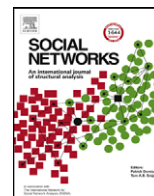




Contents lists available at [ScienceDirect](#)

Social Networks

journal homepage: www.elsevier.com/locate/socnet



Modelling the evolution of a bipartite network—Peer referral in interlocking directorates[☆]

Johan Koskinen^{a,b,*}, Christofer Edling^c

^a ReMiSS, Department of Politics and International Relations, University of Oxford, Oxford, United Kingdom

^b Nuffield College, Oxford, United Kingdom

^c School of Humanities and Social Sciences, Jacobs University, Germany

ARTICLE INFO

Keywords:

Stochastic actor-oriented models
Bipartite graphs
Interlocking directorates
Bayesian analysis
Longitudinal network data
Prediction

ABSTRACT

A central part of relational ties between social actors is constituted by shared affiliations and events. The action of joint participation reinforces personal ties between social actors as well as mutually shared values and norms that in turn perpetuate the patterns of social action that define groups. Therefore the study of bipartite networks is central to social science. Furthermore, the dynamics of these processes suggests that bipartite networks should not be considered static structures but rather be studied over time. In order to model the evolution of bipartite networks empirically we introduce a class of models and a Bayesian inference scheme that extends previous stochastic actor-oriented models for unimodal graphs. Contemporary research on interlocking directorates provides an area of research in which it seems reasonable to apply the model. Specifically, we address the question of how tie formation, i.e. director recruitment, contributes to the structural properties of the interlocking directorate network. For boards of directors on the Stockholm stock exchange we propose that a prolific mechanism in tie formation is that of peer referral. The results indicate that such a mechanism is present, generating multiple interlocks between boards.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

A bipartite network defines a graph on a node set that represents two distinct types of social entities and where there may be relational ties only between nodes that are of different types. The social entities may be individuals and social clubs, authors and journal articles, directors and corporate boards, etc. Since bipartite networks make the conceptual link between social actors and social groups explicit, the study of bipartite networks is of great theoretical importance for understanding the duality of persons and groups (Breiger, 1974; Doreian, 1979; Feld, 1981; Freeman and White,

1993; Pattison and Robins, 2002). The nodes and ties in bipartite networks may be researchers citing papers (e.g., Small, 1973); people participating in social events (Breiger, 1974); partners in crime (Frank and Carrington, 2007); or directors on corporate boards (Mizruchi, 1996). This paper is focused on the last case, directors and boards. A so-called interlocking directorate is a bipartite board-to-director network consisting of vertices representing corporate boards and directors. A tie exists between board i and director j if j is a member of board i .

Most previous research on bipartite networks has focused on one of the two unimodal networks where two nodes of one type are considered tied if they share at least one alter of the other type. The obvious benefit is that methods developed for the analysis of unimodal networks applies. But as pointed out by Robins and Alexander (2004), the duality in defining for example the individuals by their affiliations with certain events or in defining the events by their ties to individuals, makes it hard to give priority to one type of node over another (Breiger, 1974; Galaskiewicz et al., 1985; Breiger and Pattison, 1986). Transforming a bipartite network to a unimodal network always means a loss of information (Borgatti and Everett, 1997). In particular, the strength of a tie is lost if we ignore the number of nodes of one type that two nodes of another type share. It can also be shown that a completely random bipartite network may give rise to quite interesting but altogether spurious structural features in the unimodal representations (see also Newman et al., 2001).

[☆] An earlier version of this paper was presented at the workshop on “The Interdependence of Networks and Behaviour Psychological, Sociological, Political and Economic Perspectives” in Melbourne in 2008, at which we received many useful suggestions. This work has benefitted from helpful comments and suggestions by Malcolm Alexander, Peng Wang, and Garry Robins. We gratefully acknowledge the financial support from the Swedish Council for Working Life and Social Research and from Riksbankens Jubileumsfond. The research of J. Koskinen was partly done while being a Research fellow at the School of Behavioural Science, University of Melbourne. Support is also acknowledged for part of the work of Koskinen from US National Institutes of Health (NIH 1R01HD052887-01A2).

* Corresponding author at: ReMiSS, Department of Politics and International Relations, University of Oxford, Manor Road, Oxford, OX1 3UQ, United Kingdom.
Tel.: +44 1865 278524; fax: +44 1865 278725.

E-mail addresses: johan.koskinen@nuffield.ox.ac.uk (J. Koskinen), c.edling@jacobs-university.de (C. Edling).

Because of the inherent dynamic nature of social networks, a wide variety of statistical models have been proposed for studying network evolution over time. While several methods assume that changes are made in discrete steps from one moment to the next (Katz and Proctor, 1959; Wasserman and Iacobucci, 1988; Sanil et al., 1995; Robins and Pattison, 2001), considerable advantages may be had from modelling longitudinal social networks in continuous-time (cf Snijders, 1996, 2001). The early models for longitudinal social networks using continuous-time Markov chains (Holland and Leinhardt, 1977a,b; Wasserman, 1980a,b; Leenders, 1995) built on the assumption that the dyads evolved independently of each other, something which excludes the exploration of dependencies on larger structures than the dyad. In response to this Snijders proposed a flexible and empirically testable class of stochastic actor-oriented (or actor-based) models that draw on the assumption that actors strive towards organising their social ties in a utility maximizing manner (Snijders and van Duijn, 1997; Snijders, 2001, 2005, 2006).

However, so far the evolutionary network models have not been adapted to the study of the evolution of bipartite networks, although there are some longitudinal studies of the evolution of interlocking directorates. Galaskiewicz and Wasserman (1981) fit a discrete-time model to the unimodal corporate network implied by interlocks in a small bipartite network. Like previously proposed models it also assumed dyad-independence. In their study they focused primarily on the corporations and they distinguished between different industries and used a qualified version of board membership that gives directed interlocks. Mizruchi and Stearns (1988) model the appointment of financial directors on non-financial boards for a small selection of boards. Studying the dichotomous event “hiring” or analysing the evolution of the induced unimodal network, prevents investigation of the strength of interlocks. Similarly, in these two models changes are largely explained through corporate specific covariates (size, solvency, profitability, etc.) rather than in terms of structural aspects of the bipartite (and unimodal) network(s), i.e. endogenous self-organising principles.

Here we propose a model for studying the evolution of bipartite networks that draws on Snijders’ work on stochastic actor-oriented networks. We analyse data on corporate boards for all firms traded on the Stockholm Stock Exchange’s primary list between 1996 and 2005 (Edling and Sandell, 2001; Bohman, 2006). In our application, we focus on the processes of appointment to and departure of directors from boards based on a theory of action in the context of interlocking directorates. Central to this concept is the occurrence of multiple interlocks, which we elaborate on in Section 3.

We proceed by describing the statistical model. This is first done using a general model formulation with an inference strategy for the evolution of bipartite graphs that follows the model proposed in Koskinen and Edling (2004) that is in turn an extension of the stochastic actor-oriented models for directed unimodal graphs of Snijders (2001). We then discuss peer referral in interlocking directorates, in Section 3, before detailing the modifications needed for analysing interlocking directors in Section 4. We then fit models with structural features and with additional controls for director attributes in Section 5. To test the goodness of fit for the model, in Section 6 we compare forecasts from the model with observed data. We finish the paper with some concluding remarks.

2. Model formulation and estimation

We consider a bipartite graph on a fixed vertex set V that is the union of two disjoint subsets A and B , $m = |A|$, $n = |B|$. We denote by N the subset of pairs of vertices that constitute the set of possible relational ties between vertices. In the case of a bipartite graph $N =$

$A \times B$, there can only be ties between vertices from different classes. The aim is to model the evolution through time of the random set of edges $E \subseteq N$. The (generalised) adjacency matrix is a collection $x = (x_e : e \in N)$ of $|N| = nm$ tie-variables $1\{e \in E\}$, indicating whether the tie is present or not, and we denote by \mathbf{X} the space of all possible realisations of x . For $x, y \in \mathbf{X}$ we define the Hamming metric as

$$\|x - y\| = \sum_e |x_e - y_e|.$$

The model is a continuous-time Markov chain $\{x(t)\}_{t \in \mathbb{R}}$ on \mathbf{X} , which may be defined in terms of its embedded chain (or walk on the $|N|$ -cube) as in Koskinen and Snijders (2007). The transition probabilities in the embedded chain are given by a function

$$\eta(\theta, x, y) \tag{1}$$

and the time spent in $x \in \mathbf{X}$ exponentially distributed with rate function

$$\lambda(\theta, x) \tag{2}$$

where θ is a vector of parameters. It is assumed that $\eta(\theta, x, y) \geq 0$ for $\|x - y\| = 1$ and 0 otherwise. If we let $\Delta_e x$ denote the array that has element e equal to $1 - x_e$ but that is equal to x for all other elements, $e^* \in N \setminus \{e\}$, the process may be described as remaining in a state x for a period of time that is exponentially distributed, after which element e is changed to $1 - x_e$ with probability $\eta(\theta, x, \Delta_e x)$. For a digraph with $N = V^{(2)}$ this is exactly the way the stochastic actor-oriented model for unimodal graphs is defined in Snijders (2001) for the purpose of simulation.

This is a framework that covers a wide variety of different models for network evolution. The expression (1) can often be described with a logistic function, derived from evaluation of utility functions, and Snijders (2006) showed how many models of network evolution may be couched in the form of stochastic actor-oriented models. To express the bipartite evolution model in terms of an actor-oriented model we follow Koskinen and Edling (2004) and make the assumption that only vertices in A are considered to be actors so that given that a change is made, it is one of the vertices in A , say i , that decides to make a change, and if i changes its relation to j there is a jump from x to $\Delta_{ij}x$ (we use ij as a notational shorthand for (i,j)).

In the standard stochastic actor-oriented model (Snijders, 2001), all actors i independently make changes at a rate $\lambda_i(\theta, x)$, given the current state x . This gives the rate (2) as $\lambda(\theta, x) = \sum_i \lambda_i(\theta, x)$. Using well-known properties of the exponential distribution it can be shown that, given that a change is made, the probability that it is actor i that makes this change is $\lambda_i(\theta, x)/\lambda(\theta, x)$. Consequently, we may express (1) as a product

$$\eta(\theta, x, \Delta_{ij}x) = \frac{\lambda_i(\theta, x)}{\lambda(\theta, x)} p_i(\theta, x, j),$$

where

$$p_i(\theta, x, j) \tag{3}$$

is the conditional probability that the process is going to jump from x to $\Delta_{ij}x$, given that a change is made to $(x_{ij})_{j=1}^m$ for i fixed.

