Journal of the Royal Statistical Society Series A: Statistics in Society, 2024, **187**, 967–988
https://doi.org/10.1093/jrsssa/qnae016
Advance access publication 15 March 2024
Original Article

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS
Journal of the Statistics Society
Series A
Statistics in Society

A goodness of fit framework for relational event models

Viviana Amati[1], Alessandro Lomi[2] and Tom A. B. Snijders[3,4]

[1]Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Milan, Italy
[2]Faculty of Economics, Università della Svizzera italiana, Lugano, Switzerland
[3]Department of Statistics and Nuffield College, University of Oxford, Oxford, UK
[4]Department of Sociology, University of Groningen, Groningen, The Netherlands

Address for correspondence: Viviana Amati, Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Via Bicocca degli Arcimboldi 8, Milan 20126, Italy. Email: viviana.amati@unimib.it

Abstract

We introduce a novel procedure to assess the goodness of fit in relational event models. Building on existing auxiliary variable approaches developed in network modelling, the procedure involves a comparison between statistics computed on observed relational event sequences and statistics calculated on event sequences simulated from the fitted model. We argue that the internal time structure of the relational mechanisms assumed to generate the observations under the model is an important aspect of the fit of a model to observed relational event sequences. We establish the empirical value of the proposed goodness of fit approach in an analysis of data that we collected on collaborative patient-referral relations among healthcare organizations. The illustrative case study that we develop reveals distinctive features of relational event models that have been ignored or overlooked in received empirical studies.

Keywords: auxiliary variables, goodness of fit, healthcare organizations, interorganizational networks, relational event models, statistical models for networks

1 Introduction

Data generated by social interaction often take the form of a tuple (i, j, t, w) recording information on the behaviour that unit i (the 'sender'—or source of action) directs towards unit j (the 'receiver'—or target of action) at time t (Bianchi et al., 2024; Butts, 2008; Perry Wolfe, 2013). When available, information on qualities of the event connecting i and j may be summarized by the weight w (Brandes et al., 2009).

Behavioural units of interest may be represented by individuals (Stadtfeld & Block, 2017), groups of individuals (Niezink & Campana, 2023), corporate actors, such as formal organizations (), or even countries (Brandes et al., 2009). Examples of interactive behaviour defined at these various levels of analysis include conversations between individuals (Gibson, 2005), coordinated actions among groups of individuals (Bright et al., 2023), resource exchange between organizations (Vu et al., 2017), and cooperation between countries (Stadtfeld et al., 2017).

Relational event models (REMs) have proven particularly useful in empirical situations involving the analysis of time-stamped social interaction data (Butts, 2008; Butts et al., 2023). The main goals of REMs are to (i) determine which individual characteristics of i and j are more likely to facilitate (or impede) their interaction and (ii) identify configurations of past events in which i and j have been involved that make future interaction between them more (or less) likely to occur (). Cooperation and conflict in crowdsourced productions (Lerner & Lomi, 2020a), coordination among emergency response teams (Butts, 2008), interaction in online learning environments (Vu

Received: March 27, 2023. **Revised:** November 3, 2023. **Accepted:** January 23, 2024

et al., 2015), and animal social networks (Kings et al., 2023; Tranmer et al., 2015) are representative examples of recent studies demonstrating the diversity of empirical interests that REMs are able to sustain.

Initially proposed by Butts (2008), and later extended and refined by Brandes et al. (2009), Perry and Wolfe (2013), and Stadtfeld (2012), REMs involve the specification of effects to capture various forms of time dependence among events. Forms of dependence typically encountered in empirical studies of social networks include reciprocity—the tendency of symmetric relations to be more likely to be observed between two actors (Dakin & Ryder, 2020)—and path-shortening—the tendency of actors connected to one or more common thirds to become directly connected (Bianconi et al., 2014; Newman & Park, 2003; Robins et al., 2009; Snijders & Steglich, 2015).

More established statistical models for networks typically focus on the transition between mutually exclusive relational states such as, for example, 'connected' and 'disconnected'. REMs depart from this tradition of network modelling by taking as input information on sequences of observed time-ordered events, rather than the presence or absence of concurrent network ties encoded into an adjacency matrix (Bianchi & Lomi, 2022; Butts, 2008, 2009; Butts et al., 2023). Consequently, REMs involve multiple forms of time dependence linking events in a sequence. Consider reciprocity, for example. For an event directed from behavioural unit i to j observed at time t to contribute to reciprocity, an event flowing in the opposite direction (from j to i) must have occurred at some time t^- prior to t. The order in which these events happen and the time elapsing between them matter because the possibility of observing 'reciprocity' within any given time window depends on the speed of the underlying process of 'reciprocation', which is not directly observed (Bianchi et al., 2024).

This argument implies that the network-like effects typically included in empirical specifications of REMs have an internal temporal structure that makes it conceptually difficult to interpret the estimate of their associated parameters as conventional 'effects' in comparable event history models. Continuing with our example, depending on the relative speed of reciprocation, reciprocity may be immediate, or it may become observable only with a delay. In the typical case, there will be a distribution of observed 'time to reciprocity'—in other words, of waiting times before a symmetrizing event occurs. This is true more generally for all network-like effects computed on time-stamped data, such as transitive closure. In this case, the antecedent two-path configuration $i \to h \to j$ must be already formed when the path-shortening event $i \to j$ occurs. What matters in these examples is that in the 'event world' of REMs, the analogues of network effects have an internal temporal structure determined by the sequential nature of relational event data (Bianchi et al., 2022).

While the fact that event times are random should not be surprising, its implications for how network 'ties' can be constructed by aggregating time-dependent sequences of relational events are less obvious. Typically, information on the internal time distribution of local configurations of events is lost when events are aggregated into network ties over conventional time periods (Bianchi & Lomi, 2022).

One of the primary motivations for developing a relational event framework for directed social interaction processes (Perry & Wolfe, 2013) has been the possibility of including information on the timing of relational events that may 'lie beneath' network ties (Butts & Marcum, 2017) to account for the sequential constraints that shape—and at the same time emerge from social interaction (Gibson, 2005). What REMs add to our ability to analyse data with complex (temporal) dependencies is the possibility of incorporating information on the timing of individual relational events embedding individual behaviour in time-varying networks of dependence relations (Butts et al., 2023). The timing of events and their temporal ordering represent the most distinctive elements in the specification and evaluation of REMs as models for data (). For this reason, it is surprising that available approaches to assessing the goodness of fit (GOF) of REMs do not explicitly consider the timing of the event as a dimension of interest.

Next to model-based assessments of fit, such as hypothesis testing, or comparison of likelihood-based information criteria, it may be helpful in data analysis to assess GOF by comparing the implications of the estimated model with the observed data without leaning on the model in the comparison. A parallel non-network data are that, for regression analysis, next to hypothesis testing and assessment of R^2, there is a helpful role for tests of normality. This is especially important for network data, because of the strong statistical dependencies involved, and the lack of general principles for how to approximately tackle these dependencies.

Hunter et al. (2008) and Lospinoso and Snijders (2019) proposed and developed a general approach for assessing the fit of network models where the estimated model is used to simulate hypothetical observations with the same structure as the observations. Auxiliary statistics are then computed for the simulated hypothetical and the observed data, and their degree of similarity is used to assess the fit of the model. Since network data are notoriously multifaceted (Robins et al., 2009), the auxiliary statistics chosen for a particular case can be very diverse, depending on the context and purposes of the study. It is especially helpful to use auxiliary statistics that are not directly used for estimating the parameters of the model, but provide an important complementary view of the network data. The reasoning behind this claim is that the ability of the model to fit dimensions of the data that are not explicitly represented in the empirical model specification provides useful information on the GOF of the model.

In REMs, the internal time distribution of network mechanisms assumed to generate the observations is of fundamental importance, but cannot be directly represented by a specific 'effect' in the empirical model specification. The approach based on the simulation of auxiliary variables from the model as originally proposed by Hunter et al. (2008) might provide a particularly stringent test of the GOF for REMs. This does not replace significance tests or the use of information criteria, but is an addition providing another perspective.

The main contribution of this paper is to extend this approach to REMs by providing a set of principled operational criteria that may be generally adopted in empirical research to assess the extent to which assumptions underlying fitted REMs are consistent with the underlying data generating mechanisms.

In current empirical practice, assessing the GOF of REMs relies primarily on conventional information-based criteria such as the Akaike information criteria (AIC) or the Bayesian information criteria (BIC) (Butts, 2008; Lerner & Lomi, 2020a), which compare models with different specifications to a baseline model and therefore constitute relative fit measures useful for model comparison. Other approaches apply the traditional methods of GOF for event history models, such as martingale and Schoenfeld residuals (Perry & Wolfe, 2013; Vu et al., 2017).

Less frequently, prediction error (Stadtfeld & Block, 2017) and backward forecasting based on comparison of actual and simulated event sequences (Brandenberger, 2019) have been adopted to assess the GOF for REMs. None of these widely adopted procedures provides a test of the consistency of the model with the internal time structure of the mechanisms that are assumed to generate the observations.

The empirical value of the approach that we propose is tested in the context of healthcare—an empirical setting that illustrates vividly the importance of timing in social processes. Reactions of patients to treatment are highly sensitive to the timing of treatment itself (Gupta et al., 2012; Redelmeier & Bell, 2007). When the quality of care depends on the timing of treatment (Nallamothu et al., 2005), and the timing of treatment depends on the quality of coordination between healthcare organizations and ensure continuity of care (Gittell et al., 2000), then the time structure of coordination mechanisms among healthcare providers is of crucial importance in interhospital patient referral and transfer (Amati et al., 2019). Similarly important is the ability of a model to reproduce with high fidelity the internal timing of the relational mechanisms it postulates. With few recent exceptions (Bianchi et al., 2022), empirical studies have not established the adequacy of a network model in terms of its ability to reproduce with accuracy the internal time structure of the theoretical mechanisms it postulates.

After this general introduction, the article is organized as follows. In the next section, we establish the minimal notation necessary to summarize REMs and understand their inferential logic. In Section 3, we define some of the main statistics typically included in empirical model specifications. In Section 4, we adopt and adapt to REMs existing approaches developed for assessing the GOF of different statistical models for networks. In Section 5, we develop our illustrative empirical case to demonstrate the empirical value of our approach. The article concludes with a discussion section about the promises and limitations of the method proposed for assessing the GOF of REMs.

2 Relational event models

2.1 Notation

Let $\mathcal{N} = \{1, \ldots, n\}$ be the set of nodes in a network. The triplet $e = (i, j, t, w)$ denotes a relational event from node $i \in \mathcal{N}$ to node $j \in \mathcal{N}$, with weight w, observed at time $t \in \mathcal{T}$, $\mathcal{T} \subseteq \mathbb{R}_{\geq 0}$.

