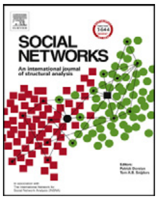
Social Networks 34 (2012) 506–514

Contents lists available at SciVerse ScienceDirect

Social Networks

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



Group based trajectories of network formation and dynamics

Sharique Hasan

Stanford University, Graduate School of Business, 655 Knight Way, Stanford, CA 94305, United States

ARTICLE INFO

Keywords:
Networks
Dynamics
Longitudinal

ABSTRACT

In this paper, we propose the application of a semi-parametric statistical methodology called Group-Based Developmental Trajectory Analysis to studying the dynamics of social networks. We begin with a discussion of theoretical problems in network analysis that may benefit from this approach. Next, we describe the methodology and how it can be applied to dyadic network data as well as aggregated node level data. We then demonstrate the methodology by analyzing the Newcomb Fraternity and the van de Bunt student data sets. Finally, we conclude with a discussion of potential directions for further research.

1. Introduction

The repeated interactions of individuals in a social setting create social networks. For example, entering college freshmen become friends with students in the same dormitory. As time progresses, some friendships are stressed and other friendships end. New relationships also emerge as people take classes across campus, attend parties that attract members of different social spheres, or are introduced to new acquaintances through their current friends. Many friendships also endure for long durations of time; a random student one meets at a campus-wide icebreaker can remain a close friend throughout life. These scenarios, describing the temporal status of a relation between two individuals, are *trajectories* (Nagin, 1999; Elder, 1985).

Dyadic trajectories describe only one aspect of network dynamics. Dyadic states also aggregate into measures that describe how actors are embedded in the larger network. For example, the in-degree centrality of an actor, measured by the number of alters who nominate him as a friend, quantifies one dimension of how this individual relates to others in the network. Because new ties can be created or dissolved, the centrality of individuals in social networks can also change. Therefore, the sequence of states describing the centrality of an actor in the social network can also be represented as a trajectory. Like dyadic trajectories, *trajectories of nodal position* are also likely to follow different temporal patterns; one trajectory might depict an individual whose centrality increases over time, while another trajectory might depict an individual whose centrality decreases.

In general, we may be justified in assuming that not all trajectories are likely to follow the same pattern of emergence, persistence, and decay. Consider friendship trajectories. Not all friendships persist, nor do they all decay. There are likely to be friendships that form and endure, other friendships that form and then end, and yet other potential friendships that never form. Nonetheless, it would be imprudent to take this logic to the extreme and argue that each trajectory is idiosyncratic; that is to say that there are no systematic temporal patterns across observations. It is more likely, however, that there is a finite set of possible trajectories that sufficiently describe the temporal evolutions of (friendship) ties. Finding this set of trajectories in the multi-wave social network data constitutes the first step in applying a trajectory perspective to studying network dynamics. Doing so allows us to describe the dynamics of social networks in two important ways. First, we are able to describe the shape of trajectories; second, we are able to quantify the relative frequency of the trajectories we observe.

The existence of a finite set of trajectories in the evolving social network raises a second question: are there systematic differences in the characteristics of the observations (i.e. dyads or nodes) that suggest that they are more likely to follow one trajectory rather than another? With respect to dyadic trajectories, we might ask whether friendships between individuals who share the same ethnicity are more likely to form and survive, rather than form and then decay. Similarly, for trajectories of nodal position we may ask whether certain individual characteristics, such as personality traits, are associated with sustained levels of high centrality over time, rather than fleeting popularity.

Thus, our task is twofold. First, we will need to discover in the multi-wave network data ‘groups’ of dyads or nodes that follow similar trajectories. Second, we will need to find, when appropriate, systematic differences in the characteristics of the observations that predict membership in the different ‘trajectory

---

E-mail address: hasan_sharique@gsb.stanford.edu

0378-8733/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.socnet.2012.04.001

groups.' In this article, we accomplish these two tasks by applying a semi-parametric statistical methodology called Group Based Developmental Trajectory Analysis (DTA) to multi-wave social network data. This methodology, pioneered by Nagin (1999), is commonly used to study the dynamics of criminal or delinquent behavior. The first aspect of the DTA methodology corresponds to the first task; it segments the panel data into a set of finite 'groups' which show similar patterns in the temporal dynamics of focal states and then estimates generalized linear models describing these patterns or trajectories. The second aspect of the methodology relates characteristics of the observations (dyads or nodes) to membership in each 'trajectory group.'

