Social Psychology 67, 653–63.

Friedkin, N.E., 1998. A Structural Theory of Social Influence. Cambridge University Press, Cambridge.

Haynie, D.L., 2001. Delinquent peers revisited: Does network structure matter? American Journal of Sociology 106, 1013–57.

Hedström, P., 2005. Dissecting the Social: On the Principles of Analytical Sociology. Cambridge University Press, Cambridge.

Holland, P.W., Leinhardt, S., 1971. Transitivity in structural models of small groups. Comparative Groups Studies 2, 107–124.

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

Holland, P.W., Leinhardt, S., 1981. An exponential family of probability distributions for directed graphs (with discussion). Journal of the American Statistical Association 76, 33–65.

Holme, P., Edling, C.R., Liljeros, F., 2004. Structure and time evolution of an Internet dating community. Social Networks 26, 155–174.

Huisman, M.E., Snijders, T.A.B. 2003. Statistical analysis of longitudinal network data with changing composition. Sociological Methods and Research 32, 253–287.

Huisman, M., Steglich, C. 2008. Treatment of non-response in longitudinal network data. Social Networks 30, 297–309.

Hummon, N.P., 2000. Utility and dynamic social networks. Social Networks 22, 221–249.

Jaccard, P., 1900. Contributions au problème de l'immigration post-glaciaire de la flore alpine, Bulletin de la Société Vaudoise des Sciences Naturelles 37, 547–579.

Jackson, M.O., Rogers, B.W., 2007. Meeting strangers and friends of friends: How random are social networks? American Economic Review 97, 890–915.

Jin, E.M., Girvan, M., Newman, M.E.J., 2001. Structure of growing social networks. Physical review E, 64, 046132.

Katz, L., Proctor, C.H., 1959. The configuration of interpersonal relations in a group as a time-dependent stochastic process. Psychometrika 24, 317–327.

Knecht, A., 2008. Friendship Selection and Friends' Influence. Dynamics of Networks and Actor Attributes in Early Adolescence. PhD dissertation, University of Utrecht.

Koskinen, J., 2004. Essays on Bayesian Inference for Social Networks. PhD Dissertation, Department of Statistics, Stockholm University.

Koskinen, J.H., Snijders, T.A.B., 2007. Bayesian inference for dynamic network data. Journal of Statistical Planning and Inference 13, 3930–3938.

Kossinets, G., Watts, D.J., 2006. Empirical analysis of an evolving social network. Science 311, 88–90.

Leenders, R.T.A.J., 1995. Models for network dynamics: A Markovian framework. Journal of Mathematical Sociology 20, 1–21.

Lindenberg, S., 1992. The method of decreasing abstraction. In: Coleman, J.S., Fararo T.J. (Eds.), Rational choice theory: Advocacy and Critique, Sage, Newbury Park, pp. 3–20.

Long, J.S., 1997. Regression Models for Categorical and Limited Dependent Variables. Sage Publications, Thousand Oaks, CA.

Macy, M.W., Willer, R., 2002. From factors to actors: Computational Sociology and agent-based modelling. Annual Review of Sociology 28, 143–166.

Maddala, G.S., 1983. Limited-dependent and Qualitative Variables in Econometrics (third ed.), Cambridge University Press, Cambridge.

Marsili, M., Vega-Redondo, F., Slanina, F., 2004. The rise and fall of a networked society: A formal model. Proceedings of the National Academy of Sciences USA 101, 1439–1442.

McFadden, D., 1973. Conditional logit analysis of qualitative choice behavior. In: P. Zarembka (Ed.), Frontiers in econometrics, Academic Press, New York, pp. 105–142.

McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a Feather: Homophily in Social Networks. Annual Review of Sociology 27, 415–44.

Merton, R.K., 1968. The Matthew effect in science. Science 159, 56–63.

Morris, M., Kretzschmar, M., 1995. Concurrent partnerships and transmission dynamics in networks. Social Networks 17, 299–318.