In addition to information on relational event sequences, information on monadic and dyadic attributes of the units (or 'nodes') may also be available. Monadic attributes describe individual characteristics of the units. Dyadic attributes are characteristics of pairs of units and may be expressed, for example, as functions of monadic attributes. We denote by v the $n \times p$ case by variable matrix of p monadic covariates. Each column of v is a variable mapping the set of units $\mathcal{N}$ to the set of attribute values $\mathcal{V}$. We denote by w an $n \times n \times q$ array of q dyadic covariates. Each matrix of the array represents a variable mapping the set of dyads $\mathcal{N} \times \mathcal{N}$ to the set of attribute values $\mathcal{W}$. Monadic and dyadic covariates can be constant over time or time-dependent. We denote time dependence by $v(t)$ and $z(t)$, respectively.

The collection of triplets (i, j, t), denoted by the set $E_i = \{(i, j, t) \mid j \in \mathcal{N}, j \neq i, t \in \mathcal{T}\}$, represents the sequence of relational events initiated by i over the time period $\mathcal{T}$. We denote the relational event data by $E = \{E_i \mid i \in \mathcal{N}, v(t), z(t)\}$.

Henceforth, we shall follow the convention that upper-case letters represent random variables and lower-case letters denote their realizations.

2.2 A marked point process model for directed interaction

A number of models for relational events have been proposed during the last decade (Brandes et al., 2009; Butts, 2008; Lerner & Lomi, 2023; Perry & Wolfe, 2013; Stadtfeld & Block, 2017; Vu et al., 2011). While the GOF approach that we propose applies, in principle, to any of these models, we focus our attention on the model developed by Vu et al. (2011, 2015). The advantage of this model for our current purposes is that model specification and comparison are greatly facilitated by its adoption in prior empirical studies of interhospital patient referral (Lomi et al., 2014; Vu et al., 2017). The second advantage of adopting this model is its explicit connection to the well-established theory of stochastic point processes (Cox & Isham, 1980).

The model assumes that the observed sequence of relational events is the outcome of a multivariate marked point process (MPP) composed of a simple point process describing the occurrence of events over time, and a mark, defining characteristics of the events. The simple point process models the occurrence of the relational events initiated by i, while the mark distribution describes the receiver j.

Formally, the multivariate marked temporal point process is modelled as a multivariate count process denoted by $N(t) = \{N_i(t) \mid i \in \mathcal{N}\}$, with univariate component $N_i(t) = |E_i(t)|$—representing the number of events initiated by i up to time t.

The model makes three major assumptions. First, only one relational event can occur at any given point in time (Lerner & Lomi, 2022). Second, censoring depends solely on the covariates included in the model (Blossfeld et al., 2014). Third, the univariate count processes $N_i(t)$ are adapted for all i to the same history of the process denoted by H_{t^-} with t^- a time that is infinitesimally smaller than t (Vu et al., 2011).

Under these conditions, the MPP is defined by its conditional intensity function (Aalen et al., 2008; Jacobsen, 2006) taking the form

$$\lambda_i(t, j \mid H_{t^-}) = \lambda_i(t \mid H_{t^-}) \cdot p_i(j \mid H_{t^-}), \tag{1}$$

and representing the influence of the history of past events on the expected occurrence of future events.

The first factor $\lambda_i(t \mid H_{t^-})$ in equation (1) is called *ground intensity function*. It models the simple point process and describes the rate at which a unit i 'sends' (or 'emits') a new event. The letter i in the subscript indicates that units can initiate events at a different rate, and thus the ground intensity function accounts for heterogeneity between units. The ground intensity function is specified as a Cox proportional hazard model:

$$\lambda_i(t \mid H_{t^-}) = \lambda_0(t) \cdot \exp(\theta' s(t, i)), \tag{2}$$

with $\lambda_0(t)$ a positive-valued function called baseline rate, θ' the vector of parameters, and $s(t, i)$ the vector of statistics. The statistics $s(t, i)$ represent the endogenous and exogenous variables that might affect the rate. Endogenous variables depend on previous events, while exogenous variables on attributes of the unit i initiating the event.

The second term $p(j \mid H_{t^-})$, called the *mark distribution*, probabilistically determines the receiver of the relational event initiated by i, i.e. it models the mark of the event. The mark distribution is multinomial

$$p_i(j \mid H_{t^-}) = \frac{\exp(\beta' s(t, i, j))}{\sum_{h \in \mathcal{R}_j(t)} \exp(\beta' s(t, i, h))}, \tag{3}$$

with β' the vector of parameters and $s(t, i, j)$ the vector of statistics. The statistics are counts of event local configurations encoding the mechanisms that might determine the marks and may depend on previous events as well as on monadic and dyadic covariates. The term $\mathcal{R}_j(t)$ is the risk set, i.e. the set of potential receivers of the event initiated by i. Under the assumption that events are non-reflexive, the risk set is defined as $\mathcal{R}_j(t) = \{j \mid j \in \mathcal{N}, j \neq i\}$, otherwise $\mathcal{R}_j(t) = \mathcal{N}$.

Examples of the statistics of the ground intensity and mark distribution are described in Tables S1.1 and S1.2, online supplementary material. Endogenous statistics are aggregations of previous events. It is often assumed that recent events are more relevant to the occurrence of the next event than events that happened far away in the past. The relevance of a past event is determined by a decay function. Different decay functions have been proposed. Butts (2009) defines weights that decay with the order of the events. Brandes et al. (2009) suggested using an exponential decay, while Vu et al. (2017b) a power-law decay depending on factors describing the rate of decay. Stadtfeld and Block (2017) adopted a time-window approach, whereby past events in a predefined time window of fixed width have weight 1, while events outside the window have a null weight. A mixture of these two approaches has been used to distinguish between the temporal relevance of short-term and long-term events (Lomi & Bianchi, 2024; Vu et al., 2015; Zappa & Vu, 2021). The choice between the different decay functions is often determined by a mix of empirical and theoretical considerations (Bianchi & Lomi, 2022).

Different estimation methods may be adopted, depending on the specification of the baseline rate $\lambda_0(t)$. If a functional form for $\lambda_0(t)$ is assumed, then the parameters of the model may be estimated via maximum likelihood. If the function $\lambda_0(t)$ is left unspecified, and the model is semi-parametric (Vu et al., 2017), estimation via partial likelihood approach is possible.

Interpreting REM estimates is not straightforward, even though the rate function and the mark distribution are defined in terms of a Cox regression and a multinomial logit model, respectively. The standard interpretation of the parameters as risk ratios and odds ratios has—at best—a heuristic value due to the structural correlation among the statistics. Thus, in the example that we present below, we focus on the significance and the sign of the coefficients. Significant and positive (negative) values of a parameter indicate that the probability of an event increases (decreases) with the value of the corresponding statistic, thereby providing evidence in favour of (against) the hypothesis that the corresponding mechanism contributed to generate the observations.

3 GOF for REMs

3.1 Related work

When is a REM an adequate model for data?

Extant research offers only an indirect answer to this question—an answer that typically depends on contingent empirical concerns. Our objective in this work is to propose a general analytical framework to evaluate the GOF of REMs that applies widely to a variety of practical research problems and underlying empirical model specifications (Butts, 2017; Stadtfeld et al., 2017).

Earlier work typically adopted relative measures of GOF defined in terms of likelihood-based information criteria such as the AIC and the BIC (Butts & Marcum, 2017) to compare fitted (full) models to suitable null models. While useful for model comparison and selection, an approach that relies on a single numerical indicator to assess the ability of a model accurately to represent multifaceted network data generated multiple dependencies is unlikely to support a fully satisfactory model evaluation strategy (Hunter et al., 2008; Lospinoso & Snijders, 2019).

In an analysis of REMs fitted to email communication data, Perry and Wolfe (2013) rely on deviance to assess model specification. In this study, the model GOF is evaluated in terms of the reduction of residual deviance due to the sequential introduction of additional factors in the empirical specification. Perry and Wolfe (2013, p. 835) recognize that deviance may not be an ideal measure of GOF because it depends entirely on the estimated parameters, and they propose inspection of a normalized version of the martingale residuals as alternative GOF diagnostic (Therneau et al., 1990).

As discussed in Vu et al. (2017), when REMs are estimated via Cox proportional hazard, a common GOF test is based on the analysis of the Schoenfeld residuals (Schoenfeld, 1982) to detect sources of a possible violation of the proportionality assumption (Grambsch & Therneau, 1994). While not originally conceived as a way to assess a model's GOF, this approach may assist in identifying specific sources of violation of proportionality assumptions.

Occasionally, the fit of REMs has been evaluated in terms of predictive accuracy. The probability of an event is computed conditional on the estimated model. If the predicted probability is larger than a conventional threshold, then the model is assumed to predict the occurrence of the event. The proportion of events predicted to occur over the events that could have occurred is used as a measure of the fit of the model (Brandenberger, 2019; Butts et al., 2007; Stadtfeld & Block, 2017). Since model predictions are sensitive to the choice of threshold, which is—to some extent—arbitrary, the GOF is evaluated visually by looking at the curve obtained by plotting the proportion of correctly predicted events associated with different threshold values. Recent work suggests that predictive accuracy may not be the most useful measure to evaluate the 3 of REMs (Schecter & Quintane, 2021).

Assessing the GOF of REMs presents difficulties that are common in statistical models for network data adopted to estimate the conditional probability of either observing the presence of network ties (Snijders et al., 2006) or change in network ties (Block et al., 2018). For this reason, current best practices in evaluating network models may provide a useful starting point to develop valuable procedures to assess the GOF of REMs.

Assessing the GOF of statistical models for networks often relies on a model evaluation procedure originally introduced for exponential random graph models (Hunter et al., 2008; Robins et al., 2009), and later adapted and extended to stochastic actor-oriented models by Lospinoso and Snijders (2019). According to this procedure, a good model should be able to explain (or 'fit') features of network data that have not been included in the empirical model specification and are not the explicit target of the fit optimization that is carried out by the parameter estimation. The value of these features (called auxiliary variables) in the data is compared with their distribution implied by the fitted model. Auxiliary variables frequently considered of empirical interest or theoretical importance may include, for example, the out-degree and in-degree distribution, the triad census, and the distribution of geodesic distances (Lospinoso & Snijders, 2019; Robins et al., 2005). Similar distributions can be computed for REMs as suggested by Brandenberger (2019).

3.2 Assessing the GOF of REMs

Similarly to other statistical models for networks, at the heart of REMs is the assumption that dependencies linking behavioural units are the primary force driving the formation of network structures (Hunter et al., 2008; Pattison & Robins, 2002). Unlike other network models where dependencies are represented in terms of concurrent network ties, REMs represent dependencies as emerging directly from sequences of observed events ordered in time (Butts & Marcum, 2017). While timing plays no special role in available models for relational states, it is a defining element in models for relational events.