We have organized the remainder of this paper into four sections. A brief theoretical discussion focusing on the dynamics of interpersonal ties and network positions is presented in the first section. We then describe the Developmental Trajectory Analysis methodology and provide a discussion of its usefulness for studying dynamic social networks; several limitations are also outlined. In the third section, we demonstrate the utility of DTA by applying it to the classic Newcomb Fraternity data (Newcomb, 1961) and the van de Bunt student data (Snijders et al., 2010; Van de Bunt et al., 1999). These case studies are intended to demonstrate how DTA can be used to reveal the dynamics of nodal positions as well as the dynamics of dyads in social networks. Finally, we conclude by discussing the implications of using DTA for studying evolving social networks.

2. Trajectories and network dynamics

Although much of the networks literature has focused on understanding the structure and implications of social networks at the cross-section, researchers have increasingly begun to study network dynamics (Snijders and Doreian, 2010). Incorporating the dimension of time in the study of network structure raises several theoretical and empirical questions. For example, consider the cross-sectional study of tie presence and absence. Cross-sectional research suggests that actor similarity or homophily increases the likelihood that a tie between two actors is present. Homophily operates along several dimensions including geographic proximity, ethnicity, age, religion and education (McPherson et al., 2001). Additionally, endogenous structural processes within the network also predict the state of a tie. These processes include reciprocity (Carley and Krackhardt, 1996), balance (Doreian et al., 1996; Cartwright and Harary, 1956), as well as structural equivalence (White et al., 1976).

By considering time, questions about the causes of tie presence and absence can be re-framed as ones of tie creation, maintenance, and decay. To answer these questions, researchers have developed new approaches and analyzed data that yield important insights. Most prominent among these approaches is the actor-based model of network dynamics (Snijders, 2001; Snijders et al., 2010). Within this modeling framework, the analyst is able to understand both attribute-based (exogenous) and structure-based (endogenous) effects on the dynamics of social networks. Studies using this methodology have found support for actor preferences for same-sex homophily in friendship, dyadic reciprocity, transitivity, as well as other exogenous and endogenous preferences in tie formation in dynamic social networks (Snijders and Doreian, 2010).

The actor-based approach to analyzing longitudinal network data provides an important framework for understanding the mechanisms by which actors evaluate their network configuration as well as their subsequent decisions to change it (Snijders et al., 2010). That is to say, it is particularly useful for understanding the determinants of tie 'transitions.' Yet, understanding how these transitions string together to create 'trajectories' is an important area of focus itself (Elder, 1985). A tie trajectory can be conceptualized as a sequence of tie transition decisions embedded in time. In a network with dichotomous ties, there are four possible transitions: no tie to a no tie ($0 \rightarrow 0$), no tie to a tie ($0 \rightarrow 1$), tie to a tie ($1 \rightarrow 1$), and tie to a no tie ($1 \rightarrow 0$). While many possible trajectories can emerge from these transitions, we may expect to see only a handful of actual trajectories in a social network. Assuming an empty network in the first period, the three possible classes of trajectories are (a) 'null trajectories' where ties never exist, (b) 'monotonic' trajectories where tie probability increases or decreases over time, and (c) 'non-monotonic' trajectories where the probability of a tie increases and then decreases. Within trajectory classes 'b' and 'c' different variations are possible if the timing of the creation and deletion transitions are varied.