Given that actor i makes a change, i changes the relation to $j \in B$, such that the “utility” $u_i(\theta, x, j, t)$ is maximised. The “utility” $u_i(\theta, x, j, t)$ captures the desirability to i of changing x_{ij} , of going from x to $\Delta_{ij}x$. The stochastic actor-oriented framework for network evolution allows for many aspects of the change to be taken into consideration (Snijders et al., 2007) but we settle for a simple form

$$u_i(\theta, x, j, t) = f_i(\theta, \Delta_{ij}x, x) + \varepsilon_i(x, j, t),$$

where the first term on the right hand side is a systematic component and the second, $\varepsilon_i(x, j, t)$, is a random component. The

systematic part of the utility function represents what we model and is meant to capture the regularities in the data. The random components represent the non-systematic chance aspects of the decision situation that we do not explicitly model. Typically these are assumed to be independently identically distributed according to a type 1 extreme value distribution,¹ which is a convenient choice for which a form of (3) is computationally equivalent to a conditional logistic regression.

2.1. Estimation

Given T realisations $x(t_0), x(t_1), \dots, x(t_{T-1})$ from the model defined by (1) and (2), the purpose is to perform inference for the model parameters θ . In describing this process we may without loss of generality assume that we only have two observations, $x(t_0)$ and $x(t_1)$, since the Markov property means that we may treat each pair of observations separately.

In order to write the likelihood function for θ given $x(t_0)$ and $x(t_1)$, let $y = y(s_0), y(s_1), y(s_2), \dots, y(s_{H+2k})$ be a realisation of the evolution from $x(t_0)$ to $x(t_1)$. Note that if $y(s_{H+2k}) = x(t_1)$ and $y(s_0) = x(t_0)$, for $H = \|x(t_0) - x(t_1)\|$, the number of changes must be $H + 2k$, for $k \in \{0, 1, 2, \dots\}$, because when the H necessary changes have been made to transform $x(t_0)$ into $x(t_1)$, any extra k unnecessary changes have to be reversed. Had a path y been observed then the likelihood function would have been given by

$$L(\theta; y) = \exp \left\{ - \sum_{h=1}^{H+2k+1} (s_h - s_{h-1}) \lambda(\theta, y(s_{h-1})) \right\} \times \prod_{h=1}^{H+2k} \eta(\theta, y(s_{h-1}), y(s_h)) \lambda(\theta, y(s_{h-1})) \quad (4)$$

where $s_{H+2k+1} = t_1$. The likelihood given observed data, the observed data likelihood, is obtained when (4) is marginalised with respect to y , conditional on the endpoints $x(t_0)$ and $x(t_1)$. From a Bayesian perspective, all relevant information about the model parameters is given by the posterior distribution, $\pi(\theta | x(t_0), x(t_1))$, and with constant prior distributions it is proportional to the observed data likelihood. In this case, neither the posterior distribution nor the observed data likelihood, are analytically tractable. We may however simulate from the joint posterior $\pi(\theta, y | x(t_0), x(t_1))$ which is proportional to (4) (multiplied by a proper prior distribution $\pi(\theta)$ if available). A sample from this distribution may be obtained using Markov chain Monte Carlo (MCMC) by creating a Markov chain that alternates between drawing from $\pi(\theta | y)$ and $\pi(y | \theta, x(t_0), x(t_1))$, both of which are proportional to (4) (for details see Koskinen and Snijders, 2007).

We return to model estimation in Section 4, when applying this model to analyse peer referral in interlocking directorates.

3. Interlocking directorates and peer referral—the action mechanism

This section presents an action mechanism for the actor-oriented model of network evolution. We specify and interpret the model based on two interacting mechanisms: peer referral and homosociality (or cognitive gender homophily). Before we discuss these mechanisms, we wish to briefly underscore the importance of explicitly considering the bipartite structure in the analysis of interlocking directorates.

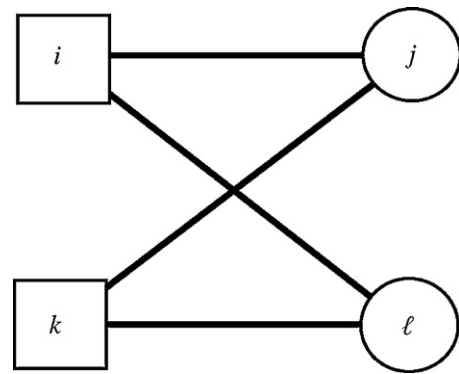


Fig. 1. A four-cycle (C_4): multiple overlaps create a multiple tie between board i and board k .

One can easily transform a bipartite board-to-director network into two one-mode symmetric networks: The inter-personal, director-to-director network consists of the directors and the ties between them. Two directors share a tie if they are members of the same board. Similarly, in the board-to-board network, the boards are the vertices and there are ties between pairs of boards that share a board member. However, some important information is lost by such transformations.

The multiple interlock as illustrated in Fig. 1 (C_4) occurs when two boards of directors share (at least) two directors (also Four-cycle, Pattison and Robins, 2002; Wang et al., 2009; or (2,2) biclique, Borgatti and Everett, 1997). In the following graphs circles and squares correspond to directors and boards respectively and a line connecting a director with a board means that the director is a member of the board. If any one of the ties, say $\{i, j\}$, is missing from the subgraph in Fig. 1, the board i is only indirectly linked to j through a path of length 3. We refer to this structure as a (bipartite) three-path (L_3). Should director j be appointed to the board i , such that the tie $\{i, j\}$ is created, the three-path is transformed into the four-cycle depicted in the figure. The addition of this edge does not alter any of the one-mode networks as the two directors were already connected through the other board and the two boards were connected through the director that sat on both boards. Thus in terms of the two one-mode networks these two subgraphs L_3 and C_4 are equivalent.

Most analyses of interlocking directorates have focused on the one-mode (binary) representations, and multiple interlocks (i.e. instances of C_4) have in general not been distinguished from single interlocks. The reason for not investigating multiple interlocks can stem from the choice of model (e.g., Mizruchi and Stearns, 1988), or lack of data (e.g., Barnes and Ritter, 2001). However, a small number of studies focus specifically on the prevalence of multiple interlocks, in particular the broken tie literature (Ornstein, 1982, 1984; Palmer, 1983). Whereas these early studies found that a severed interlock was more likely to be replaced if this interlock was part of a multiple interlock, Palmer et al. (1986) suggested that this association was spurious and that the true determinant was the level of embeddedness. They referred to the production of an interlock that corresponds to the transition from a three-path to a four-cycle as “reinforcement”. The process of reinforcing already present ties between boards as an expression of the firms’ embeddedness can also be taken down to the level of directors. Whereas we focus on the social process of board-to-board tie reinforcement, the phenomena have also gained some attention in corporate governance literatures, and the conclusion seems to be that mutual interlocks are not strategically motivated (Carpenter and Westphal, 2001; Fich and White, 2003, 2005).

Robins and Alexander (2004) mention personal referral as a possible process that lead to multiple interlocks. When a board is

¹ The assumption of independence across $j \in B$ for each i may be relaxed by positing a multivariate normal distribution instead (Koskinen, 2004).

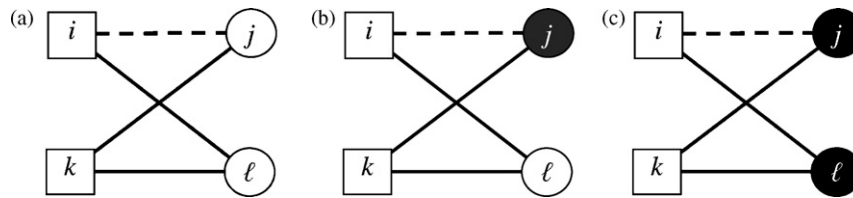


Fig. 2. Bipartite attribute dependent configurations and the removal/creation of the tie $\{i,j\}$. Structures including dashed tie where dark (director) node indicates node having attribute and light node indicates absence of attribute. For attribute being female (W) labels are: (a) male homogenous four-cycle (C_{MM}); (b) heterogeneous four-cycle (C_{MW}); (c) female homogenous four-cycle (C_{WW}).

looking to appoint a new board member, a present member may recommend a person with whom he or she has had previous experience of working with on another board. If this personal referral is endorsed by the board a multiple interlock is created.

This practice can be expressed in terms of an action mechanism in which ego puts a preference on alter when it comes to recommending new directors to the board of which ego is a member. Explaining the formation of this preference goes beyond the scope of this paper, but we can suggest some plausible explanations worthy of future exploration. An obvious candidate is that directors have more first-hand information on directors whom they have worked with, and that they consider first-hand information more trustworthy than second-hand information. Another suggestion is that working together on a board gives rise to a sense of community among board members that would make fellow board members first choice candidates for filling vacant positions. A related idea, assuming that multiple board assignments are attractive, is that directors who know each other from working together recommend each other reciprocally. A more sophisticated explanation could be that directors are bound by silent agreements and codes of conduct stemming from a sense of membership in collectives such as “old boy’s networks” and “inner circles” (Mills, 1956; Useem, 1984) that tell them to look out for their peers.

Applying the mechanism of peer referral, the interlock itself (manifested by the two-star) is not desirable for the firm. The interlock may be desirable to the director who acts merely as an unconstrained individual or to the board that acts as a collective representing a firm that strives to be central (for purposes of information gathering or for other purposes) in the board-to-board network. Boards and directors are also unlikely to strive towards establishing reach in the bipartite network as manifested in a desire to create three-paths. If peer referral of co-members is expressing the individual action constrained by the setting and informed by group norms, then it is a mechanism that is central to the evolution of the network. The empirical manifestation of this action mechanism is the creation of four-cycles and, on the network level, a high bipartite clustering. The clustering one might observe in the two-unimodal networks, as manifested in triangulation as well as short path-ways, are emergent features of the processes in operation in the bipartite evolution. Perpetuating these behaviours reinforces the norms behind it—these are part of the dynamics, the mechanism is both the cause and the effect.

Naturally, personal referral does not necessarily result in four-cycles in the board-to-director network. A board member may well recommend a person on other grounds than that they have worked together on another board, and the recommended person might not be a member of the (relevant) population. Underlying social structures are important in the recruiting process.

If peer referral is indeed a central mechanism to the evolution of the network this may reflect underlying allegiances in which, we may hypothesise, women are not included. If women are not part of the old boy’s networks and not generally treated as peers by men we would expect that the recruitment of women does not follow the same rules as for men (e.g., Israeli and Talmud, 1997). This

would then manifest itself in a low frequency of peer referrals that involved women. In order to investigate this we may distinguish different types of four-cycles.

In Fig. 2 there are three types of four-cycles that may result from ℓ referring j . In Fig. 2a a man is referring another man creating what may be termed a male-homogeneous four-cycle. In Fig. 2c a woman is referring another woman creating what may be termed a female-homogeneous four-cycle. An example of a heterogeneous four-cycle is given in Fig. 2b, where a man refers a woman. When women are not included in the peer practices outlined above we expect instances of heterogeneous four-cycles to be rare. In our data there is not enough information to fully separate the different four-cycles wherefore we settle for including the interaction effect of women and four-cycles, essentially counting the number of structures Fig. 2(b) and (c). The mirror image of Fig. 2(b), where a male is referred by a woman we deem to be of little interest and too rare to be empirically testable.