Networks 'effects' computed from sequences of time-stamped relational event data have an explicit time extension. This is the case because it typically takes some time for an event sequence to unfold and give rise to patterns of dependence with theoretically interesting self-organizing properties.

For example, some delay is likely to occur between one event flowing from i to j at time t, and the symmetrizing event flowing in the opposite direction at time $t' > t$. When the difference between t and t' is not fixed but varies across events, time to reciprocation will follow a distribution of waiting times where some events are reciprocated quickly, while some other events may be reciprocated after a longer time (Bianchi et al., 2022). Events that are not reciprocated within the observation period are considered censored observations (Bianchi et al., 2022).

Similarly, it takes time for any two-path event sequence $(i \to h \to j)$ observed before time t to generate a path-shortening event $(i \to j)$ that produces triadic closure at some future time $t' > t$. Like reciprocity, time to transitive closure is likely to give rise to a distribution of 'time-to-closure' spells whereby a path-shortening event may close an open two-path quickly and contribute to clustering of events in the shorter term, while some other two paths may remain open and operate only over a longer term. As it was the case for reciprocity, two-path event sequences that remain open at the end of the observation period may be considered censored observations. For any given observation period of fixed length, the internal clock of network mechanisms such as reciprocity and transitivity reflects the relative speed at which these mechanisms operate to affect the occurrence of future events (Bianchi et al., 2022; Stadtfeld & Block, 2017).

The stylized examples we discussed suggest not only that the generating mechanisms incorporated in REMs have internal time extension but also that they are likely to play themselves out over different time scales (Butts, 2009)—and at different speeds. Perhaps the most ambitious promise of REMs is to 'reveal how sequences of interactions operate within short and longer time intervals' so that REMs might 'open new insights into the relative timing of social network mechanisms' (Stadtfeld & Block, 2017, p. 319).

For this reason, a particularly appropriate and stringent test for the GOF of a REM—however specified—involves an assessment of the ability of the model to reproduce the internal timing of the dependence mechanisms it postulates. Models consistent with the internal time structure of the network mechanisms generating the observations capture an essential aspect of REMs that distinguishes them from available statistical models for networks: the possibility of exploiting information produced by the exact timing of events.

In the section that follows, we develop this line of argument and link it to the auxiliary variable approach developed by Hunter et al. (2008) to assess the GOF of REMs. Clearly, the importance of capturing with accuracy the internal time structure of relational mechanisms may vary across empirical settings. Yet, understanding the differences in relative speed at which the various mechanisms operate is an issue of general theoretical importance that network models have generally been unable to address (Butts, 2009). As we explain below, and as we discuss in greater detail in the illustrative part of the paper, we suggest that—other conditions being equal—a model able to reproduce these differences with high fidelity is a more desirable model for relational event data than a model that does not.

3.3 Auxiliary statistics

For the purpose of developing an auxiliary variable approach to the evaluation of the GOF for REMs, we restrict attention to three network mechanisms whose relevance has been documented extensively in empirical network research: reciprocation and path-shortening via transitive and cyclic closure (Robins et al., 2009). These mechanisms are known to play an essential role in empirical research on social (Newman & Park, 2003) and other kinds of networks (Milo et al., 2002; Newman et al., 2002; Robins et al., 2005). The same mechanisms also figure prominently in theories of, and empirical research on interorganizational relations—the specific context of the empirical illustration that we develop in this paper (Atouba & Shumate, 2010; Kitts et al., 2017; Laumann & Marsden, 1982; Lomi & Pattison, 2006).

The proposed auxiliary statistics describe the distribution of the internal time of a specific mechanism. Define the ordered set of times where events $i \to j$ occur by $\{t_1, \ldots, t_M\} = \{t \mid (i, j, t) \in E_i\}$ and let t_0 be the first time in $\mathcal{T}$.

The first internal timing we consider is the time to reciprocation. The condition for reciprocation of an event $i \to j$ is the existence of (at least) one previous event $j \to i$ in the opposite direction. For the times $t_m \in \{t_1, \ldots, t_M\}$ (with $1 \le m \le M$), define $E^{rec}_{ji}(t^-_m) = \{(j, i, t') \in E_j \mid t_{m-1} < t' < t_m\}$ as the set of all earlier events $j \to i$ that have not yet been reciprocated at t_m. If this set is empty, event (i, j, t_m) does not contribute to the computation of the auxiliary statistic. For t_m where this set is non-empty, the time to reciprocation for t_m is $t_m - \min\{t' \mid (j, i, t') \in E^{rec}_{ji}(t^-_m)\}$.

The second internal timing of interest is the time to transitive closure. The condition for transitive closure of an event $i \to j$ is the existence of (at least) one two-path[1] $i \to h \to j$. For the times $t_m \in \{t_1, \ldots, t_M\}$ (with $1 \le m \le M$) and for every $h \in \mathcal{N}, h \neq i, j$, define $E^{tr}_{ihj}(t^-_m) = \{(i, h, t') \in E_i, (h, j, t'') \in E_i \mid t_{m-1} < t', t'' < t_m\}$ as the set of all two paths $i \to h \to j$ that have not yet been close at t_m. If this set is empty, event (i, j, t_m) does not contribute to the computation of the auxiliary statistic. For t_m where this set is non-empty, the time to transitive closure for t_m is $t_m - \min\{t' \mid (i, h, t') \in E^{tr}_{ihj}(t^-_m),\ t'' \mid (h, j, t'') \in E^{tr}_{ihj}(t^-_m)\}$.

The condition for cyclic closure of an event $i \to j$ is the existence of (at least) one two-path $j \to h \to i$. For the times $t_m \in \{t_1, \ldots, t_M\}$ (with $1 \le m \le M$) and for every $h \in \mathcal{N}, h \neq i, j$, define $E^{cy}_{jhi}(t^-_m) = \{(j, h, t') \in E_j, (h, i, t'') \in E_h \mid t_{m-1} < t', t'' < t_m\}$ as the set of all two paths $j \to h \to i$ that have not yet been close at t_m. If this set is empty, event (i, j, t_m) does not contribute to the computation of the auxiliary statistic. For t_m where this set is non-empty, the time to cyclic closure for t_m is $t_m - \min\{t' \mid (j, h, t') \in E^{cy}_{jhi}(t^-_m),\ t'' \mid (h, i, t'') \in E^{cy}_{jhi}(t^-_m)\}$.

The definition and computation of the network statistics hinge on several assumptions. By using the minimum we assume that the clock regulating the internal timing of a given mechanism starts ticking as soon as the antecedent (or predecessor) configuration emerges. Moreover, we do not account for the recurrence of multiple conditions before the mechanism is completed, e.g. situations in which multiple events $j \to i$ occurred before observing $i \to j$. Appendix S3, online supplementary material, provides the analogous definitions of the statistics for the cases in which it is crucial to account for repeated events or when the clock should start ticking from the time in which the most recent antecedent (or predecessor) configuration emerges.

The second assumption is that the clock that controls the time needed for an event to turn a predecessor configuration (for example, any two-path sequence) into its associated successor configuration (for example, a triangular sequence) stops ticking as soon as the event occurs.[2] When this happens, the antecedent configuration is removed from the sample, and the presence of subsequent events is ignored in the computation of the auxiliary statistics.

The last assumption concerns the definition of the sets $E^{rec}_{ij}(t^-_m)$, $E^{tr}_{ihj}(t^-_m)$, and $E^{cy}_{jhi}(t^-_m)$, which includes all the events occurred up to time t_m and have not (or not yet) components of more complex local configuration. Clearly, the composition of this set of events might vary depending on the context, case study, and type of relational events.

Table 1 illustrates the computation of the internal time for reciprocity, transitive closure, and cyclic closure, defined as the time elapsed between the formation of the condition and that of an event leading to a given configuration. The internal time for other network effects frequently included in empirical model specifications, for example, sending and receiving balance (Vu et al., 2017), may be defined similarly.

The definition of the internal time of the mechanisms is the component of the auxiliary statistics. Let $\phi_M(m)$ be the empirical distribution of the internal time of a specific mechanism. The auxiliary statistics for each mechanism are the deciles of the internal time empirical distribution computed using the generalized inverse of the empirical distribution function.

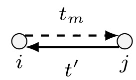
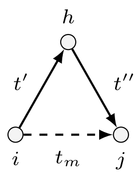
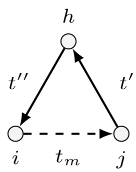
3.4 Implementation

Building on Hunter et al. (2008), we evaluate the fit of a REM by comparing the distributions of the auxiliary statistics computed on the observed data with the distributions of the auxiliary statistics computed on sequences of events simulated from the estimated model. Given an estimated model, the sequence of events is generated from the MPP model described in Section 2.2 by conditioning on the times of the observed events and repeating those two steps for each event:

[1] Following the mathematical definition of transitivity, we assume that the order of the events leading to the two-path is irrelevant, i.e. we do not distinguish between the case in which $i \to j$ is created before $h \to j$ and the case in which $h \to j$ is created before $i \to j$. However, statistics accounting for the order of events in the predecessor configuration can be computed if theoretically relevant or empirically important for understanding the phenomenon under investigation.

[2] The notions of predecessor and successor configurations are adapted, loosely, from the mathematical theory of discrete dynamical systems (Wolfram, 1984).

Table 1. Illustration and definition of the internal time of basic network mechanisms for relational event data

Reciprocation		$t_{rec} = t_m - \min\{t' \mid (j, i, t') \in E^{rec}_{ji}(t_m^-)\}$
Transitive closure		$t_{tr} = t_m - \min\{t' \mid (i, h, t') \in E^{tr}_{ihj}(t_m^-),\ \ t'' \mid (h, j, t'') \in E^{tr}_{ihj}(t_m^-)\}$
Cyclic closure		$t_{cy} = t_m - \min\{t' \mid (j, h, t') \in E^{cy}_{jhi}(t_m^-),\ \ t'' \mid (h, i, t'') \in E^{cy}_{jhi}(t_m^-)\}$

1. Given the time of the next event, determine the unit initiating the event with probability $p = \exp(\hat{\theta}' s(t, i)) / \sum \mathcal{R}_i(t) \exp(\hat{\theta}' s(t, i))$ based on the ground intensity function in equation (2) and with $\hat{\theta}$ the estimate for θ. The set $R_i(t)$ is the set of units at risk of initiating an event at time t.
2. Select the receiver of the event sampling from the set $\mathcal{R}_{ij}(t) \setminus i$ using the multinomial probability $p_i(j, H_{t^-})$ in equation (3) obtained by substituting β' with its estimate $\hat{\beta}'$.

These two steps are repeated until one event is generated for each observed event time.