With this typology in mind, we can ask two elementary questions regarding tie trajectories. First, we can ask what classes of trajectories actually exist; second, we can attempt to quantify the relative frequency of these trajectories. Both questions are theoretically relevant. While many friendships can be described in terms of current states (e.g. 'is a friend' or 'is not a friend') many relations are described in terms of the experienced dynamic within that relation. For example, a person can be considered an 'old friend' or a 'new friend'; the first relation may imply early timing of tie creation while the second may imply late timing. Other relations may be relegated to categories such as 'used to be a friend.' These informal descriptions of dynamic relations have quantitative counterparts in the trajectory classes above. The 'used to be a friend' category implies a non-monotonic trajectory where a friendship relation was created but then it deteriorated after some time. Other mappings are also possible, and a focus on trajectories as a complement to transitions may be useful for answering questions about the existence and prevalence of certain trajectory classes within different social settings (e.g. schools vs. firms), different types of relations (e.g. friendship vs. advice), and periods of time (e.g. weeks vs. years).

Aside from identifying tie trajectories, understanding the characteristics – exogenous and endogenous – associated with each trajectory class may also be fruitful. For example, we can ask whether there are theoretically distinct types of homophily associated with monotonic vs. non-monotonic friendship or advice trajectories. While relating these factors to trajectories in a model-based framework is outside the scope of this paper, it is still possible to provide useful descriptions of such correlations.

2.1. Trajectories of nodal positions

Tie trajectories describe the state of a dyadic relation over time. A second type of trajectory – trajectories of nodal position – describe the temporal changes in the relative position of a node; these trajectories may also be an important focus of study. Social network researchers have developed several measures of 'centrality' to quantify the relative position of an actor in a social network. The measures range from local tie counts (degree centrality) to more complex measures such as betweenness and closeness that quantify the degree to which nodes act as bridges or can easily reach other nodes, respectively. Researchers who study network centrality have generally focused on three areas: quantifying positions in social network structure (see Freeman, 1979, 1977), understanding the determinants of positions (e.g. Ibarra, 1992) and understanding the implications of positions for important outcomes such as career progression or innovation (Burt, 2006).

While most research exploring these questions has been conducted at the cross-section, longitudinal studies are also appearing in the literature. In a recent paper by Moody et al. (2011) popularity trajectories were studied among samples of students in 28 communities across the United States. The findings suggest that there is significant change in popularity (in-degree) across the time period they studied and that these popularity trajectories affected the likelihood of substance abuse. Similarly, Bendersky and Shah (2010) find different trajectories of status (in-degree) in a sample of MBA students and differential impact of such trajectories on performance. This growing interest in 'nodal trajectories' has raised important issues for network theory – namely a conceptualization of 'position in social structure' as dynamic rather than static. It is therefore possible to ask some basic questions about these trajectories: what kinds of trajectories exist, how common are they, what factors predict membership in each trajectory and what are the effects of trajectory membership on relevant outcomes.

From a theoretical perspective, the first two questions can be thought of as temporal extensions to the quantification of centrality in static networks. This focus on 'dynamic positions' may be illuminating for several reasons. The existing research on network dynamics does seem to suggest intra-individual variability in network position over time; people's centrality can increase, decrease or remain stable. Furthermore, much like we assign social meaning to measures of static centrality (e.g. greater in-degree means popularity), similar socially descriptive labels can be assigned to trajectories of nodal positions. Colloquial descriptions of the 'social climber' (low to high) or the 'tragic hero' (high to low) are common labels for trajectories of nodal positions. Such trajectories are meaningful social descriptions of actors; they may, in fact, be more informative about how actors are treated and valued than static descriptions that focus on only on current or past position in the social network. Moreover, understanding the determinants of these trajectories as well as their impact on the actions and outcomes of actors may also be theoretically interesting.

In the next section, we provide a brief overview of a semi-parametric statistical methodology called Group-Based Developmental Trajectory Analysis or DTA (Nagin, 2005, 1999). We believe that DTA is a useful tool for studying the dynamics of social networks. The methodology is presented in this article as a descriptive, rather than analytic, procedure. Network data has a significant amount of interdependence across observations and dealing with this interdependence – temporal and relational – is outside the scope of this paper. Thus, DTA should be considered a complement, rather than a substitute, to existing approaches to analyzing dynamic network data that provide approaches for dealing with the independence issues inherent in network analysis.