If there is a negative tendency towards creating heterogeneous four-cycles, this is exactly an example of how “homosocial reproduction” operates—men prefer to socialise with men and promote other men (Kanter, 1977; Stafstudd, 2006). If homosocial reproduction drives the evolution indicating that women are excluded from the peer practices of the boards this may well indicate further imbalances.

To parse out whether peer referral is indeed an active mechanism or whether the prevalence of four-cycles is simply a result of accumulative interlocks requires a dynamic model. We furthermore need to distinguish between preservation of multiple ties and the creation of multiple ties, the latter which indicates peer referral. If the mechanisms investigated here are active and if their application is reinforced through repeated use, then this is a process through time. If cross-sectional snap-shots were investigated it would not be possible to say whether peer referral is a key ingredient in board appointments or if higher than expected number of four-cycles at each cross-section is simply inherited from previous time-points. The process is also by definition dynamic since the actions of the individuals within the board are not independent of their environment, the context. The action by one board furthermore reshapes the context for other boards.

4. Modelling and estimating peer referral in interlocking directorates

Before we define the specific forms of (1) and (2), given in Section 2, the longitudinal study of interlocking directorates requires a few further modifications of the model specification. These are modifications that stem both from the nature of interlocking directorates as well as from computational considerations that have to be taken into account when dealing with large datasets.²

² Other modifications are needed for other types of bipartite data. For example, in the case of co-offending networks, offenders choose offences to participate in but the offences themselves are transitional and not independently defined as in

4.1. Agency

As argued in the previous section we observe actions taken by the boards but assume that this is the result of actions of individual directors making up the boards. For the purpose of fitting the model the board itself is considered the actor. Given a choice between allocating choice to board or directors the former seems more plausible.³

Since it is assumed that the actors (now, boards) chose the course of action that maximises the utility to board, and that the maximisation is carried out over the entire set of actions, we also assume that the board is acting on perfect information. The random components allow for deviations from the systematic part of the utility function but the random component is assumed to be independent of model quantities and therefore it cannot be used to account for differential, fuzzy, knowledge of different “parts” of the decision space. The Markovian nature of the model assume that there is no strategic action in the sense that a board is prepared to take a loss now in order to gain in the future.

4.2. Differentiating between recruitment and release

Although we only aim at providing a rough approximation of the mechanism central to the evolution of the interlocking directorates, the difference in nature between recruiting a director and releasing a director has to be addressed (in the broken tie studies and in e.g., Mizruchi and Stearns (1988), the initiation of a tie is clearly distinguished from the dissolution of a tie). We may for example expect that in many of the instances where a director leaves a board this is not a consequence of a choice on the part of the board but is more likely to be the result of retirement, illness, etc. To distinguish between the different processes, we let

$$f_i(\theta, \Delta_{ij}x, x) = \begin{cases} f_{i,1}(\theta, \Delta_{ij}x), & \text{if } x_{ij} = 1 \\ f_{i,0}(\theta, \Delta_{ij}x), & \text{if } x_{ij} = 0. \end{cases} \quad (5)$$

This is akin to the gratification function in Snijders (2001), in that the satisfaction with a new state may depend on the previous state.

The number of seats on a board of directors is not likely to fluctuate much from year to year. To control for the effect of board size in the model outlined above would require a complicated rate function. This would also have to be combined with sensibly chosen functions of the board size in $f_i(\theta, \Delta_{ij}x, x)$ but the differences in board sizes between boards, and the small number of choices for releasing a board member in comparison to the number of available choices leads to numerical instabilities.⁴ For the present dataset for instance, a board in 1996 with six board members has six choices of directors to release and a choice of 1238 directors to recruit. To resolve this problem we introduce a binary random variable $v_i(\phi, x)$ that indicates whether the board i decide to recruit or release a director. This is included in the model as follows: given that i decide

to make a change to the board composition $(x_{ij})_{j=1}^m$, and $v_i(\phi, x) = v$, (3) is set to

$$p_{i,v}(\theta, x, j) = \begin{cases} \frac{\exp(f_{i,v}(\theta, \Delta_{ij}x))}{\sum_k \exp(f_{i,v}(\theta, \Delta_{ik}x))1\{x_{ik} = v\}}, & \text{if } x_{ij} = v \\ 0, & \text{otherwise} \end{cases}$$

The jump probabilities in (3) may now be written as mixtures

$$p_i(\theta, x, j) = Pr\{v_i(\phi, x) = 1\}p_{i,1}(\theta, x, j) + (1 - Pr\{v_i(\phi, x) = 1\})p_{i,0}(\theta, x, j).$$

We settle for a simple expression for the probabilities

$$Pr\{v_i(\phi, x) = 1\} = \begin{cases} \phi, & \text{if } \sum_j x_{ij} > 0 \\ 0, & \text{otherwise} \end{cases}$$

We are not going to specify the probabilities $Pr\{v_i(\phi, x) = 1\}$ any further since the corresponding variables are ancillary for the purpose of parameter inference in the estimation scheme (as we shall see further on).

4.3. Latent walk length

In the basic model formulation, given that $x(t_0)$ changes into $x(t_1)$, we know that at least H changes have taken place. In addition there might have been $k \in \{0, 1, 2, \dots\}$ superfluous changes that were then reversed. For friendship networks it is plausible to assume that a number of such superfluous changes have been made but perhaps less so for interlocking directorates.⁵ We could introduce strong prior information to reflect the fact that we believe k to be small but the size of the dataset makes it convenient to perform inference conditioned on $k = 0$. The consequences for inference are that we only need to employ “Move type 3” (in the terminology of Koskinen and Snijders, 2007, i.e. permuting the order of the changes) in the updating of latent walks and, that the probabilities $Pr\{v_i(\phi, x) = 1\}$ cancel in the evaluation of (4) for updating the latent walks.

As for all the models in the stochastic actor-oriented framework for network evolution, had the exact timing of the changes been observed sequential logistic regression could have been used (Butts, 2008). In the case investigated here the exact times are however not available. In principle there are fixed points in time when the new boards are officially appointed⁶ (annual grand meetings) but (notwithstanding the fact that they are not recorded here) from a substantive point of view the decisions are not instantaneous. The process of appointment may be said to be one going on over a longer period of time, during which the search for directors goes on. The search for directors to nominate can take many expressions as it is not legally regulated (Bohman, 2009). Sometimes, but far from always board nomination committees are used (ibid.). At the time of official endorsement by the shareholders the boards have already been de facto appointed (there are exceptions, instances when the annual grand meeting reject the recommendations by the board, but these are rare).

the case of boards of directors. Similarly, in co-authorship networks, the event, an article, is not defined independently of its authors.

³ A model where both boards and directors initiate change may lead to identification issues since we would not be able to determine whether the decision was made by the board or the director. Furthermore, if one is prepared to accept the notion that (to all intents and purposes) the board operates with relative autonomy, there still remains to explain the complex collective action of board room. While some of this may be addressed in the current model framework for our purposes we are satisfied with the fact that we have a model for how the individual action mechanism is reflected in the actual board decisions. An alternative would be to frame appointment as a negotiating process which is close to the utility based models for edge-based processes proposed in Snijders (2006).

⁴ Previously Koskinen and Edling (2004) tried to solve this by including a quadratic function of the number of current board members in $f_i(\theta, \Delta_{ij}x, x)$.

⁵ For the EIES friendship network k was found to be of the order 5–10 (Koskinen and Snijders, 2007).

⁶ If this was thought to be the evolutionary process of the network then a model where changes are made all at once could be applied (Mayer, 1984). If a dyad independent model is assumed and the times of appointment are observed a discrete-time model may be defined (Galaskiewicz and Wasserman, 1981). Note that they also deal with the directed unimodal board-to-board network. Both of these approaches would require a different framework for decision-making and might also require a level of detail that we do not aspire to and in the interest of introducing the bipartite evolution model we do not want to unnecessarily limit the generality of the approach.

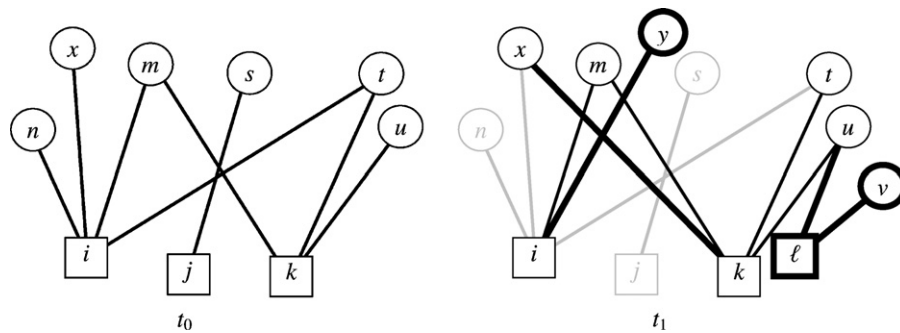


Fig. 3. Boundary issues for the evolution of bipartite graphs: between t_0 and t_1 , four ties have disappeared, four ties have been created, one board is born and one board has died. As a consequence and due to the fact that directors that are not on boards do not belong to the graph, two directors died and two were born.

4.4. Data and boundary considerations

The definition of node sets when studying evolution of bipartite networks may depend on the application but in most cases the boundary specifications differ from those for unimodal network analysis. As an example of the difficulties in the longitudinal study of interlocking directorates consider the networks in Fig. 3. There are two (hypothetical) realisations of a bipartite network at time t_0 and t_1 , and for the two cross-sectional networks it is assumed that a node of either type can only be part of the network if it is tied to a node of another type. Tie dissolution and tie creation thus redefine the set of nodes (at t_1 greyed lines denotes ties or vertices that have disappeared and thick lines denotes ties or vertices that have appeared). There are three instances where ties are dissolved. When $\{i, n\}$ is dissolved, n disappears from the bipartite graph since i is the only board n is a member of. The node t does not disappear when $\{i, t\}$ is dissolved since t is also a member of board k . The node x is only a member of board i and consequently would have disappeared had the tie $\{k, x\}$ not been created. The fact that $\{i, y\}$ is created means that y is added to the population. If a node dies like j , it takes all nodes with it, like s , that are not tied to any other node. Analogously, the birth of a board may bring with it nodes that were not already part of the population, like in the case of $\{l, v\}$ and v .