The distribution of the auxiliary statistics is computed from the observed data and the simulated sequences of events to assess the model fit. The deciles of the empirically observed distributions of the statistics are then compared.

Following established network modelling practice (Hunter et al., 2008), we adopt the approach of Lospinoso and Snijders (2019) and evaluate the fit graphically using violin plots and by means of the Mahalanobis distance-based Monte Carlo GOF test.

For each auxiliary statistic, the empirical distribution is represented using violin plots, one for each decile (Hintze & Nelson, 1998). If the model fits the data well, the observed deciles should not be extreme in the distribution of the auxiliary statistics generated under the estimated model. By 'not extreme', we mean that the observed 10-quantiles should lie between the 5% and 95% percentiles of the simulated values. In the following, we refer to this range as the 90% confidence interval. The GOF test uses the Mahalanobis distance to measure the divergence of the simulated deciles and those observed as the test statistic. Large values of the Mahalanobis distance indicate the extent to which simulated and observed deciles differ. A large value would suggest poor model fit. We refer interested readers to the paper of Lospinoso and Snijders (2019) for the technical details.

The graphical procedure and the Mahalanobis distance-based Monte Carlo GOF test require that the sequence of events can be generated from the estimated model.

4 Empirical illustration

4.1 Setting

The quality of health care services depends crucially on their timing. Because 'The consequences of adverse events cannot always be offset by working harder on subsequent days' (Redelmeier & Bell, 2007, p.1164), outcomes of treatment are highly sensitive to the timing of treatment itself (Gupta et al., 2012). The same treatment administered at a different time point may be associated with widely different patient outcomes (Powers, 2020).

If the quality of care depends on the timing of treatment (Seymour et al., 2017), and if the timing of treatment depends on the quality of coordination (Hoffer Gittell, 2002), then it should be obvious why the time structure of coordination mechanisms between health care organizations may be of paramount importance in the case of interhospital patient transfer (Amati et al., 2019; Berta et al., 2022; Zachrison et al., 2022).

Interorganizational collaboration and coordination among healthcare organizations is a major area of empirical application for network models (Iwashyna et al., 2009; Landon et al., 2012; Lomi & Pallotti, 2012; Mascia et al., 2017).

Network event models have provided valuable insights into the relational processes of collaborative patient transfer and sharing relations. Interhospital patient transfer relations generate data with the canonical relational event structure: a sender hospital (i) transfers a patient at time t to a receiver hospital (j) in an attempt to provide a collaborative solution to a difficult clinical case. When multiple hospitals do the same, they give rise to a network of patient transfer events—a common feature of regional systems of healthcare (Lee et al., 2011). Questions asked in empirical studies typically concern the role of geographical distance (Mascia et al., 2017), ownership (Zachrison et al., 2022), competition (Lomi & Pallotti, 2012), and the effect quality, size, and cost differentials on interhospital patient transfer rates (Kellermann & Ackerman, 1988; Lomi et al., 2014).

The importance of relational coordination among health care organizations to ensure continuity of care is well recognized both in clinical practice (O'Malley & Cunningham, 2009), as well as health care management research (Hoffer Gittell, 2002). In turn, this general recognition has attracted increasing interest in the antecedents and consequences of interorganizational relations within the field of healthcare (Bolton et al., 2021).

4.2 Data

The data consist of the complete set of 3,778 interhospital patient transfer events observed between all the 35 hospitals (public and private) located in the Italian region Abruzzo observed over a four-year period from 01.01.2005 to 31.12.2008. The sample contains only elective transfer events. Interhospital transfer events involving critically ill patients are excluded. The data were provided by the regional Agency of Public Health, which is responsible for collecting and managing discharge data to assess the performance of regional hospitals.

Additional information was collected over and above the time series of interhospital referral events. The administrative local units [henceforth local health unit (LHU)] to which a hospital belongs facilitates the occurrence of events because the hospitals located in the same LHU refer to the same management; the geographical distance between hospitals in kilometres is a proxy for transportation costs and risks of transferring a patient; the number of beds and the mean occupancy rate describe the capacity of the hospitals and the use of available capacity. All the variables, but the distance among hospitals, are monadic and time-varying due to administrative changes.

Table 2 provides a summary of the covariates included in the model we estimate in the empirical part of the paper. A detailed description of the event data and the hospital-specific covariates can be found in Amati et al. (2019, 2021). The set of cases needed for estimation (Lerner & Lomi, 2020b) includes the complete set of 128,452 non-events—i.e. all the possible interhospital patient transfer events that could have happened during the observation period but did not.

4.3 Model specification and estimation

To illustrate the GOF procedure, we specify two models with an identical ground intensity function but a different mark distribution. The first model includes only basic dyadic and exogenous statistics. The second model accounts for more complex (triadic) dependencies. In the empirical part of the paper, we report estimates of a model based, in part, on prior empirical studies (Vu et al., 2017). The specification we adopt is kept at a relatively low level of complexity because we want to focus attention on the main analytical purpose of the illustration.

The ground intensity function in (2) is specified using the out-degree, out-intensity, and last sending statistics. The out-degree controls for the heterogeneous tendency of hospitals to initiate patient transfer events based on the number of their hospital partners. Being the average recency per partner, the out-intensity statistic accounts for both the number of hospital partners and the

Table 2. Description of the covariates

Variable	Description	Type	Range	Mean	s.d.
Local health unit (LHU)	Membership of single hospitals to the different administrative units in which the region is partitioned	Nominal	1–6	–	–
Distance (geo.dist)	Geographical distance (kilometres)	Continuous	2–146	69.0	28.8
Hospital size (n.beds)	Total number of staffed beds	Count	[20–661]	155.4	138.6
Occupancy rate (occ.rate)	Proportion of beds occupied	Continuous	[5–217][a]	74.5	23.9

Note. [a]It is not uncommon for some hospital to operate above installed capacity, particularly during periods in which cost-cutting measures, downsizing strategies, and restructuring plans are being implemented.

recency of the events initiated by i. The last sending accounts for the tendency of hospitals to initiate a transfer event when an event was recently initiated in the past.

For the mark distribution in (3), we included statistics that reflect standard micro-mechanisms that explain network dynamics. We start by describing the endogenous network mechanisms. The first mechanism is preferential attachment reflecting the tendency of hospitals to transfer patients to 'popular' receiver hospitals—hospitals chosen by many others as partners. Preferential attachment is represented by the in-degree and the in-intensity statistics. While the in-degree statistic counts the number of hospitals transferring patients to the receiving hospital, the in-intensity accounts for the average number of received transfer events for sending hospital.

The second and third mechanisms are dyadic. The last receiving accounts for the tendency of hospitals to receive a transfer event when an event was recently received in the past. Repetition represents relational inertia—the tendency of hospitals to transfer patients to the same hospitals selected as partners in the past. Reciprocation describes the preferential tendency of hospitals to select reciprocating partner hospitals as receivers.

The last two mechanisms relate to triadic closure—the tendency of hospitals sharing one or more alters to directly connect. Triadic closure may be the outcome of transitivity or tendencies towards cyclic closure (Lerner & Lomi, 2017). Transitive closure occurs when a two-path $i \to h \to j$ is closed by the event $i \to j$, while cyclic closure concerns the two-path $j \to h \to i$.

Exogenous (control) covariates include: (i) joint membership in the same LHU to control for the preferential tendency of hospitals to refer patients to other hospitals within the same administrative area; (ii) capacity (number of beds) to control for the hospital size; and (iii) availability (occupancy rate) to control for the differential attractiveness of larger hospitals and hospitals with capacity available, respectively. Finally, we include geographical distance to control for the tendency of hospitals to transfer patients to nearby hospitals.

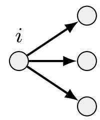
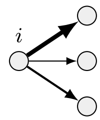
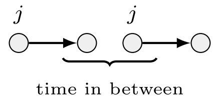
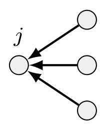
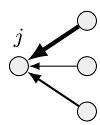
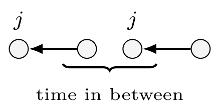
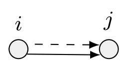
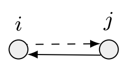
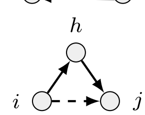
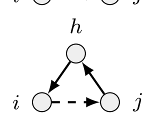
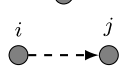
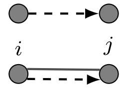
We refer readers to Table 3 for the mathematical definitions and graphical representations of the relevant statistics. For the analysis, we use the specification of the decay function f based on the power-law decay.

Table 4 reports the model estimates. All parameters, with the exception of those related to the occupancy rate, are significant at conventional levels. These results suggest that the occupancy rate of receiver hospitals does not affect interhospital patient flows significantly.

For the ground intensity function, the estimates of the out-degree and out-intensity parameters indicate that hospitals have a diverse propensity to initiate transfer events. The positive values indicate that hospitals initiating a higher number of patient transfer events are more likely to initiate new patient transfer events in the future. The negative value of the last sending parameter suggests that hospitals are more likely to initiate a new referral event if they recently initiated a past event.

For the mark distribution, the positive parameters associated with the in-degree and in-intensity statistics suggest that hospitals receiving events initiated by many other hospitals or receiving many transfer events are more likely to be the receiver of the next event. The estimates provide clear evidence of preferential attachment. The negative value of the parameter related to the recent receiving statistic indicates that referral events are more likely to flow to recently receiving hospitals. We also observe that there is evidence of repetition and reciprocation as indicated by the

Table 3. Statistics for the conditional intensity function

Statistic	Representation	Formula
Out-degree		$\sum_{j\neq i} \mathbb{I}_{\{N_{ij}(t^-)>0\}}$
Out-intensity		$\frac{\sum_{j\neq i}\sum_{e=1}^{N_{ij}(t^-)} f(t_e)}{\sum_{j\neq i}\mathbb{I}_{\{N_{ij}(t^-)>0\}}}$
Recent sending		$t - \max_{e\in E_i} t_e$
In-degree		$\sum_{k\neq j}\mathbb{I}_{\{N_{kj}(t^-)>0\}}$
In-intensity		$\frac{\sum_{k\neq j}\sum_{e=1}^{N_{kj}(t^-)} f(t, T_{kj}^e, \alpha)}{\sum_{k\neq j}\mathbb{I}_{\{N_{kj}(t^-)>0\}}}$
Recent receiving		$t - \max_{e\in E_i} t_e$
Repetition		$\sum_{e=1}^{N_{ij}(t^-)} f(t_e)$
Reciprocation		$\sum_{e=1}^{N_{ij}(t^-)} f(t, T_{ji}^e, \alpha)$
Transitive closure		$\sum_{h\neq i,j} g(\sum_{e=1}^{N_{ih}(t^-)} f(t_e),\ \sum_{e=1}^{N_{hj}(t^-)} f(t_e))$
Cyclic closure		$\sum_{h\neq i,j} g(\sum_{e=1}^{N_{hi}(t^-)} f(t_e),\ \sum_{e=1}^{N_{jh}(t^-)} f(t_e))$
Monadic attribute		$v_j(t)$
Matching		$\mathbb{I}_{\{v_i(t)=v_j(t)\}}$
Dyadic attribute		$z_{ij}(t)$

Note. In the formulas, $N_{ij}(t^-)$ is the number of relational events from unit i to unit j up to, but not including time t; $f(t_e)$ is the decay function accounting for the temporal relevance of previous events t_e; v and z are monadic and dyadic covariates. The symbol $\mathbb{I}$ denotes the indicator function taking a value of 1 if the condition between brackets at the subscript is true; otherwise it is 0. The function $g(\cdot)$ denotes the geometric mean of the quantities between brackets.

corresponding positive parameters. Thus, patient transfer events to previous partners and reciprocating past transfer events are more likely to occur in the future.