3. Developmental trajectory analysis

A 'developmental trajectory' models the progression of a state, behavior, or action over time Nagin (2005). Examples of states that constitute dyadic tie trajectories are the presence or absence of a tie, the strength of a tie, or the count of a dyadic event (e.g. email sent). Likewise, the measurements that constitute trajectories of nodal positions include the centrality of an actor i at periods of time t. The Group-Based Developmental Trajectory Analysis (DTA) methodology assumes that sample of observations consists of a mixture of groups that follow a finite set of relatively homogeneous trajectories (Nagin, 2005). There are two aspects of the DTA methodology. In the first, a set of homogeneous 'clusters' or 'groups' of trajectories is discovered through a maximum likelihood procedure performed using a general quasi-Newton procedure (Jones et al., 2001). If covariates describing the observations are available, the DTA methodology can simultaneously relate characteristics of observations to membership in these trajectory groups. DTA has been applied in various settings; most prominently, it has been used to study criminal and aggressive behavior over the life course (Nagin and Tremblay, 1999) as well as the adoption of information technology (Zheng et al., 2005).

The methodology assumes that the researcher has collected data consisting of temporal sequences of observations for a focal state or behavior. In network analysis, these observations can be at the level of the dyad, the node, or the triad. For dyadic observations, denoted Y_{ijt}, i represents the sender, j the receiver, and t the time period in which the tie exhibits this state such that $Y_{ij} = \{Y_{ij1}, \ldots, Y_{ijT}\}$. Node level observations can be denoted as Y_{it}, with i representing the focal actor and t representing the time period in which the measurement of node position was taken such that $Y_i = \{Y_{i1}, \ldots, Y_{iT}\}$. Though there are no formal guidelines on the number time periods required for estimating trajectory models, the minimum number is clearly two; however, three or more periods are required if one would like to estimate trajectory models with non-linear terms.

As mentioned earlier, the methodology identifies a set of trajectories in the data and determines the relative frequency of each of these trajectories. Current implementations of DTA are able to model trajectories using three distributional assumptions. The binary logit implementation can be used to model the time varying dichotomous state of ties; the censored normal, continuous measures of centrality or tie strength; the zero-inflated Poisson, degree centrality or interaction counts. We begin the application of DTA by specifying the number, denoted as k, of trajectory 'groups' we would like to estimate in the data and the order of the polynomial degree for each of the k trajectory groups. Nagin (2005, pp. 66–67) argues that selecting the number of groups is the primary analytic task and the "choice of the order of the trajectory for each group is of less importance." Taking a logit model as our example, we then estimate a set of regressions corresponding to each of the k groups:

$$p^k(y_{ijt}) = \frac{exp(\beta_0^k + \beta_1^k T + \beta_2^k T^2)}{1 + exp(\beta_0^k + \beta_1^k T + \beta_2^k T^2)} \tag{1}$$

Next, we take the results of this estimation, the $p^k(y_{ijt})$, and for each sequence Y_{ij} compute the product over the T time periods, such that:

$$P^k(Y_{ij}) = \prod^{T} p^k(y_{ijt}) \tag{2}$$

Eq. (2) represents the probability of observing a sequence of states given their membership in group k. Summing across all groups we have the unconditional probability of observing a sequence Y_{ij}, which can be written as $P(Y_{ij}) = \sum^K \pi_k P^k(Y_{ij})$. Here the π_k represents the proportion of sequences that are classified into group k. These elements are then combined in the likelihood function represented by Eq. (3):

$$L = \prod^{N} \sum^{K} \pi_k P^k(Y_{ij}) \tag{3}$$

Maximizing the likelihood function allows us to estimate the parameters that represent the relative frequency of each of the trajectory groups and the shape of each trajectory. A parameter, π_k, is estimated which is the proportion of sequences that are categorized in each group. The second set of estimated parameters is the coefficients on the regression models; in our case, these are the logistic regressions β_0^k, β_1^k, β_2^k in Eq. (1). After estimating the coefficients, the posterior probability that a sequence Y_{ij} belongs in group k can be estimated using Eq. (4).

$$\hat{P}(k|Y_i) = \frac{\hat{P}(Y_i|k)\hat{\pi}_k}{\sum^K \hat{P}(Y_i|k)\hat{\pi}_k} \tag{4}$$