We use what may be termed a pair-wise union rule, according to which for two points in time t_0 and t_1 , the node set is considered to consist of all the nodes that are present at at least one of those occasions. By analysing the data for years pair-wise rather than for the entire period 1996–2005 we minimise the errors that may occur as a result of node set discrepancies as well as reduce the data down to a manageable size—the period 1996–2005 has 429 unique boards and 3177 unique directors but the largest network according to the union rule is 2001–2002 that has 306 boards and 1675 directors.

The population of boards necessarily changes over time as a result of mergers, acquisitions, businesses going out of business and the introduction of new corporations. The number of births and deaths each year are given in Table 1. For example we see that out of the 220 boards of corporations listed in 1996, 20 had died in 1997 and another 26 boards had been born. By the pair-wise union rule this means that the number of boards in the population in the period 1996–1997 is $220 + 26 = 246$ and the size of the disjoint union is $220 - 20 = 200$. Since 40 boards are born in 1998, the number of boards in the period 1997–1998 is 266.

The derived population of directors also changes as a result of board deaths and births as well as dissolution and creation of ties. The effect of the pair-wise union rule on the bipartite director degree distribution may be seen in Table 1. The degree distribution when the data for 1996 has been transformed according to the union rule is identical to the original data except for the number of directors without board positions, the number of directors with

degree 0. These 216 individuals are the directors that are members of at least one corporate board in 1997 but not members of any board in 1996. The data at 1997 are augmented by 211 directors with zero degree when 1997 is used as the second time-point. These are the individuals that that are members of at least one corporate board in 1996 but not members of any board in 1997.

The percentage of female directors increases from 4.8% in 1996 to 18% in 2005. As an example of how the population of directors is affected by the pair-wise union rule consider the percentage of female directors in 1996 and in 1997, which is 4.8 and 4.3 respectively. Since all female directors that are board members for at least one time-point are included in the population for 1996–1997, the percentage of female directors in the data used is in-between, namely 4.7. The effects on the average number of board positions for men and women, respectively, from applying the pair-wise union rule are also given in Table 1, as are the bipartite director degree distributions for the original data and the transformed data.

In some instances the pair-wise union rule may come with some drawbacks. Some director departures may in fact reflect board deaths rather than release of a director. Births of directors may involve something of a circular definition; if a director is not on any board in 1996 but on a board in 1997, the board may choose that director according to the model because this director is going to be chosen in the future by that board. It would be better to model births of directors separately and in principle it is straightforward to do so in the Bayesian framework but in view of parsimony we refrain from modelling this. An alternative would be to start only with directors that are board members. Huisman and Snijders (2003) deal with changing composition by letting the births and deaths of nodes be time-varying but exogenously defined events.

When a director retires, dies or in another way leaves the population altogether, this may mean the director leaves several board positions simultaneously, which violates the model assumptions and may introduce additional dependencies between the decision processes of the boards. This is likely to affect the dissolution process more than the appointment process.

4.5. Statistics

To define the systematic functions in (5), we introduce a set of parameters $\theta_1, \dots, \theta_p$, that are used to weight together counts $z_{i1}(x), \dots, z_{ip}(x)$ of configurations in a bipartite graph to determine the (systematic) contribution to the satisfaction of the board i with x . When differentiating between creating a tie and dissolving a tie, we may write

$$f_{i,v}(\theta, \Delta_{ij}x) = \theta_{1,v}z_{i1,v}(\Delta_{ij}x) + \dots + \theta_{p,v}z_{ip,v}(\Delta_{ij}x) \tag{6}$$

where $v = x_{ij}$. The vector of parameters here is assumed to be unrestricted. When a parameter is positive this means that boards tend to prefer to make changes that lead to increases

Table 1
Bipartite director degree distributions. The order of the degree distributions for a year is: when year used as t_1 , original data (in bold face), when year used as t_0 .

	Deaths/births	Number of positions										Average position			
		0	1	2	3	4	5	6	7	8	9	10	Men	W	%W
1996	(220)	-	799	120	45	34	11	15	2	2	0	0	1.46	1.11	4.8
	23	216											1.21	0.93	4.7
1997	26	211	-	796	129	47	27	14	12	7	1	0	1.22	0.91	4.7
	20	276											1.47	1.19	4.3
1998	40	173	-	864	154	56	28	21	7	5	1	0	1.16	0.85	4.8
	20	291											1.46	1.20	4.7
1999	38	193	-	965	147	58	32	21	6	4	1	0	1.16	0.88	5.1
	30	394											1.42	1.24	4.7
2000	44	262	-	1085	172	49	38	11	7	4	0	0	1.23	0.99	5.1
	19	297											1.42	1.24	4.7
2001	33	227	-	1137	183	64	28	16	3	1	3	0	1.08	0.84	5.2
	26	239											1.42	1.24	4.7
2002	17	311	-	1063	188	61	27	12	8	2	2	1	1.17	1.18	5.3
	18	240											1.36	1.27	5.5
2003	3	273	-	1042	188	51	28	11	8	3	0	0	1.12	1.07	5.3
	10	233											1.36	1.27	5.7
2004	5	231	-	1050	182	55	28	8	6	2	1	1	1.17	1.02	6.1
	14	175											1.36	1.27	5.7
2005	4	213	-	1023	173	66	17	7	5	3	1	0	1.17	1.02	6.1
													1.36	1.27	5.7

in (or preserves in the case of dissolution) the number of configurations corresponding to that parameter, given everything else.

In the case of cross-sectional data, the Markov assumptions of Frank and Strauss (1986), as applied to bipartite networks implies that the only sufficient statistics are the bipartite star statistics (Skvoretz and Faust, 1999; Wang et al., 2009). Four-cycles may also be derived from social circuit outcome dependence assumptions (Snijders et al., 2006) as shown in Wang et al. (2009). Here our primary interest is in peer referral and the creation of four-cycles, which is why we let the four-cycle be one of the configurations we include in (6). As illustrated in Fig. 4(c), a four-cycle is created when a tie $\{i,j\}$ is added by i to a three-cycle. If there is a positive coefficient corresponding to the number four-cycles for creating ties then we know that there is a preference for appointing directors that is consistent with peer referral. If in addition the coefficient corresponding to the number of four-cycles for dissolving ties is zero or negative then we know that the preference for appointing direc-

tors that is consistent with peer referral is not only an artefact of boards wanting to have strong ties to other boards. The systematic function is a linear function of the number of four-cycles (ceteris paribus) not primarily because the value of multiple interlocks is assumed to be additive but because there is a positive (negative) preference for recruiting a director if he or she has been referred by someone on the board. Occasionally, the addition of a tie may create more than one additional four-cycle and one may then ask if the effect in those cases should be additive—does it matter if a director is referred by more than one present director? Here we assume that it does but alternatives where only the sign (addition of four-cycle or not) is considered could well prove useful.

The four-cycle is built up of bipartite director two-stars (S_{B2}) (Fig. 4a) and a bipartite three-path (L_3) (Fig. 4b). To ascertain that the preference for creating four-cycles is not merely due to a preference for establishing interlocks (S_{B2}) or long paths (L_3), we need to include counts of these as well. Some other counts described in Skvoretz and Faust (1999), Wang et al. (2009), and Robins and

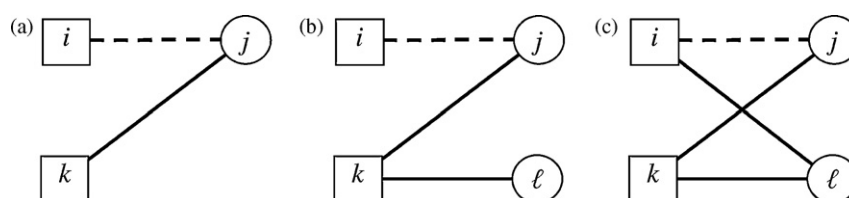


Fig. 4. Bipartite configurations and the removal/creation of the tie $\{i,j\}$. Structures including dashed tie: (a) bipartite director two-star (S_{B2}); (b) bipartite three-path (L_3); (c) bipartite four-cycle (C_4).

Table 2
Parameter estimates (posterior means and standard deviations) for longitudinal model fitted pair-wise to corporate boards on the Stockholm stock exchange. Structural effects only.

	Effect					
	Dissolution			Creation		
	Dir. Two-star	Three-Path	Four-Cycle	Dir. Two-star	Three-Path	Four-Cycle
1996–1997	-0.409 (0.159)	0.045 (0.014)	0.052 (0.073)	-0.477 (0.205)	-0.002 (0.016)	1.046 (0.076)
1997–1998	0.219 (0.266)	-0.022 (0.023)	0.227 (0.092)	-1.110 (0.203)	0.021 (0.017)	1.019 (0.077)
1998–1999	-0.118 (0.198)	0.014 (0.017)	0.028 (0.074)	-0.981 (0.202)	0.012 (0.017)	1.437 (0.084)
1999–2000	-0.175 (0.221)	0.017 (0.02)	0.162 (0.117)	-0.896 (0.256)	-0.039 (0.022)	1.295 (0.112)
2000–2001	-0.220 (0.182)	0.023 (0.017)	-0.096 (0.115)	0.430 (0.223)	-0.122 (0.019)	1.509 (0.109)
2001–2002	0.172 (0.162)	-0.002 (0.016)	0.182 (0.110)	-0.084 (0.255)	-0.073 (0.021)	1.685 (0.087)
2002–2003	-0.058 (0.181)	0.004 (0.016)	0.047 (0.089)	-2.390 (0.522)	0.085 (0.044)	1.220 (0.102)
2003–2004	-0.309 (0.352)	0.032 (0.030)	0.235 (0.151)	-2.225 (0.551)	0.063 (0.044)	1.303 (0.150)
2004–2005	-0.273 (0.231)	0.023 (0.022)	0.286 (0.136)	-3.003 (0.460)	0.122 (0.036)	1.626 (0.150)

Alexander (2004) such as the number of edges and bipartite board two-stars are not relevant if we condition on walk length. For example, conditional on the event that the board is going to add a director, all potential new board members will yield the same contribution to the count of bipartite board two-stars. The corresponding parameter is hence not estimable.

In a first model only the three structural features illustrated in Fig. 4 are included. In a second model we also include several director attributes as covariates. In the cases where these are time-varying, they are defined in relation to time-points t_0 and t_1 for the relevant pair of years.

4.6. Covariates

Many studies have pointed to the importance of the relation between the corporate board and the CEO (Zajac and Westphal,

1994; Fich and White, 2003) for the performance of the board and the company. The position of CEO offers an opportunity for forging alliances and gaining experience and insight. Fich and White (2005) also put the focus on the overrepresentation of CEOs in multiple interlocks. Here we do not include any interaction for CEO and multiple interlocks but include an indicator of whether a director has any CEO-experience. This covariate indicates if a director has been the CEO for a corporation for a year previous to the second time-point t_1 .

We define chair-experience in an analogous fashion to CEO-experience, namely if a director has been chairman of a board in the year previous to the second time-point. The chair has a central position in the board of directors and someone with chair-experience may therefore be regarded as having insight into the corporations on which he or she has served. Having chairman experience may also signal status and add to one's reputation.