The estimates for the closure statistics indicate evidence against transitive closure and evidence of cyclic closure. Jointly interpreted, the two coefficients suggest the absence of a hierarchical ordering in the interhospital network. Finally, the parameters of the exogenous statistics indicate

Table 4. Model estimation for the patient transfer events observed over the period 01.01.2005–31.12.2008

	Model 1			Model 2		
	Est.	s.e.	Sig.	Est.	s.e.	Sig.
Ground intensity function						
Out-degree	0.613	0.021	***	0.613	0.021	***
Out-intensity	0.421	0.012	***	0.421	0.012	***
Last sending	−1.233	0.133	***	−1.233	0.133	***
Mark distribution						
In-degree	0.087	0.043	*	0.087	0.046	
In-intensity	0.061	0.025	*	0.098	0.0003	***
Last receiving	−3.029	0.231	***	−2.930	0.053	***
Repetition	0.247	0.011	***	0.249	0.012	***
Reciprocation	0.122	0.017	***	0.112	0.015	***
Transitive closure				−0.069	0.024	**
Cyclic closure				0.166	0.035	***
Local health unit (LHU) match	2.118	0.765	***	0.749	0.0.024	***
Number of beds	1.273	0.060	***	1.298	0.061	***
Occupancy rate	0.054	0.046		0.073	0.047	
Geographical distance	−0.731	0.037	***	−0.727	0.038	***

Note. *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001.

that transfer events between hospitals affiliated with the same LHU, directed to hospitals with a high capacity and located nearby, are more likely. These empirical results are generally consistent with those reported by extant research on interhospital patient transfer (Lomi et al., 2014; Stadtfeld et al., 2016; Vu et al., 2017; Zachrison et al., 2022).

4.4 Assessing model fit

We evaluated the fit of Model 1 and Model 2 in Table 4 using the procedure described in Section 3. In the current work, we chose to illustrate the method using the time elapsed since the first occurrence of the conditioning event because the number of repeated events before observing reciprocation or closure is small (mean 5 and median 4 for transitive closure and mean 4, median 3 for reciprocity). For each estimated model, we simulated sequences of relational events following the approach described in Section 3.4. We calculated the auxiliary statistics for the observed and generated sequence of events and computed the time distribution to reciprocity, transitive closure, and cyclic closure statistics.

Figure 1 reports the observed density distribution of the auxiliary statistics (solid line) and that obtained by pooling the simulated distributions for Model 1 (dotted line) and Model 2 (dashed line), respectively. The support of the density distributions has been limited to the third quantile of the observed distribution for visualization purposes. The same plots for the entire support and the descriptive statistics of the distributions are reported in Section S2, online supplementary material. Comparison between the models indicates that the distributions generated from Model 1 are characterized by larger deviations from the observed ones due to higher variation and heavier right tails. On the contrary, the distributions from Model 2 are closer to the observed ones.

We use the deciles (or the 10-quantiles) and the Mahalanobis distance-based Monte Carlo GOF to compare the simulated and empirical distributions. Figures 2 and 3 report the violin plots of the deciles of the distributions of the auxiliary statistics for 1,000 simulated sequences of transfer events from Model 1 and Model 2, respectively. In each graph, the x-axis reports the order of the decile, while the y-axis is the standardized value of the quantile. The values are standardized

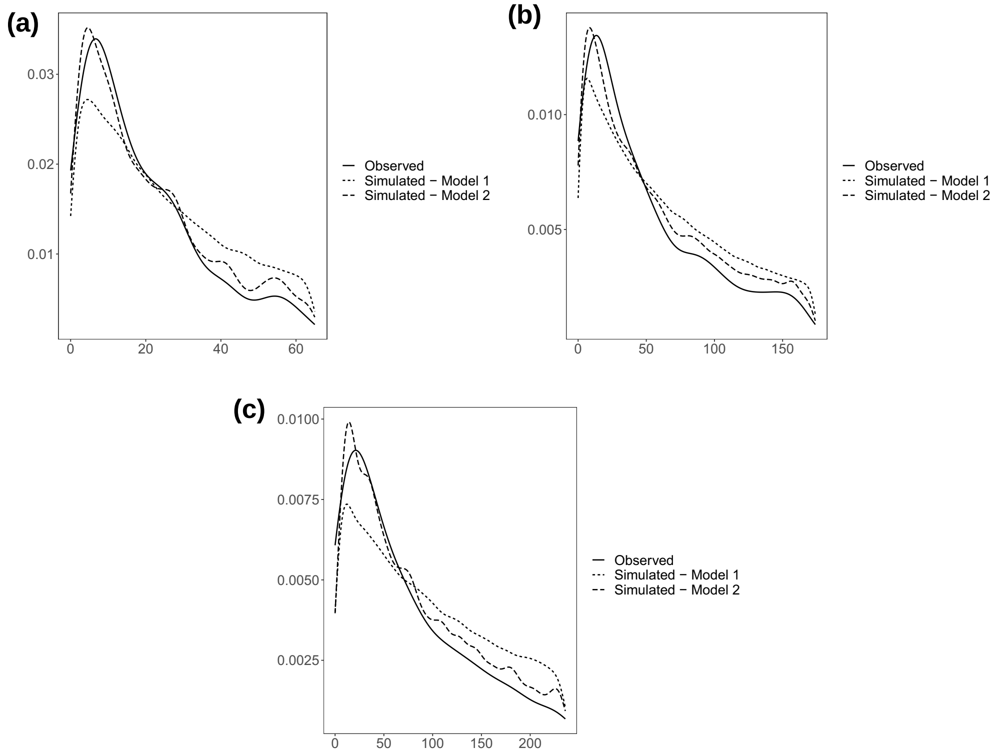


Figure 1. Density distribution of the internal time for reciprocation (a), transitive (b), and cyclic closure (c) for the patient transfer data and the simulated sequence of events from Model 1 and Model 2. The support of the distributions is limited to the third quartile for visualization purposes.

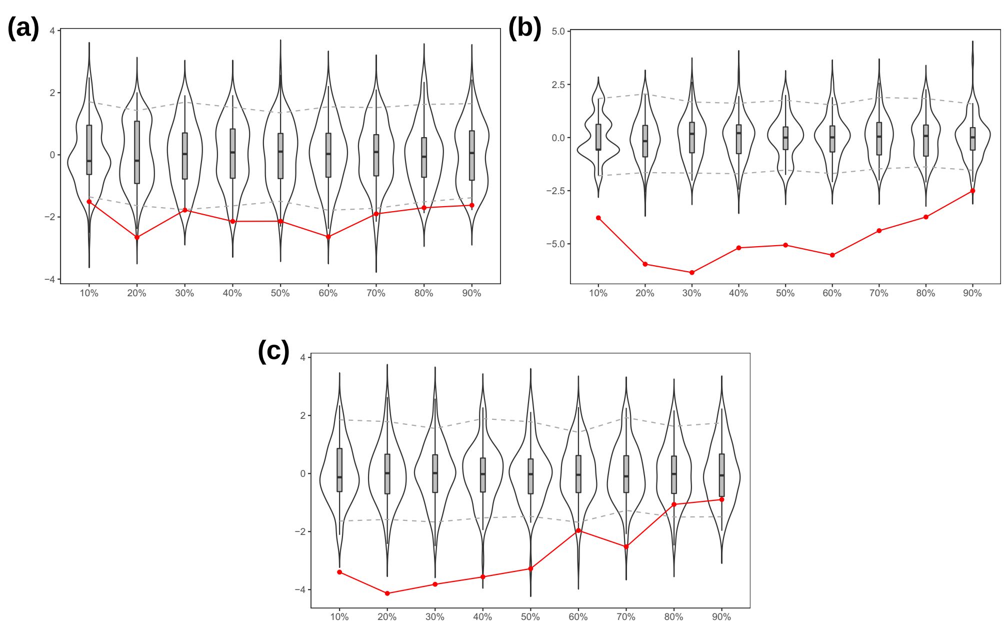


Figure 2. Goodness of fit (GOF) for Model 1 including only dyadic statistics. The violin plots refer to the internal time for reciprocation (a), transitive (b), and cyclic closure (c). Dashed lines are the 90% confidence interval for the deciles. The solid line and dots denote the observed deciles. The p-values of the Mahalanobis distance-based Monte Carlo GOF test are lower than 0.01 for all the auxiliary statistics.

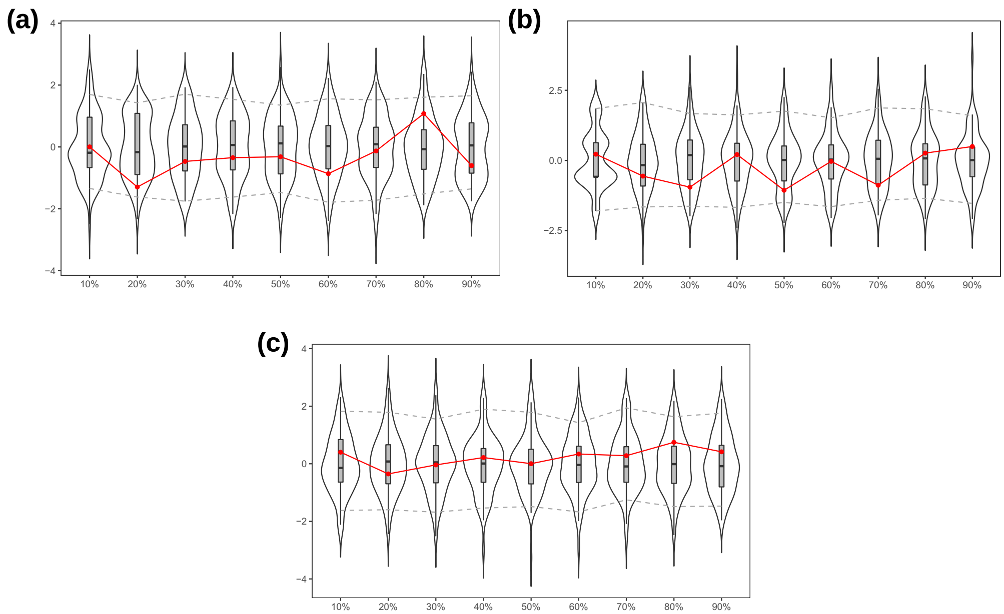


Figure 3. Goodness of fit (GOF) for Model 2 including the statistics for transitive and cyclic closure. The violin plots refer to the internal time for reciprocation (a), transitive (b), and cyclic closure (c). Dashed lines are the 90% confidence interval for the deciles. The solid line and dots denote the observed deciles. The p-values of the Mahalanobis distance-based Monte Carlo GOF test are 0.124 for reciprocation, 0.276 for transitive closure, 0.194 for cyclic closure.

using the mean and standard deviations of the simulated deciles so that the violin plots are on a common scale. Each violin plot depicts the box plots and the distribution (line around the box plot) of the deciles for the simulated sequence of events, the corresponding observed count (dots), and the 90% confidence intervals (dashed line). The model fits the data well when the solid line falls within the dashed lines.