Newcomb, T.M., 1962. Student Peer-Group Influence. In Sanford, N. (Ed.), The American College : A Psychological and Social Interpretation of the Higher Learning. Wiley, New York.

Newman, M.E.J., 2002. Assortative mixing in networks. Physical Review Letters 89, 208701.

Pearson, M.A., Michell, L., 2000. Smoke rings: Social network analysis of friendship groups, smoking and drug-taking. Drugs: Education, Prevention and Policy 7, 21–37.

Price, D. de S., 1976. A general theory of bibliometric and other advantage processes. Journal of the American Society for Information Science 27, 292–306.

Robins, G., Snijders, T.A.B., Wang, P., Handcock, M., Pattison, P., 2007. Recent developments in exponential random graph ($p*$) models for social networks. *Social Networks* 29, 192–215.

Schweinberger, M., 2008. Statistical modeling of network dynamics given panel data: goodness-of-fit tests. Submitted for publication.

Skyrms, B., Pemantle, R., 2000. A dynamic model of social network formation. PNAS 97, 9340–9346.

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

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

Snijders, T.A.B., 2002. Markov chain Monte Carlo estimation of exponential random graph models. Journal of Social Structure, 3.2.
http://www.cmu.edu/joss/content/articles/volume3/Snijders.pdf

Snijders, T.A.B., 2005. Models for longitudinal network data. In: Carrington, P.J., Scott, J., Wasserman, S. (Eds.), Models and Methods in Social Network Analysis, Cambridge University Press, New York, pp. 215–247.

Snijders, T.A.B., 2006. Statistical Methods for Network Dynamics. In: Luchini, S.R. et al. (Eds.), Proceedings of the XLIII Scientific Meeting, Italian Statistical Society, CLEUP, Padova (It.), pp. 281–296.

Snijders, T.A.B., Koskinen, J.H., Schweinberger, M., 2009. Maximum Likelihood Estimation for Social Network Dynamics. Submitted.

Snijders, T.A.B., Steglich, C.E.G., Schweinberger, M., 2007. Modeling the co-evolution of networks and behavior. In: van Montfort, K., Oud, H., Satorra, A., Longitudinal models in the behavioral and related sciences, Lawrence Erlbaum, Mahwah, NJ, pp. 41–71.

Snijders, T.A.B., Steglich, C.E.G., Schweinberger, M., Huisman, M., 2008. Manual for SIENA version 3.2. Groningen: ICS, University of Groningen; Oxford: Department of Statistics, University of Oxford,
http://stat.gamma.rug.nl/snijders/siena.html.

Steglich, C.E.G., Snijders, T.A.B., Pearson, M., 2009. Dynamic networks and behavior: Separating selection from influence, Submitted for publication.

Udehn, L., 2002. The changing face of methodological individualism. Annual Review of Sociology 8, 479–507.

van de Bunt, G.G., Groenewegen, P., 2007. An actor-oriented dynamic network approach: the case of interorganizational network evolution. Organizational Research Methods 10, 463–482.

van de Bunt, G.G., van Duijn, M.A.J., and Snijders, T.A.B., 1999. Friendship networks through time: An actor-oriented statistical network model. Computational and Mathematical Organization Theory 5, 167–192.

Wasserman, S., 1979. A stochastic model for directed graphs with transition rates determined by reciprocity. In: Schuessler, K.F. (Ed.), Sociological Methodology 1980, Jossey-Bass, San Francisco, pp. 392–412.

Wasserman, S., K. Faust. 1994. Social Network Analysis: Methods and Applications. Cambridge University Press, New York and Cambridge.

Wasserman, S., Iacobucci, D., 1988. Sequential social network data. Psychometrika 53, 261–282.

Wasserman, S., Pattison, P.E., 1996. Logit models and logistic regression for social networks: I. An introduction to Markov graphs and p^*. Psychometrika 61, 401–425.