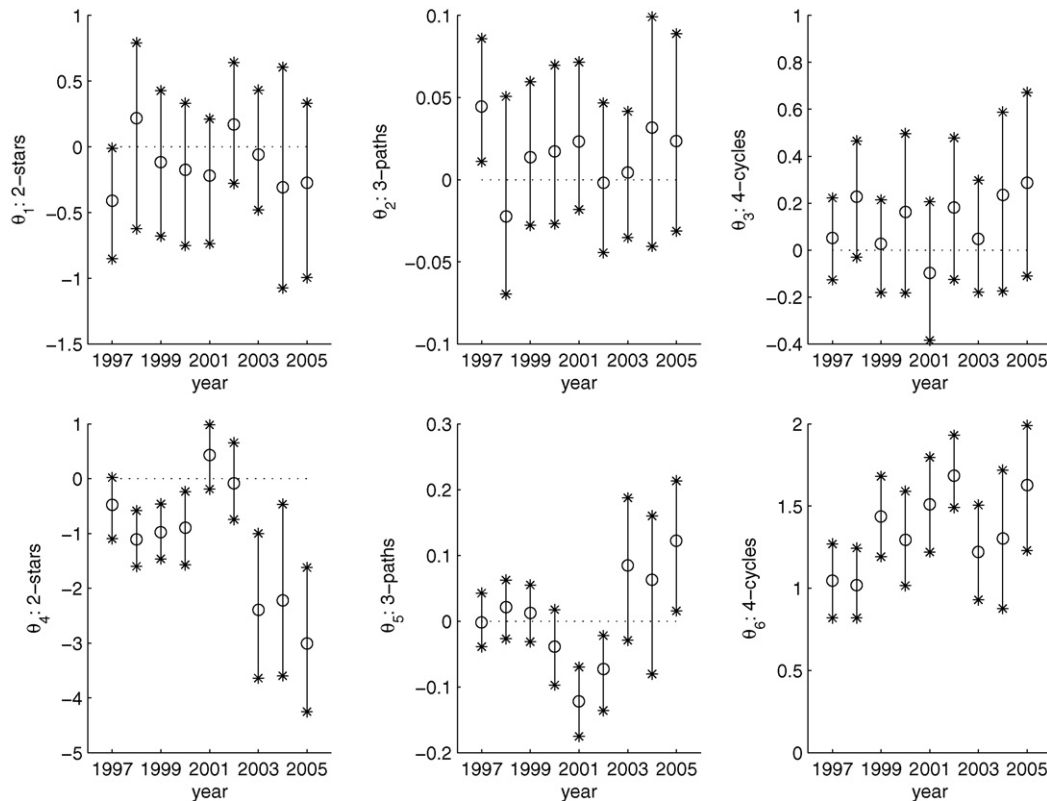


Fig. 5. Probability intervals (.99 equal tail probability) for parameters across time in the purely structural model. Parameters θ_1 through θ_3 correspond to two-stars (S_{B2}), three-paths (L_3), and four-cycles (C_4), for the dissolution of ties and parameters θ_4 through θ_6 correspond to two-stars (S_{B2}), three-paths (L_3), and four-cycles (C_4), for the creation of ties.

Table 3

Parameter estimates (posterior means and standard deviations) for longitudinal model fitted pair-wise to corporate boards on the Stockholm stock exchange. Structural effects, director attribute effects and the interaction of sex and peer referral for the *dissolution* of a tie. NB: "*" indicates that marginal posterior is improper.

	Effects dissolution							
	Dir. Two-star	Three-Path	Four-Cycle	CEO exp.	Chair exp	Seniority	Woman	Int. Wom. four-cycle
1996–1997	−0.280 (0.158)	0.033 (0.013)	0.101 (0.066)	0.276 (0.151)	−0.043 (0.168)	−0.093 (0.070)	−0.234 (0.283)	0.468 (0.762)
1997–1998	0.230 (0.151)	−0.023 (0.013)	0.252 (0.071)	0.579 (0.173)	−0.003 (0.151)	−0.130 (0.074)	−0.042 (0.321)	1.017 (0.701)
1998–1999	−0.151 (0.178)	0.017 (0.014)	0.041 (0.085)	0.118 (0.148)	0.052 (0.162)	−0.217 (0.070)	−0.262 (0.345)	−0.120 (0.576)
1999–2000	−0.086 (0.192)	0.009 (0.018)	0.143 (0.121)	−0.076 (0.126)	−0.305 (0.123)	−0.183 (0.062)	0.290 (0.433)	2.742 (0.849)*
2000–2001	−0.092 (0.222)	0.012 (0.021)	−0.104 (0.127)	0.000 (0.158)	0.138 (0.131)	−0.128 (0.085)	0.501 (0.299)	1.596 (1.558)*
2001–2002	0.227 (0.199)	−0.010 (0.019)	0.214 (0.101)	0.151 (0.127)	−0.238 (0.123)	−0.068 (0.061)	−0.015 (0.273)	1.268 (1.873)
2002–2003	−0.075 (0.206)	0.007 (0.018)	0.067 (0.118)	0.060 (0.117)	0.050 (0.133)	−0.267 (0.058)	0.919 (0.347)	1.349 (0.977)
2003–2004	0.071 (0.283)	0.002 (0.024)	0.103 (0.139)	−0.034 (0.121)	0.162 (0.165)	−0.352 (0.073)	0.188 (0.277)	5.605 (2.887)*
2004–2005	−0.404 (0.283)	0.035 (0.025)	0.310 (0.184)	−0.009 (0.176)	0.198 (0.189)	−0.130 (0.077)	0.068 (0.237)	0.089 (0.375)

Table 4

Parameter estimates (posterior means and standard deviations) for longitudinal model fitted pair-wise to corporate boards on the Stockholm stock exchange. Structural effects, director attribute effects and the interaction of sex and peer referral for the *creation* of a tie.

	Effects creation							
	Dir. Two-star	Three-Path	Four-Cycle	CEO experience	Chair experience	Seniority	Woman	Int. wom. four-cycle
1996–1997	−0.599 (0.202)	0.008 (0.017)	1.002 (0.073)	−0.106 (0.128)	0.359 (0.183)	−0.247 (0.046)	−0.270 (0.240)	0.362 (0.536)
1997–1998	−1.360 (0.161)	0.048 (0.014)	0.971 (0.084)	−0.136 (0.139)	0.663 (0.131)	−0.219 (0.048)	−0.208 (0.217)	−0.698 (0.801)
1998–1999	−0.796 (0.288)	0.002 (0.023)	1.354 (0.090)	−0.532 (0.177)	0.226 (0.178)	−0.282 (0.042)	−0.344 (0.230)	0.429 (0.413)
1999–2000	−0.554 (0.198)	−0.063 (0.019)	1.406 (0.096)	−0.341 (0.145)	0.106 (0.154)	−0.347 (0.045)	−0.033 (0.189)	−0.111 (0.639)
2000–2001	0.236 (0.271)	−0.097 (0.026)	1.527 (0.096)	−0.501 (0.147)	0.343 (0.161)	−0.267 (0.048)	−0.252 (0.222)	−0.451 (0.505)
2001–2002	−0.715 (0.219)	−0.011 (0.019)	1.662 (0.089)	−0.359 (0.148)	0.028 (0.129)	−0.219 (0.044)	0.373 (0.171)	−0.721 (0.709)
2002–2003	−1.698 (0.382)	0.039 (0.029)	1.163 (0.110)	−0.141 (0.128)	0.143 (0.207)	−0.315 (0.062)	0.866 (0.119)	0.161 (0.154)
2003–2004	−2.149 (0.379)	0.066 (0.029)	1.388 (0.121)	−0.490 (0.149)	0.059 (0.227)	−0.302 (0.050)	0.337 (0.138)	−0.872 (0.254)
2004–2005	−2.854 (0.452)	0.114 (0.037)	1.580 (0.133)	−0.452 (0.199)	0.139 (0.179)	−0.220 (0.063)	0.081 (0.171)	−1.200 (0.816)

Seniority may be important as both marker of position in status hierarchies as well as a control for natural demographic changes. Here we used age and to improve numerical stability, age is standardised for each time period.

The female representation on company boards has long been a hotly debated issue, and we have already alluded to the occurrence of and change in female board representation in the study period. But does increasing representation mean increasing interlocking when controlling for other factors? Are women fully integrated in the corporate structure and are they as powerful as men in terms of being central in the network? The indicator of whether a director is a woman or not also serves as the lower order term for the interaction effect of sex and four-cycles. As previously stated, the interaction term counts the number of structures corresponding to Fig. 2(b) and (c).

5. Results

For setting initial parameter values and variance–covariance matrix for the proposal distribution we followed Koskinen and Snijders (2007) and obtained the Hessian of and parameter values that maximised the likelihood for arbitrarily chosen shortest paths. The Metropolis–Hastings algorithm was implemented with a total of 4000 iterations and a burn-in period of 1000 iterations.

Summaries for the posterior of the purely structural model are given in Table 2 and comparisons of interval estimates are given in Fig. 5. The processes of appointments and releases are quite different. Apart from a tendency to preserve three-paths in 1996–1997, there are no structural effects in the release.

There is no desire (or an aversion in 2000–2002) to create bipartite director two-stars when recruiting directors, everything else given. Creating interlocks is thus not desirable to the board per se. Since the coefficient for three-paths is negative or zero there is no desire to link to boards (via directors) with many directors or desire to reach directors indirectly (to shorten path-ways in the bipartite

graph in an efficient manner; it appears to be a negative effect in 2000–2001 and in 2001–2002 and a positive effect in 2004–2005, otherwise no effect).

There is a marked desire to create four-cycles and in combination with the fact that there is no tendency towards preserving four-cycles, this is strong evidence of peer referral operating.

While parameter estimates may not be directly comparable over time due to changes in network compositions, the trends in Fig. 5, reveal that the evolution appears to go through three different phases. In 1996–2000 the model seems homogeneous with respect to the (negative) preference for interlocking directors (two-paths), three-paths and peer referrals; in 2000–2002, the clustering increases; in 2002–2005 the process returns to something resembling the first phase but with greater variation (less certainty in the parameter estimates).

Adding director attributes does not change the structural effects to any greater extent. Summaries of the posteriors for the model with director attributes are given in Table 3 for dissolution of ties and in Table 4 for creation of ties. Interval estimates for the parameters are given in Fig. 6.

There is a positive CEO-effect for dissolution of ties in 1997–1998 but not in any other year suggesting that it is not of great importance to retain directors with CEO-experience. For four periods there is a negative CEO-effect on recruitment. Looking at the distribution of the number of board appointments for directors with and without CEO-experience (table not reproduced here) there is no clear discernible difference other than that the proportion of directors with two or more seats is slightly larger for directors with CEO-experience and also that the proportion of isolate directors is higher for this group (at the second time-point). The latter means that many of the directors that leave the population have CEO-experience but few of the directors entering the population have CEO-experience (details may be obtained from the authors).