Figure 2 indicates that Model 1 is not far from reproducing the distribution of the time to reciprocation, but dramatically fails to capture the time to transitive and cyclic closure. While the observed values of the deciles for the time to reciprocation are quite close to the limits of the 90% confidence interval, those for the transitive and cyclic closure are far off. This indicates that Model 1 does not have a good fit because it produces estimates that are inconsistent with the internal time distribution of the data generating mechanisms. As a confirmation of this, the Mahalanobis distance-based Monte Carlo GOF returns p-values lower than 0.01 for all the auxiliary statistics, suggesting that the simulated time distributions do not resemble the observed ones. In particular, the simulated distributions have larger deciles than those observed, thereby implying that the simulated distributions have a heavier tail than the observed distribution as depicted in Figure 1.

When the statistics for cyclic closure and transitive closure are included in the model (Model 2), the distributions of the time to reciprocation, transitive, and cyclic closure are well reproduced (Figure 3). The observed deciles lie in the 90% confidence interval, and the average of the simulated deciles is close to the corresponding empirical values. This suggests that Model 2 has a good fit for the selected auxiliary variables. The Mahalanobis distance-based Monte Carlo GOF tests return p-values larger than 0.05 for all the auxiliary statistics, supporting the statement that the simulated time distributions resemble the observed one.

We conclude this section by observing that the specification of Models 1 and 2 relates to the mark distribution rather than the ground intensity function that governs the timing of the effects. This suggests that out-degree and out-intensity effects provide sufficient information to explain the initiation of transfer events. At the same time, more than dyadic statistics are needed to explain the choice of receiving partner hospital.

5 Discussion and conclusions

When does a REM provide an adequate representation of data produced by continuously observed social interaction? Addressing this question provided the main motivation for this study.

Clearly, many answers are possible, depending, among other things, on the aspects of the model and the data that are considered most empirically interesting or theoretically relevant in any particular circumstance. We focused on developing a general procedure to evaluate the GOF for REMs around what is typically considered their most distinctive feature, namely their ability to incorporate information on the timing of the observations.

The auxiliary variable approach to the GOF of network models is now both well established (Hunter et al., 2008) as well as broadly adopted (Lospinoso & Snijders, 2019). Yet, we are not aware of studies that have adapted this approach to examine the ability of REMs to reproduce the internal time structure of the mechanisms that are assumed to generate the observations.

The main motivation for developing this perspective is grounded in the factual observation that effects in REMs have an internal time extension determined by the sequential ordering of the events that are actually observed (Bianchi et al., 2022).

The approach to the GOF for REMs proposed in this paper takes the distinctiveness of REMs into account and distinguishes itself from likelihood-based methods (AIC, BIC) and prediction-based methods generally adopted in extant empirical research.

Methods in the former class (likelihood-based) typically involve model comparison with the objective of establishing the 'best' model relative to the other, and are most suited for model selection rather than global GOF. Unlike likelihood-based methods, the objective of the approach based on auxiliary statistics is not to provide a comparative evaluation of models within a given class of models, but to assess the GOF of the model by comparing its implications directly with the data. Methods in the latter class (prediction-based) evaluate the fit using the model's predictive accuracy and address questions about how well the model performs at predicting events. Unlike methods based on prediction accuracy, according to the auxiliary statistics approach that we have proposed, a model has a good fit if it can reproduce the internal time structure of the relational mechanisms it postulates. For this reason, we think that our methodological proposal may be better grounded in a more theoretical understanding of social networks and social relations as processes situated in time and space (Abbott, 2001; Bearman et al., 1999; Gibson, 2005; Kitts et al., 2017; Mische & White, 1998; Moody, 2002).

We believe that representing with accuracy the internal time structure of relational processes is at least as important as predicting with accuracy the next event (Brandenberger, 2019) or the number of events that happened relative to those that could have happened but did not (Butts et al., 2007; Stadtfeld & Block, 2017). According to the argument we have developed in this paper, a model inconsistent with the internal timing structure of the relational mechanisms that it postulates cannot be considered a satisfactory model for relational event data.

We have tested the practical value of our methodological argument in the context of data on collaborative care among hospitals—an empirical setting where the timing of events is generally considered of particular relevance. We have shown that the ability of the model to reproduce the internal time structure of the effects of theoretical interest is sensitive to changes in empirical model specification. In models strictly defined in terms of dyadic dependence, parameter estimates do not produce a consistent representation of the timing of observed event sequences. Full models incorporating basic patterns of extra-dyadic dependence present in the data can reproduce with significantly higher fidelity the internal time structure of the network mechanisms assumed to have generated the observed data.

The identification of the auxiliary variables was driven by a combination of empirical (sample-specific) and theoretical (sample-independent) considerations. We focused on mechanisms of transitive and cyclic closure because prior studies instruct us that these forms of extra-dyadic dependence are present in the specific empirical setting we have examined (Amati et al., 2019; Lomi et al., 2014; Zachrison et al., 2022) and in interorganizational networks more generally (Lomi & Pattison, 2006).

As we expect to be typical in empirical studies, the selection of the auxiliary variables in the example we have presented comes from a combination of knowledge about extant empirical results (Kitts et al., 2017) and general indications coming from theories of interorganizational relations

(Powell et al., 2005). Forms of triadic closure play an important role in the analysis of network data more generally (Robins et al., 2009; Snijders et al., 2006). For this reason, we expect that our proposed analytical strategy will apply beyond the boundaries of our illustrative case study. Additional auxiliary statistics, such as, for example, the frequency of repetition and the number of non-reciprocated events may be considered important in specific applications, and nothing prevents their future development and implementation.

From a practical point of view, however, the choice of the auxiliary variables is necessarily context-dependent, at least to some extent. Different auxiliary variables may be appropriate in different empirical settings. For example, in the analysis of e-mail data where some events are simultaneous (Perry & Wolfe, 2013), the ability of the model to capture time to message response (or 'reciprocity') may be more relevant than the ability of the model to capture the internal time structure of more complex closure mechanisms. On the other hand, the timing of closure mechanisms may be crucial in studies of interpersonal interaction within task-oriented teams (Leenders et al., 2016). Triadic configurations associated with structural balance are essential to understanding polarization and coordination in teams and hence deserve particular attention in that context (Lerner & Lomi, 2020a). Cyclic closure has been found to play an important stabilizing role in financial markets (Lomi & Bianchi, 2024). Given the relative recency of REMs, we do not yet have sufficient experience that may be distilled into a small set of rules of thumb or best practices to guide empirical research. As empirical experience accumulates, we believe that it will become possible to establish simple experience-based rules of thumb, or decision heuristics, to guide the empirical assessment of the GOF of REMs fitted to data produced by continuously observed social interaction processes.

Our work involves a number of explicit assumptions. For example, one event contributes only once to the computation of the relevant statistics in our models. While it is clearly possible to experiment with alternative assumptions, this assumption is not unreasonable as it avoids the need to develop additional assumptions on time dependence in—and stability of network effects. The way we choose to compute the inter-event time is non-unique in the sense that alternatives are conceivable. Consider reciprocity, for example. An obvious alternative to what we have proposed could have been to select the last $i \to j$ event before $j \to i$ and consider all the previous events as censored reciprocation spells.

Deserving mention in this closing section are the potential problems inherent in censoring, which—while of limited import in our empirical case—are generally difficult to rule out completely in observational studies. Under conditions of non-random censoring, the problem is that the duration of the spells included in the sample (inter-event time in our case) and those in the population may differ significantly. While censoring patterns do not affect the partial likelihood function directly (Efron, 1977), censoring may lead to a loss of efficiency in the estimates of the parameters in the Cox model (Qin & Shen, 2010). We think that this potential problem is unlikely to invalidate the conclusions of the study because we have no reason to believe that the length of the spells fully observed within the sample differs significantly from the spells initiated during the weeks immediately before the start of the sample observation period, or concluded in the weeks immediately following its end.

We introduced the auxiliary statistics only for sequences of exact time-stamped events. However, the approach we proposed may be adopted also when only the time-order information is available by defining auxiliary statistics that account for the order of events. For instance, for reciprocation, we propose to use the number of events originating from j since the first event from i to j instead of the time duration. Appendix S3, online supplementary material, provides examples of auxiliary statistics for ordered events that may be specified when the structure of the data requires it. Similarly, the auxiliary statistics are computed at the network level since they are aggregated over all the actors. Building on Stadtfeld and Block (2017), analogous auxiliary statistics can be defined at the actor level to reveal actor-level heterogeneity, and the presence of specific actors exhibiting distinctive patterns of relational behaviour. This strategy would be particularly valuable for assessing the GOF of REMS specified according to general principles of actor-oriented modelling (Snijders, 2017; Stadtfeld et al., 2017). It might also be used to identify outliers, i.e. actors behaving significantly differently from others.

Our work suggests that REMs may assist in addressing explicitly issues of timing that are implicit in—but not absent from other network modelling frameworks. The internal time extension of network effects that we have discussed makes the interpretation of a single parameter associated with a process with a time-dependent intensity hard to decipher. Consider reciprocity, for example. Qualitative evidence is available that the time distribution of symmetrizing events is highly skewed with a very long tail (Perry & Wolfe, 2013): the majority of reciprocated events happen quickly, but some may happen after considerable delay. When this is the case, it is unclear how and when reciprocation will affect the probability of observing the next event in a sequence. The additional complexity inherent in the possibility that different relational mechanisms may be regulated by different internal time clocks further limits our ability to provide a meaningful quantitative interpretation for the estimates of REMs.