Table 1: Parameter estimates of friendship evolution models, with standard errors and two-sided p-values.

	Model 0			**Model 1**			**Model 2**			**Model 3**		
	estim.	s.e.	p	estim.	s.e.	p	estim.	s.e.	p	estim.	s.e.	p
objective function												
outdegree	-1.41	0.43	< 0.001	-1.67	0.38	< 0.001	-1.62	0.34	< 0.001	-1.59	0.33	< 0.001
reciprocity (evaluation)	1.34	0.22	< 0.001	1.42	0.20	< 0.001	1.45	0.20	< 0.001	0.71	0.48	0.14
reciprocity (endowment)										1.42	0.80	0.076
transitive triplets	0.23	0.03	< 0.001	0.21	0.03	< 0.001	0.21	0.03	< 0.001	0.20	0.03	< 0.001
transitive ties	0.74	0.21	< 0.001	0.74	0.23	0.001	0.78	0.22	< 0.001	0.67	0.20	< 0.001
3-cycles	-0.32	0.10	< 0.001	-0.26	0.09	0.005	-0.25	0.09	0.007	-0.22	0.10	0.028
indegree popularity (sqrt)	-0.18	0.12	0.13			0.23			0.23			0.29
outdegree popularity (sqrt)	-0.40	0.24	0.093	-0.56	0.23	0.013	-0.62	0.20	0.002	-0.61	0.20	< 0.001
outdegree activity (sqrt)	0.01	0.07	0.93			0.18			0.12			0.19
sex (M) ego	0.35	0.13	0.008	0.39	0.13	0.002	0.44	0.14	0.002	0.41	0.13	< 0.001
sex (M) alter	0.10	0.13	0.46	0.15	0.13	0.24	0.18	0.14	0.20	0.16	0.14	0.25
same sex	0.49	0.13	< 0.001	0.54	0.12	< 0.001	0.56	0.14	< 0.001	0.56	0.13	< 0.001
primary school	0.37	0.13	0.006	0.35	0.14	0.013	0.34	0.14	0.016	0.35	0.15	0.020
rate function												
rate period 1	10.01	2.05		9.76	1.91		9.69	1.96		10.93	2.26	
rate period 2	10.67	1.85		10.44	1.83		10.28	1.82		11.13	1.93	
rate period 3	9.51	1.53		8.91	1.39		8.81	1.37		9.38	1.52	
effect sex (M) on rate							-0.42	0.28	0.13			

The p-values are based on approximate normal distributions of the t-ratios (estimate divided by standard error). When rendered for non-estimated parameters, they refer to score-type tests.

Table 2: Estimates of model for co-evolution of friendship and delinquency, with standard errors and two-sided p-values.

effect	estimate	s.e.	p
network objective function			
outdegree	-1.91	0.41	< 0.001
reciprocity (evaluation)	0.25	0.44	0.57
reciprocity (endowment)	2.10	0.81	0.010
transitive triplets	0.22	0.03	< 0.001
transitive ties	0.67	0.23	0.004
3-cycles	-0.27	0.11	0.014
outdegree based popularity (sqrt)	-0.52	0.25	0.034
sex (M) ego	0.48	0.16	0.002
sex (M) alter	0.19	0.16	0.23
same sex	0.61	0.15	< 0.001
primary school	0.46	0.18	0.010
delinquency similarity	3.22	1.66	0.053
network rate function			
network rate period 1	9.94	2.12	
network rate period 2	10.86	2.00	
network rate period 3	9.39	1.49	
delinquency linear	0.00	0.27	1.00
delinquency quadratic	0.12	0.16	0.48
sex (M)	-0.19	0.42	0.66
average similarity	6.08	3.06	0.047
behavior rate function			
delinquency rate period 1	1.50	0.70	
delinquency rate period 2	3.50	2.48	
delinquency rate period 3	2.64	1.39	

The p-values are based on approximate normal distributions of the t-ratios (estimate divided by standard error).