There is a consistent negative recruitment effect for seniority as well as for release in some periods. While there is a positive

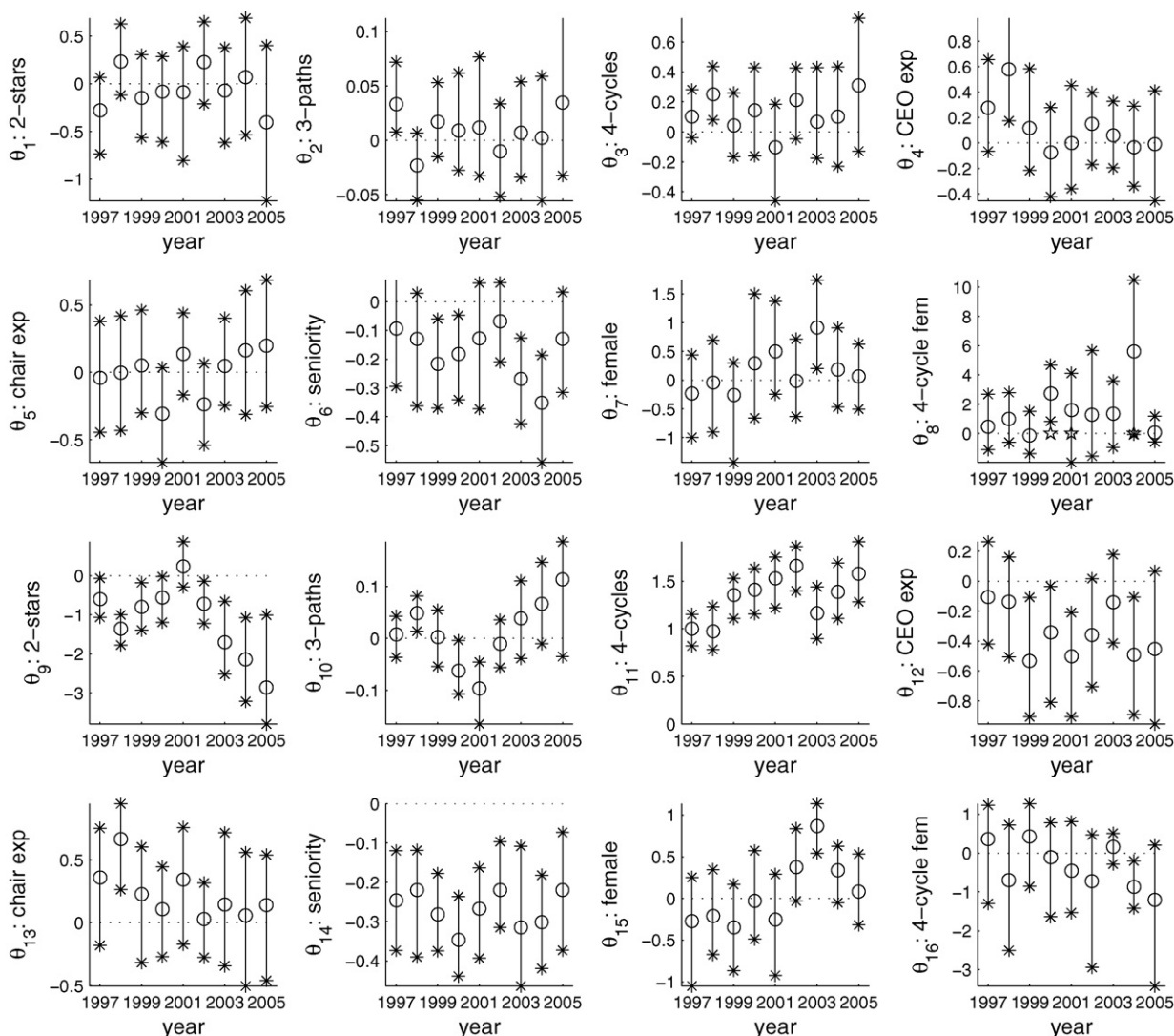


Fig. 6. Probability intervals (.99 equal tail probability) for parameters across time in model with structural effects, director attributes and interaction between sex and peer referral.

association between seniority and the number of board seats for most years cross-sectionally, the increase in the number of boards a director sits on is not positively related to age.

In 1997–1998 there is a positive effect for chair-experience but in other periods there is no attribute effect for having experience as chairman of the board.

There is no effect of sex in the period 1996–2001 but in 2001–2004 there seems to be a positive effect for recruitment (the .99-intervals just cover 0 in 2000–2001 and 2003–2004) and in 2002–2003 boards tend to want to keep female directors. This development is also reflected in Table 1, the transition from 2001 to 2002 is the year when the average number of board positions per women first reaches that of men. The following year there is also a preference for keeping these women, a tendency against firing women, which is reflected in the positive coefficient θ_7 in 2001–2002. Hence, in this period women hold more board positions than men on average. This is not explained by seniority, CEO or chairman-experience. As to how these women were recruited there seems to be a lot of uncertainty. There are not enough opportunities to create female interlocks since there were in general too few women (cf Table 1) and at the start of the series women accounted for too few ties (but from 2002 women occupy the same number of board seats as men on average).

The increased representation of women after 2001 is not accompanied by increasing peer referral. In fact, there is a tendency against referral of female directors in 2003–2004 – the posterior probability that the corresponding parameter is negative is practically 1 – and the following period the corresponding posterior probability is .95. We may tentatively conclude that while there is an increasing representation of women on the stock exchange, they are still not integrated in the male dominant power structures. Furthermore, the relative popularity of female directors is short-lived. We could speculate that there is an influx of token women in response to the threat of legislation.

We performed two robustness checks to assess the sensitivity of the model. To investigate the effects of the possible dependencies that may be induced by directors (permanently) leaving the population in a time interval, we fitted the model for each interval (t_0, t_1) excluding the directors that were isolates at t_1 (see Table 1). The differences in posteriors are overall small. For dissolution seniority in 1997–1998 has a negative effect, and in 1999–2000 and 2003–2004 no effect; and there is too little information to infer anything for the interaction of female director and four-cycles. For recruitment the effect of CEO disappears in 1998–2001 and 2003–2004; chair-experience becomes positive in 1996–1998 and 1999–2001; seniority has no effect in 2003–2004; the effect for female directors

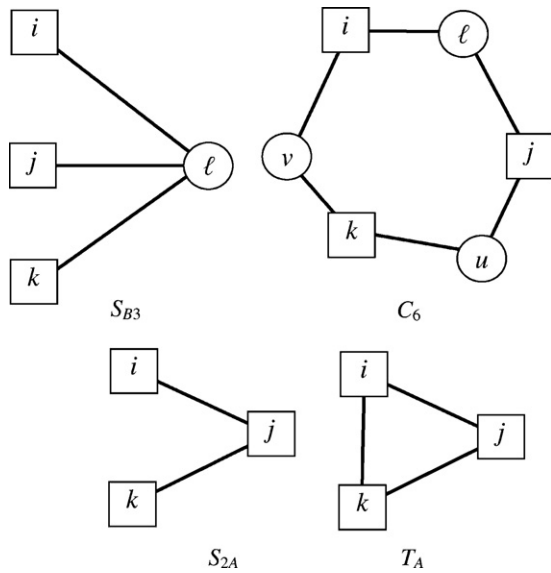


Fig. 7. A bipartite director three-star (S_{B3}), a bipartite 6-cycle (C_6), a unimodal board two-star (S_{2A}), and unimodal board triangle (T_A).

edges up above 0 and the negative effect of the interaction of female director and four-cycle weakens in 2003–2004. While the average age of isolate directors at time t_1 is consistently .5 standard deviation units higher than at time t_0 (meaning that directors that leave the population are on average older than directors that enter the population) this does not explain the negative effect of seniority. The previously mentioned demographics do cause CEO-experience to lose its effect when the isolates at the second time-point are excluded.

6. Goodness of fit

We may use Bayesian forecasting to assess the goodness of fit of a stochastic actor-oriented model (Koskinen and Snijders, 2007). If

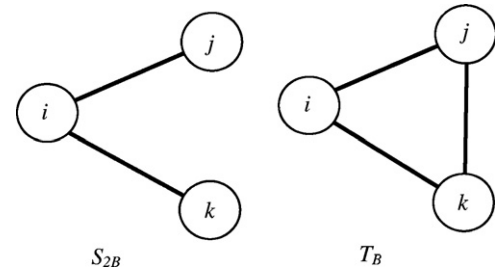


Fig. 8. A unimodal director two-star (S_{2B}), and unimodal director triangle (T_B).

the model is fit to observations $x(t_0)$ and $x(t_1)$, a third observation $x(t_2)$, which is not used in fitting the model, may be used for model evaluation. The idea is to assess how probable $x(t_2)$ is given $x(t_0)$ and $x(t_1)$, unconditional on the parameters. The distribution of $x(t_2)$ given $x(t_0)$ and $x(t_1)$ is unconditional on the parameters is the posterior predictive distribution $p(x(t_2)|x(t_0), x(t_1))$. Assuming that the model is time-homogenous, the posterior predictive probability is given by

$$p(x(t_2)|x(t_0), x(t_1)) = \int p(x(t_2)|\theta)\pi(\theta|x(t_0), x(t_1))d\theta, \tag{7}$$

in which $p(\theta)$ is the model defined by (4) marginalised with respect to the intermediate paths. It is very rarely possible to evaluate this integral analytically. Instead, we have to rely on Monte Carlo methods. In practice we simulate from the model for different parameter vectors, where we choose the parameter vectors in proportion to their posterior probability, their probability given $x(t_0)$ and $x(t_1)$. To choose these parameter vectors in proportion to their posterior probability we take the parameters from our posterior MCMC sample.