A promising solution to this problem has been recently proposed by Juozaitienė and Wit (2022), who implement a semi-parametric specification of reciprocity and triadic effects, which are then modelled directly as time-varying splines. This representation affords a direct analysis of the temporal structure of the network effects of interest.

The approach we proposed in this paper is applicable, in principle, to the evaluation of the fit of any form of REM, provided that it is possible to simulate event sequences from the model that are sufficiently proximate to the observed series. Related to this last point, and to conclude, one additional set of issues deserving mention concerns the difficulties of simulating event sequences based on the semi-parametric proportional hazard specification commonly used in empirical research. The simulation-based approach we have proposed is likely to be considerably simplified by adopting parametric models involving explicit and testable assumptions about the form of time dependence shaping the observed flow of relational events. We think that the development of parametric models will be an important part of the next step necessary to make REMs more powerful, flexible, and generally applicable to data with complex temporal dependencies produced by continuously observed social interaction.

Acknowledgments

We thank the members of the Social Networks Lab at ETH Zürich and members of the Institute of Computing at the Università of della Svizzera Italiana (USI), Lugano, for constructive feedback and comments on earlier drafts. Professors Christoph Stadtfeld (ETHZ) and Ernst Wit (USI), in particular, provided valuable criticism and advice throughout the development of this work. Our gratitude extends to Dr. Duy Vu for making available a working version of the Java code that we adopted to estimate the parameters of the models discussed in this article. Finally, we would like to thank the USI Library Services for their assistance to making our work available as an open access publication.

Conflict of interest: None of author has a conflict of interest to disclose.

Funding

Schweizerischer Nationalfonds (SNF) (Swiss National Science Foundation) number CR12I1_156229.

Data availability

The data underlying this article cannot be shared publicly due to the sensitive nature of the information they contain. The R code implementing the auxiliary statistics is available from the corresponding author upon request.

Supplementary material

Supplementary material is available online at *Journal of the Royal Statistical Society: Series A*.

References

Aalen O. O., Borgan O., & Gjessing H. K. (2008). *Survival and event history analysis*. Springer.

Abbott A. (2001). *Time matters: On theory and method*. University of Chicago Press.

Amati V., Lomi A., & Mascia D. (2019). Some days are better than others: Examining time-specific variation in the structuring of interorganizational relations. *Social Networks*, *57*, 18–33. https://doi.org/10.1016/j.socnet.2018.10.001

Amati V., Lomi A., Mascia D., & Pallotti F. (2021). The co-evolution of organizational and network structure: The role of multilevel mixing and closure mechanisms. *Organizational Research Methods*, *24*(2), 285–318. https://doi.org/10.1177/1094428119857469

Atouba Y., & Shumate M. (2010). Interorganizational networking patterns among development organizations. *Journal of Communication*, *60*(2), 293–317. https://doi.org/10.1111/jcom.2010.60.issue-2

Bearman P., Faris R., & Moody J. (1999). Blocking the future: New solutions for old problems in historical social science. *Social Science History*, *23*(4), 501–533. https://doi.org/10.1017/S0145553200021854

Berta P., Vinciotti V., & Moscone F. (2022). The association between hospital cooperation and the quality of healthcare. *Regional Studies*, *56*(11), 1858–1873. https://doi.org/10.1080/00343404.2021.2009792

Bianchi F., Filippi-Mazzola E., Lomi A., & Witt E. (2024). Relational event modeling. *Annual Review of Statistics and Its Application*, *11*. https://doi.org/10.1146/annurev-statistics-040722-060248

Bianchi F., & Lomi A. (2022). From ties to events in the analysis of interorganizational exchange relations. *Organizational Research Methods*, *26*(3), 524–565. https://doi.org/10.1177/10944281211058469

Bianchi F., Stivala A., & Lomi A. (2022). Multiple clocks in network evolution. *Methodological Innovations*, *15*(1), 29–41. https://doi.org/10.1177/20597991221077877

Bianconi G., Darst R. K., Iacovacci J., & Fortunato S. (2014). Triadic closure as a basic generating mechanism of communities in complex networks. *Physical Review E*, *90*(4), 042806. https://doi.org/10.1103/PhysRevE.90.042806

Block P., Koskinen J., Hollway J., Steglich C., & Stadtfeld C. (2018). Change we can believe in: Comparing longitudinal network models on consistency, interpretability and predictive power. *Social Networks*, *52*, 180–191. https://doi.org/10.1016/j.socnet.2017.08.001

Blossfeld H.-P., Hamerle A., & Mayer K. U. (2014). *Event history analysis: Statistical theory and application in the social sciences*. Psychology Press.

Bolton R., Logan C., & Gittell J. H. (2021). Revisiting relational coordination: A systematic review. *The Journal of Applied Behavioral Science*, *57*(3), 290–322. https://doi.org/10.1177/0021886321991597

Brandenberger L. (2019). Predicting network events to assess goodness of fit of relational event models. *Political Analysis*, *27*(4), 556–571. https://doi.org/10.1017/pan.2019.10

Brandes U., Lerner J., & Snijders T. A. B. (2009). Networks evolving step by step: Statistical analysis of dyadic event data. In *Proceedings of the 2009 international conference on advances in social network analysis and mining (ASONAM'09)* (pp. 200–205). IEEE Computer Society Press.

Bright D., Sadewo G. R. P., Lerner J., Cubitt T., Dowling C., & Morgan A. (2023). Investigating the dynamics of outlaw motorcycle gang co-offending networks: The utility of relational hyper event models. *Journal of Quantitative Criminology*, 1–43. https://doi.org/10.1007/s10940-023-09576-x

Butts C. T. (2008). A relational event framework for social action. *Sociological Methodology*, *38*(1), 155–200. https://doi.org/10.1111/j.1467-9531.2008.00203.x

Butts C. T. (2009). Revisiting the foundations of network analysis. *Science*, *325*(5939), 414–416. https://doi.org/10.1126/science.1171022

Butts C. T. (2017). Comment: Actor orientation and relational event models. *Sociological Methodology*, *47*(1), 47–56. https://doi.org/10.1177/0081175017728929

Butts C. T., Lomi A., Snijders T. A., & Stadtfeld C. (2023). Relational event models in network science. *Network Science*, *11*(2), 175–183. https://doi.org/10.1017/nws.2023.9

Butts C. T., & Marcum C. S. (2017). A relational event approach to modeling behavioral dynamics. In A. Pilny & M. S. Poole (Eds.), *Group processes: Data-driven computational approaches* (pp. 51–92). Springer International Publishing.

Butts C. T., Petrescu-Prahova M., & Remy Cross B. (2007). Responder communication networks in the world trade center disaster: Implications for modeling of communication within emergency settings. *The Journal of Mathematical Sociology*, *31*(2), 121–147. https://doi.org/10.1080/00222500601188056

Cox D. R., & Isham V. (1980). *Point processes*. (Vol. 12), CRC Press.

Dakin R., & Ryder T. B. (2020). Reciprocity and behavioral heterogeneity govern the stability of social networks. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(6), 2993–2999. https://doi.org/10.1073/pnas.1913284117

Efron B. (1977). The efficiency of Cox's likelihood function for censored data. *Journal of the American Statistical Association*, *72*(359), 557–565. https://doi.org/10.1080/01621459.1977.10480613

Gibson D. R. (2005). Taking turns and talking ties: Networks and conversational interaction. *American Journal of Sociology*, *110*(6), 1561–1597. https://doi.org/10.1086/428689

Gittell J. H., Fairfield K. M., Bierbaum B., Head W., Jackson R., Kelly M., Laskin R., Lipson S., Siliski J., Thornhill T., & Zuckerman J. (2000). Impact of relational coordination on quality of care, postoperative pain and functioning, and length of stay: A nine-hospital study of surgical patients. *Medical Care*, *38*(8):807–819. https://doi.org/10.1097/00005650-200008000-00005

Grambsch P. M., & Therneau T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, *81*(3), 515–526. https://doi.org/10.1093/biomet/81.3.515

Gupta S., Bonilla M., Valverde P., Fu L., Howard S. C., Ribeiro R. C., & Sung L. (2012). Treatment-related mortality in children with acute myeloid leukaemia in Central America: Incidence, timing and predictors. *European Journal of Cancer*, *48*(9), 1363–1369. https://doi.org/10.1016/j.ejca.2011.10.009

Hintze J. L., & Nelson R. D. (1998). Violin plots: A box plot-density trace synergism. *The American Statistician*, *52*(2), 181–184. https://doi.org/10.1080/00031305.1998.10480559

Hoffer Gittell J. (2002). Coordinating mechanisms in care provider groups: Relational coordination as a mediator and input uncertainty as a moderator of performance effects. *Management Science*, *48*(11), 1408–1426. https://doi.org/10.1287/mnsc.48.11.1408.268

Hunter D. R., Goodreau S. M., & Handcock M. S. (2008). Goodness of fit of social network models. *Journal of the American Statistical Association*, *103*(481), 248–258. https://doi.org/10.1198/016214507000000446

Iwashyna T. J., Christie J. D., Moody J., Kahn J. M., & Asch D. A. (2009). The structure of critical care transfer networks. *Medical Care*, *47*(7), 787–793. https://doi.org/10.1097/MLR.0b013e318197b1f5

Jacobsen M. (2006). *Point process theory and applications: Marked point and piecewise deterministic processes*. Springer Science & Business Media.

Juozaitienė R., & Wit E. C. (2022). Non-parametric estimation of reciprocity and triadic effects in relational event networks. *Social Networks*, *68*, 296–305. https://doi.org/10.1016/j.socnet.2021.08.004

Kellermann A. L., & Ackerman T. F. (1988). Interhospital patient transfer. *The New England Journal of Medicine*, *319*(10), 643–647. https://doi.org/10.1056/NEJM198809083191010

Kings M., Arbon J. J., McIvor G. E., Whitaker M., Radford A. N., Lerner J., & Thornton A. (2023). Wild jackdaws can selectively adjust their social associations while preserving valuable long-term relationships. *Nature Communications*, *14*(1), 5103. https://doi.org/10.1038/s41467-023-40808-7

Kitts J. A., Lomi A., Mascia D., Pallotti F., & Quintane E. (2017). Investigating the temporal dynamics of interorganizational exchange: Patient transfers among Italian hospitals. *American Journal of Sociology*, *123*(3), 850–910. https://doi.org/10.1086/693704

Landon B. E., Keating N. L., Barnett M. L., Onnela J.-P., Paul S., O'Malley A. J., Keegan T., & Christakis N. A. (2012). Variation in patient-sharing networks of physicians across the United States. *Journal of the American Medical Association*, *308*(3), 265–273. https://doi.org/10.1001/jama.2012.7615

Laumann E. O., & Marsden P. V. (1982). Microstructural analysis in interorganizational systems. *Social Networks*, *4*(4), 329–348. https://doi.org/10.1016/0378-8733(82)90017-X

Lee B. Y., McGlone S. M., Song Y., Avery T. R., Eubank S., Chang C.-C., Bailey R. R., Wagener D. K., Burke D. S., Platt R., & Huang S. S. (2011). Social network analysis of patient sharing among hospitals in Orange County, California. *American Journal of Public Health*, *101*(4), 707–713. https://doi.org/10.2105/AJPH.2010.202754

Leenders R. T. A., Contractor N. S., & DeChurch L. A. (2016). Once upon a time: Understanding team processes as relational event networks. *Organizational Psychology Review*, *6*(1), 92–115. https://doi.org/10.1177/2041386615578312

Lerner J., & Lomi A. (2017). The third man: Hierarchy formation in wikipedia. *Applied Network Science*, *2*(1):1–30. https://doi.org/10.1007/s41109-017-0043-2

Lerner J., & Lomi A. (2020a). The free encyclopedia that anyone can dispute: An analysis of the micro-structural dynamics of positive and negative relations in the production of contentious wikipedia articles. *Social Networks*, *60*, 11–25. https://doi.org/10.1016/j.socnet.2018.12.003

Lerner J., & Lomi A. (2020b). Reliability of relational event model estimates under sampling: How to fit a relational event model to 360 million dyadic events. *Network Science*, *8*(1), 97–135. https://doi.org/10.1017/nws.2019.57

Lerner J., & Lomi A. (2022). A dynamic model for the mutual constitution of individuals and events. *Journal of Complex Networks*, *10*(2), cnac004. https://doi.org/10.1093/comnet/cnac004

Lerner J., & Lomi A. (2023). Relational hyperevent models for polyadic interaction networks. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *186*(3), 577–600. https://doi.org/10.1093/jrsssa/qnac012

Lomi A., & Bianchi F. (2024). A time to give and a time to receive: Role switching and generalized exchange in a financial market. *Social Networks*, *77*, 118–128. https://doi.org/10.1016/j.socnet.2021.11.005

Lomi A., Mascia D., Vu D. Q., Pallotti F., Conaldi G., & Iwashyna T. J. (2014). Quality of care and interhospital collaboration: A study of patient transfers in Italy. *Medical Care*, *52*(5), 407–414. https://doi.org/10.1097/MLR.0000000000000107

Lomi A., & Pallotti F. (2012). Relational collaboration among spatial multipoint competitors. *Social Networks*, *34*(1), 101–111. https://doi.org/10.1016/j.socnet.2010.10.005

Lomi A., & Pattison P. (2006). Manufacturing relations: An empirical study of the organization of production across multiple networks. *Organization Science*, *17*(3), 313–332. https://doi.org/10.1287/orsc.1060.0190

Lospinoso J., & Snijders T. A. (2019). Goodness of fit for stochastic actor-oriented models. *Methodological Innovations*, *12*(3), 205979911988428. https://doi.org/10.1177/2059799119884282

Mascia D., Pallotti F., & Angeli F. (2017). Don't stand so close to me: Competitive pressures, proximity and interorganizational collaboration. *Regional Studies*, *51*(9), 1348–1361. https://doi.org/10.1080/00343404.2016.1185517

Milo R., Shen-Orr S., Itzkovitz S., Kashtan N., Chklovskii D., & Alon U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, *298*(5594), 824–827. https://doi.org/10.1126/science.298.5594.824

Mische A., & White H. (1998). Between conversation and situation: Public switching dynamics across network domains. *Social Research*, *65*(3), 695–724. https://www.jstor.org/stable/40971267

Moody J. (2002). The importance of relationship timing for diffusion. *Social Forces*, *81*(1), 25–56. https://doi.org/10.1353/sof.2002.0056

Nallamothu B. K., Bates E. R., Herrin J., Wang Y., Bradley E. H., & Krumholz H. M. (2005). Times to treatment in transfer patients undergoing primary percutaneous coronary intervention in the United States: National registry of myocardial infarction (NRMI)-3/4 analysis. *Circulation*, *111*(6), 761–767. https://doi.org/10.1161/01.CIR.0000155258.44268.F8

Newman M. E., & Park J. (2003). Why social networks are different from other types of networks. *Physical Review E*, *68*(3), 036122. https://doi.org/10.1103/PhysRevE.68.036122

Newman M. E., Watts D. J., & Strogatz S. H. (2002). Random graph models of social networks. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(Suppl. 1), 2566–2572. https://doi.org/10.1073/pnas.012582999

Niezink N. M., & Campana P. (2023). When things turn sour: A network event study of organized crime violence. *Journal of Quantitative Criminology*, *39*(3), 655–678. https://doi.org/10.1007/s10940-022-09540-1

O'Malley A. S., & Cunningham P. J. (2009). Patient experiences with coordination of care: The benefit of continuity and primary care physician as referral source. *Journal of General Internal Medicine*, *24*(2), 170–177. https://doi.org/10.1007/s11606-008-0885-5

Pattison P., & Robins G. (2002). Neighborhood-based models for social networks. *Sociological Methodology*, *32*(1), 301–337. https://doi.org/10.1111/1467-9531.00119

Perry P. O., & Wolfe P. J. (2013). Point process modelling for directed interaction networks. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *75*(5), 821–849. https://doi.org/10.1111/rssb.12013

Powell W. W., White D. R., Koput K. W., & Owen-Smith J. (2005). Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. *American Journal of Sociology*, *110*(4), 1132–1205. https://doi.org/10.1086/421508

Powers W. J. (2020). Acute ischemic stroke. *New England Journal of Medicine*, *383*(3), 252–260. https://doi.org/10.1056/NEJMcp1917030

Qin J., & Shen Y. (2010). Statistical methods for analyzing right-censored length-biased data under Cox model. *Biometrics*, *66*(2), 382–392. https://doi.org/10.1111/biom.2010.66.issue-2

Redelmeier D. A., & Bell C. M. (2007). Weekend worriers. *New England Journal of Medicine*, *356*(11), 1164–1165. https://doi.org/10.1056/NEJMe068310

Robins G., Pattison P., & Wang P. (2009). Closure, connectivity and degree distributions: Exponential random graph (p*) models for directed social networks. *Social Networks*, *31*(2), 105–117. https://doi.org/10.1016/j.socnet.2008.10.006

Robins G., Pattison P., & Woolcock J. (2005). Small and other worlds: Global network structures from local processes. *American Journal of Sociology*, *110*(4), 894–936. https://doi.org/10.1086/427322

Schecter A., & Quintane E. (2021). The power, accuracy, and precision of the relational event model. *Organizational Research Methods*, *24*(4), 802–829. https://doi.org/10.1177/1094428120963830

Schoenfeld D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika*, *69*(1), 239–241. https://doi.org/10.1093/biomet/69.1.239

Seymour C. W., Gesten F., Prescott H. C., Friedrich M. E., Iwashyna T. J., Phillips G. S., Lemeshow S., Osborn T., Terry K. M., & Levy M. M. (2017). Time to treatment and mortality during mandated emergency care for sepsis. *New England Journal of Medicine*, *376*(23), 2235–2244. https://doi.org/10.1056/NEJMoa1703058

Snijders T. A. (2017). Stochastic actor-oriented models for network dynamics. *Annual Review of Statistics and its Application*, *4*(1), 343–363. https://doi.org/10.1146/statistics.2017.4.issue-1

Snijders T. A., Pattison P. E., Robins G. L., & Handcock M. S. (2006). New specifications for exponential random graph models. *Sociological Methodology*, *36*(1), 99–153. https://doi.org/10.1111/j.1467-9531.2006.00176.x

Snijders T. A., & Steglich C. E. (2015). Representing micro–macro linkages by actor-based dynamic network models. *Sociological Methods & Research*, *44*(2), 222–271. https://doi.org/10.1177/0049124113494573

Stadtfeld C. (2012). *Events in social networks: A stochastic actor-oriented framework for dynamic event processes in social networks*. KIT Scientific Publishing.

Stadtfeld C., & Block P. (2017). Interactions, actors, and time: Dynamic network actor models for relational events. *Sociological Science*, *4*, 318–352. https://doi.org/10.15195/issn.2330-6696

Stadtfeld C., Hollway J., & Block P. (2017). Dynamic network actor models: Investigating coordination ties through time. *Sociological Methodology*, *47*(1), 1–40. https://doi.org/10.1177/0081175017709295

Stadtfeld C., Mascia D., Pallotti F., & Lomi A. (2016). Assimilation and differentiation: A multilevel perspective on organizational and network change. *Social Networks*, *44*, 363–374. https://doi.org/10.1016/j.socnet.2015.04.010

Therneau T. M., Grambsch P. M., & Fleming T. R. (1990). Martingale-based residuals for survival models. *Biometrika*, *77*(1), 147–160. https://doi.org/10.1093/biomet/77.1.147

Tranmer M., Marcum C. S., Morton F. B., Croft D. P., & de Kort S. R. (2015). Using the relational event model (REM) to investigate the temporal dynamics of animal social networks. *Animal Behaviour*, *101*, 99–105. https://doi.org/10.1016/j.anbehav.2014.12.005

Vu D., Hunter D., Smyth P., & Asuncion A. (2011). Continuous-time regression models for longitudinal networks. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, K. Q. Weinberger (Eds.), *Advances in neural information processing systems* (Vol. *24*, pp. 2492–2500). Curran Associates, Inc.

Vu D., Lomi A., Mascia D., & Pallotti F. (2017). Relational event models for longitudinal network data with an application to interhospital patient transfers. *Statistics in Medicine*, *36*(14), 2265–2287. https://doi.org/10.1002/sim.v36.14

Vu D., Pattison P., & Robins G. (2015). Relational event models for social learning in MOOCs. *Social Networks*, *43*, 121–135. https://doi.org/10.1016/j.socnet.2015.05.001

Wolfram S. (1984). Universality and complexity in cellular automata. *Physica D: Nonlinear Phenomena*, *10*(1–2), 1–35. https://doi.org/10.1016/0167-2789(84)90245-8

Zachrison K. S., Amati V., Schwamm L. H., Yan Z., Nielsen V., Christie A., Reeves M. J., Sauser J. P., Lomi A., & Onnela J. -P. (2022). Influence of hospital characteristics on hospital transfer destinations for patients with stroke. *Circulation: Cardiovascular Quality and Outcomes*, *15*(5), 362–372. https://doi.org/10.1161/CIRCOUTCOMES.121.008269

Zappa P., & Vu D. Q. (2021). Markets as networks evolving step by step: Relational event models for the interbank market. *Physica A: Statistical Mechanics and its Applications*, *565*, 125557. https://doi.org/10.1016/j.physa.2020.125557