Predicting the individual elements of $x(t_2)$ may not necessarily be very informative and therefore we choose to look at a variety of functions of the bipartite network, such as the one-mode degree distributions and bipartite clustering. Since the inference for θ given

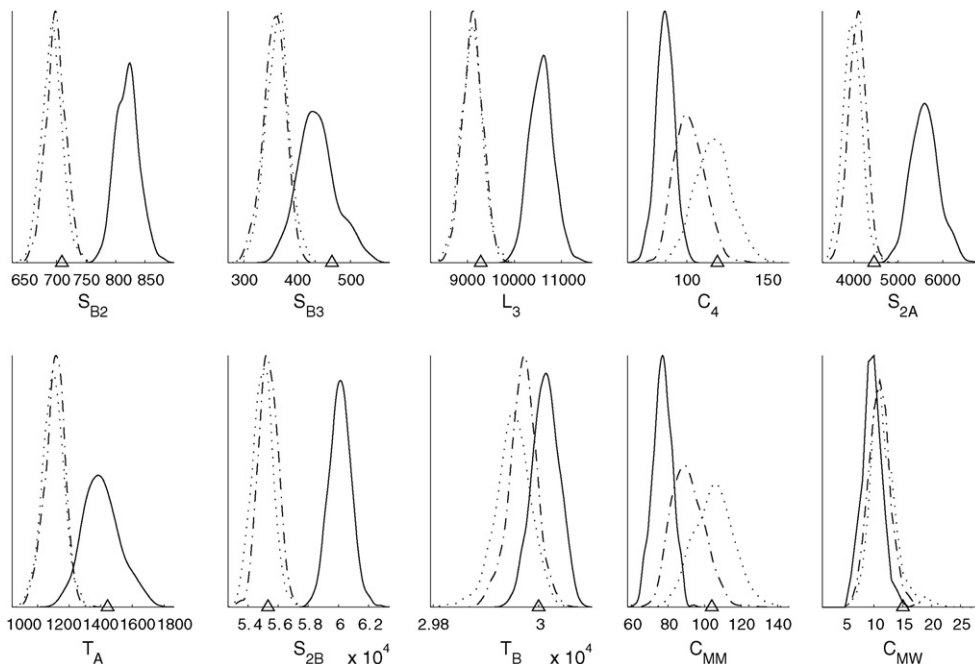


Fig. 9. Posterior predictive distributions for bipartite graph counts: observed count in 2005 (triangle); (true) posterior predictive distribution given observations in 2003 and 2004 (dashed dot); posterior predictive distribution using posterior distribution of model parameters conditional on given observations in 2004 and 2005 (dotted); base-line distribution with all parameters set to 0 (solid).

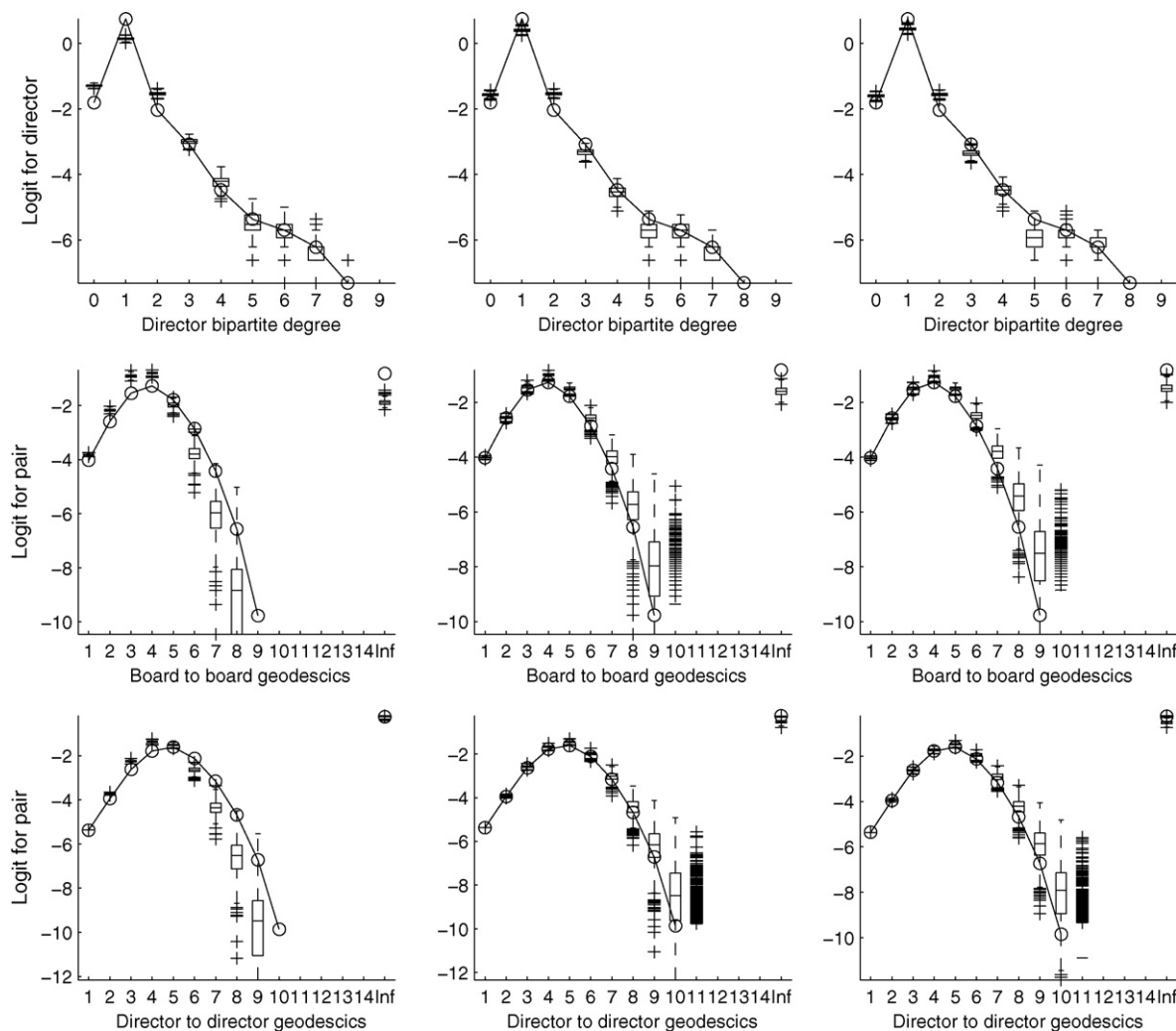


Fig. 10. Posterior predictive distributions for bipartite director degree, unimodal board geodesic distribution, unimodal director geodesic distribution on logit scale. Observed distributions in 2005 (circle solid line); columns left to right: base-line distribution with all parameters set to 0; (true) posterior predictive distribution given observations in 2003 and 2004; posterior predictive distribution using posterior distribution of model parameters conditional on given observations in 2004 and 2005.

$x(t_0)$ and $x(t_1)$ was conditioned on the walk length and thereby the sum of the variables $v_i(\phi, x)$ for each i , the posterior predictive distribution is similarly conditioned on the number of changes per board and the number of these changes that involved recruiting or releasing a director. Note that when this approximation is used then the posterior predictive distribution does incorporate information about the observation that is being used for validations and hence the distribution is not strictly speaking a true posterior predictive distribution as in (7).

We chose the observations in 2002 and 2003 as our $x(t_0)$ and $x(t_1)$ and use the posterior conditional on these two observations to draw from our approximation of (7). The results for a collection of graph functions are given in Figs. 9 and 10. For reference, the corresponding graph functions under a random walk are included. The random walk process is simulated in the same fashion as the posterior predictive distribution but with all parameters set to 0. This means that all sequences of changes that are concordant with the data structure are equally probable.

We have fixed the board sizes and hence any variation in unimodal board triangles (T_A in Fig. 7) is due to variation in overlapping director two-stars (S_{B2}). Because of structural restrictions all three of the models capture the unimodal director clustering (i.e. T_B). In the case of the random walk the ability to replicate T_B is in some

part due to over representing the number of unimodal director two-paths (S_{2B} , see Fig. 8). Consequently, if we were to compare the ratio of the number of T_B to S_{2B} for the different models, and use that as a measure of the unimodal director clustering, then the random walk would not capture the unimodal clustering.

The forecast makes an accurate prediction for the three-paths and four-cycles (Fig. 9). The random walks severely overestimates the number of three-paths and underestimates the number of four-cycles which, using the definition of Robins and Alexander (2004), means that the bipartite clustering coefficient is underestimated.

As seen in Fig. 9, the prediction model does not quite capture the number of S_{B3} (and neither does the model estimated using 2005). The random walk more closely manages to capture the number of S_{B3} . As a result the random walk also captures unimodal board triangles (T_A) better but produces too many unimodal board two-stars (S_{2A}). Note that apart from directors that sit on many boards, S_{B3} , the other distinct underlying bipartite structure that yields T_A is a 6-cycle, C_6 . On the whole, the forecast, the random walk and the fitted model do an equally good job of predicting the bipartite director degrees, or bipartite director stars, as seen in the top three panels in Fig. 10.

We can claim to have captured the substantive element of T_A , unimodal board clustering. We could possibly have improved on

the fit of S'_{B3} s but there is no substantive reason for including this in the model for a boards satisfaction with a new configuration and even less so for including 6-cycles.

Thus, there is a clustering of boards that is unexplained by the model but that is likely to be the result of heterogeneity in the popularity of directors. By the popularity of a director we refer primarily to his or her actual number of board appointments. Differences in popularity may be explained by underlying peer groups or may be reflected in or reflected by status and prestige. Differential status and popularity could be captured by including further director attributes but there could also exist a self-organising principle that explains status through reputation (Useem, 1984; Kaplan and Reishus, 1990; Zajac and Westphal, 1994; Podolny, 2005). Whether this status stems from the increased experience that comes with being central in the network or whether there was previously a peer dependent selection process so that multiple interlocks merely reflect high status and prestige in underlying social structures is an open question.

As an example of how well the model captures the other graph structures of the unimodal networks, we may compare the observed distribution of geodesic distances to the predicted (Fig. 10). The model does surprisingly well but some thought must be given to the inherent structural restrictions on the unimodal representations. This is evident in the unimodal projections of the graphs produced by the random walk in that the geodesic distributions closely resemble data. The random walk does however fail in capturing the tail of the distribution, not giving enough probability to long distances.

The prediction based on 2003–2004 largely replicates the number of four-cycles in 2005 (overall the number of predicted four-cycles is a little lower than observed and as predicted using 2005, which is a consequence of, among other things, the estimate for 2003–2004 being lower than that for 2004–2005). The random walk, by comparison, predicts far too few four-cycles. Since the random walk already predicts too few four-cycles it almost manages to predict the number of C_{MW} .

By including the base-line model we have shown that the fit of our model is not an artefact of the restrictions stemming from looking at the evolution from one point in time to another—there is sufficiently much change to highlight the processes at work in the data.

7. Discussion

We introduced a model for bipartite evolution drawing on Snijders' work on stochastic actor-oriented models, and showed how to apply this model for the study of interlocking directorates. Analysing data on firms traded at the Stockholm Stock Exchange between 1996 and 2005, we found that creating four-cycles seems desirable for corporate boards but that preserving them is not as important. We conclude that this is a sign of peer referral in the recruitment process, and that this is central to the evolution of the bipartite network. However, the increase in female directors during the early 2000s is not associated with any peer referral, suggesting that women were assigned to boards through other recruitment mechanisms. Lastly we showed by using time-series forecasting as cross-validation, that the fitted model managed to explain many features of the data that were not explicitly modelled.

In addition to the analysis conducted here there are some extensions that would be worthwhile pursuing. Although the strategic and instrumental interests of the corporations were not considered here it would be possible and useful to incorporate firm characteristics (e.g., Mizuchi and Stearns, 1988) and business group affiliation (Smangs, 2006), as well as board composition (e.g., Ruigrok et al., 2006) and additional director attributes. The

literature also contains frequent mentioning of underlying networks and these could be taken into account to some extent by including director affiliation network (clubs and affiliations from e.g., who's who). The question is whether the result would change considerably or whether these affiliations are already channelled through the social neighbourhoods and the peer referral process. A slightly different perspective would be to fully investigate the co-dependence of the board networks and other formal networks. It is a reasonable suggestion that there is co-dependence between the interlocking network and the ownership network (Bohman, 2006). To analyse either one as evolving exogenously and independent of the other would be similar to when dyadic board covariates are employed. A more substantive interest would be to consider the joint co-evolution of the interlocking directorate network and the ownership network. Needless to say any of these further explorations would require extensive data collection in addition to the data that we had available for the current analysis.

For bipartite networks other than interlocking directorates model specifications would necessarily have to be redefined to reflect differences in boundaries, type of actions, etc. We believe that the methodology as laid out here goes some way towards giving directions as to how that could be made and that this opens for plenty of future applications, as do the opportunities to relax some of the simplifying assumptions made here. In addition to these extensions we could think of fitting a joint model for evolution using change point analysis techniques or likelihood-based model selection (Snijders et al., in press); treating changes in the population using birth and death processes for directors and boards.

References

- Barnes, C.R., Ritter, E.R., 2001. Networks of corporate interlocking: 1962–1995. *Critical Sociology* 27, 192–220.
- Bohman, L., 2006. Economic action and interfirm relations: diffusion of stock repurchases on the Stockholm Stock Exchange 2000–2003. *European Sociological Review* 22, 383–396.
- Bohman, L., 2009. Interlocking directorates and director recruitment. Unpublished manuscript. Stockholm University.
- Borgatti, S.P., Everett, M.G., 1997. Network analysis of 2-mode data. *Social Networks* 19, 243–269.
- Breiger, R.L., 1974. Duality of Persons and Groups. *Social Forces* 53, 181–190.
- Breiger, R.L., Pattison, P.E., 1986. Cumulated social roles—the duality of persons and their algebras. *Social Networks* 8, 215–256.
- Butts, C.T., 2008. A relational event framework for social action. *Sociological Methodology* 38, 155–200.
- Carpenter, M.A., Westphal, J.D., 2001. The strategic context of external network ties: examining the impact of director appointments on board involvement in strategic decision making. *Academy of Management Journal* 44, 639–660.
- Doreian, P., 1979. On the delineation of small group structures. In: Hudson, H.C. (Ed.), *Classifying Social Data*. Jossey-Bass, San Francisco.
- Edling, C., Sandell, R., 2001. Social influence and corporate behaviour—a case study of interdependent decision-making in Sweden's publicly traded firms. *European Sociological Review* 17, 389–399.
- Feld, S.L., 1981. The focused organization of social ties. *American Journal of Sociology* 86, 1015–1035.
- Fich, E.M., White, L.J., 2003. CEO compensation and turnover: the effects of mutually interlocking. *Wake Forest Law Review* 38, 935–959.
- Fich, E.M., White, L.J., 2005. Why do CEOs reciprocally sit on each other's boards? *Journal of Corporate Finance* 11, 175–195.
- Frank, O., Carrington, P.J., 2007. Estimation of offending and co-offending using available data with model support. *Journal of Mathematical Sociology* 31, 1–46.
- Frank, O., Strauss, D., 1986. Markov graphs. *Journal of the American Statistical Association* 81, 832–842.
- Freeman, L.C., White, D.R., 1993. Using galois lattices to represent network data. *Sociological Methodology*, 127–146, vol. 23 23.
- Galaskiewicz, J., Wasserman, S., 1981. A dynamic study of change in a regional corporate network. *American Sociological Review* 46, 475–484.
- Galaskiewicz, J., Wasserman, S., Rauschenbach, B., Bielefeld, W., Mullaney, P., 1985. The influence of corporate-power, social-status, and market position on corporate interlocks in a regional network. *Social Forces* 64, 403–431.
- Holland, P.W., Leinhardt, S., 1977a. Dynamic-model for social Networks. *Journal of Mathematical Sociology* 5, 5–20.
- Holland, P.W., Leinhardt, S., 1977b. Social-structure as a network process. *Zeitschrift Fur Soziologie* 6, 386–402.
- Huisman, M., Snijders, T.A.B., 2003. Statistical analysis of longitudinal network data with changing composition. *Sociological Methods & Research* 32, 253–287.

- Izraeli, D.N., Talmud, I., 1997. Getting aboard: mode of recruitment and gender composition: the case of women directors in Israel. *International Review of Women and Leadership* 2, 26–45.
- Kanter, R.M., 1977. *Men and Women of the Corporation*. Basic books, New York.
- Kaplan, S.N., Reishus, D., 1990. Outside directorships and corporate performance. *Journal of Financial Economics* 27, 389–410.
- Katz, L., Proctor, C.H., 1959. The concept of configuration of interpersonal-relations in a group as a time-dependent stochastic-process. *Psychometrika* 24, 317–327.
- Koskinen, J.H., 2004. Bayesian inference for longitudinal social networks. Research report 2004:4. Stockholm University, Department of statistics. Available at <http://gauss.stat.su.se/site/pdf/RR2004.4.pdf>.
- Koskinen, J.H., Edling, C., 2004. The evolution of interlocking directorates. Paper presented at the XXIV International Sunbelt Social Network Conference, May 12–16, 2004, Portorož, Slovenia.
- Koskinen, J.H., Snijders, T.A.B., 2007. Bayesian inference for dynamic social network data. *Journal of Statistical Planning and Inference* 137, 3930–3938.
- Leenders, R., 1995. Models for network dynamics—a Markovian framework. *Journal of Mathematical Sociology* 20, 1–21.
- Mayer, T.F., 1984. Parties and networks—stochastic-models for relationship networks. *Journal of Mathematical Sociology* 10, 51–103.
- Mills, C.W., 1956. *The Power Elite*. Oxford University Press, New York.
- Mizruchi, M.S., 1996. What do interlocks do? An analysis, critique, and assessment of research on interlocking directorates. *Annual Review of Sociology* 22, 271–298.
- Mizruchi, M.S., Stearns, L.B., 1988. A longitudinal-study of the formation of interlocking directorates. *Administrative Science Quarterly* 33, 194–210.
- Newman, M.E.J., Strogatz, S.H., Watts, D.J., 2001. Random graphs with arbitrary degree distributions and their applications. *Physical Review E* 64, 026118.
- Ornstein, M.D., 1982. Interlocking directorates in Canada—evidence from replacement patterns. *Social Networks* 4, 3–25.
- Ornstein, M.D., 1984. Interlocking directorates in Canada: intercorporate or class alliance? *Administrative Science Quarterly* 29, 210–231.
- Palmer, D., 1983. Broken ties—interlocking directorates and intercorporate coordination. *Administrative Science Quarterly* 28, 40–55.
- Palmer, D., Singh, J.V., Friedland, R., 1986. The ties that bind—organizational and class bases of stability in a corporate interlock network. *American Sociological Review* 51, 781–796.
- Pattison, P., Robins, G., 2002. Neighborhood-based models for social networks. *Sociological Methodology* 2002, 301–337, Vol 32 32.
- Podolny, J.M., 2005. *Status Signals: A Sociological Study of Market Competition*. Princeton University Press, Princeton.
- Robins, G., Pattison, P., 2001. Random graph models for temporal processes in social networks. *Journal of Mathematical Sociology* 25, 5–41.
- Robins, G.L., Alexander, M., 2004. Small worlds among interlocking directors: network structure and distance in bipartite graphs. *Journal of Computational and Mathematical Organization Theory*, 69–94.
- Ruigrok, W., Peck, S.I., Keller, H., 2006. Board characteristics and involvement in strategic decision making: evidence from Swiss companies. *Journal of Management Studies* 43, 1201–1226.
- Sanil, A., Banks, D., Carley, K., 1995. Models for evolving fixed node networks—model-fitting and model testing. *Social Networks* 17, 65–81.
- Skvoretz, J., Faust, K., 1999. Logit models for affiliation networks. *Sociological Methodology*, 253–280, Vol 29 29.
- Small, H., 1973. Cocitation in scientific literature—new measure of relationship between 2 documents. *Journal of the American Society for Information Science* 24, 265–269.
- Smangs, M., 2006. The nature of the business group: a social network perspective. *Organization* 13, 889–909.
- Snijders, T.A.B., 1996. Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology* 21, 149–172.
- Snijders, T.A.B., 2001. The statistical evaluation of social network dynamics. *Sociological Methodology*, 361–395, Vol 31 31.
- 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., Schweinberger, M., in press. Maximum likelihood estimation for social network dynamics. *The Annals of Applied Statistics*. DOI: 10.1214/09-AOAS313.
- Snijders, T.A.B., Pattison, P., Robins, G.L., Handcock, M., 2006. New specifications for exponential random graph models. *Sociological Methodology* 36, 99–153.
- Snijders, T.A.B., Steglich, C.E.G., Schweinberger, M., Huisman, M., 2007. *Manual for Siena Version 3.1*. Groningen/Oxford, ICS Department of Sociology/Department of Statistics.
- Snijders, T.A.B., van Duijn, M., 1997. Simulation for statistical inference in dynamic network models. In: Conte, R., Hegselmann, R., Terna, P. (Eds.), *Simulating Social Phenomena*. Springer, Berlin, pp. 493–512.
- Stafsudd, A., 2006. People are strange when you are a stranger: senior executives select similar successors. *European Management Review* 3, 177–189.
- Useem, M., 1984. *The Inner Circle: Large Corporations and the Rise of Business Political Activity in the U.S. and U.K.* Oxford University Press, New York.
- Wang, P., Sharpe, K., Robins, G.L., Pattison, P.E., 2009. Exponential random graph (p^*) models for affiliation networks. *Social Networks* 31, 12–25.
- Wasserman, S., 1980a. Analyzing social networks as stochastic processes. *Journal of the American Statistical Association* 75, 280–294.
- Wasserman, S., 1980b. A stochastic model for directed graphs with transition rates determined by reciprocity. *Sociological Methodology* 11, 392–412.
- Wasserman, S., Iacobucci, D., 1988. Sequential social network analysis. *Psychometrika* 53, 261–282.
- Zajac, E.J., Westphal, J.D., 1994. The costs and benefits of managerial incentives and monitoring in large US corporations—when is more not better. *Strategic Management Journal* 15, 121–142.