As with any clustering procedure, the researcher is required to a priori select the appropriate number of trajectory groups and the degree of each polynomial. There are several recommended approaches for accomplishing this task. From a quantitative perspective, the DTA methodology uses the Bayesian Information Criterion (BIC) as the fit statistic. BIC is computed as:

$$\text{BIC} = \log(L) - \frac{1}{2} \times \log(n) \times r \tag{5}$$

In this formulation, $\log(L)$ is log of the maximized likelihood function, $\log(n)$ is the log of the sample size, and r is the number of parameters in the model; it is recommended that the model with the highest score is chosen. In practice, the BIC score is often not a sufficient measure for overall model appropriateness; thus, Nagin (2005, p. 77) and others recommend a more holistic treatment of model selection. Researchers are urged to apply domain understanding and study goals as additional criteria for selecting the appropriate trajectory model.

In addition to identifying trajectories, DTA also supports linking covariates to trajectory group membership. This is conventionally done through the use of a multinomial regression model whereby the probability of group membership is modeled as linear combination of covariates that describe characteristics of each of the sequences. For example, in understanding the factors correlated to membership in a trajectory of nodal position, the researcher may be interested in understanding whether males are more likely, as compared to females, to belong to one trajectory group rather than another. There are also less formal ways of describing the characteristics of trajectory group members. Tables of summary statistics are often another way of providing such descriptions.

It is perhaps useful to place DTA in the networks literature as well as among other latent class approaches in general. In a broader sense, the DTA methodology is a latent class statistical approach for longitudinal data. Cross-sectional latent class approaches already have an established place in network analysis; for instance, classic work on block modeling White et al. (1976) and more recent latent space approaches for determining groups in cross-sectional social network data are other such examples (Hoff et al., 2002; Davis and Carley, 2008). DTA departs from many similar approaches in network analysis because of its focus on identifying groups with a temporal dimension rather than groups within a cross-section.

The DTA methodology is also similar to more traditional cluster analysis methods such as k-means clustering or hierarchical clustering (Kaufman and Rousseeuw, 1990; Dumenci and Windle, 2001). DTA also differs from cluster analysis in two important respects (Dumenci and Windle, 2001). First, while cluster analysis can be used to determine temporal trajectories – traditional clustering approaches, unlike DTA, are agnostic to the temporal dimensions of the data. DTA considers the temporal dimension by modeling the trajectories as polynomial functions of time. The individual vectors of network observations $Y_i = \{Y_{i1}, \ldots, Y_{iT}\}$ to be clustered are not considered random elements of $\mathbb{R}^T$ with independent components. In the DTA methodology, the components depend on each other as values of a polynomial at each point in time. Second, while clustering approaches may also find comparable groups in the data, the DTA methodology provides a general framework for deciding on the number of groups using the BIC, probabilistically assessing the likely membership of observations in groups, modeling the shape of each trajectory using different distributional assumptions, and linking co-variates to group membership. Furthermore, a study by Dumenci and Windle (2001) found that cluster analysis was able to find distinct clusters in longitudinal data when there were large differences in the initial values of trajectories; however, in their analyses cluster analysis was unable to detect distinct clusters if differences existed in the shapes of trajectories. This latter difference, is in many ways just as interesting, if not more, than large differences in initial values. However, in practice, traditional clustering approaches and DTA are often comparable in the ability to partition the data into distinct groups. Finally, DTA is akin to approaches such as generalized growth mixture modeling (GGMM) proposed by Muthen (2001) as well as the mixture latent Markov models proposed by Vermunt et al. (2008). Some discussion of the difference between DTA, GGMM and other trajectory approaches is presented in Nagin (2005).

Although the DTA methodology is useful for describing multi-wave social network data, there are several important limitations to consider when interpreting the results. Chief among these are the independence assumptions in DTA that are known to be violated by the network data (see for approaches to handling interdependence in cross-sectional social network data Robins et al., 2007; Dekker et al., 2007). Specifically, the DTA methodology models temporal dependence as a polynomial function of time conditional on group membership. This modeling strategy assumes temporal dependence as a polynomial function of time, but ignores more complex stochastic dependence over time and dependence between actors. However, incorporating temporal and structural dependencies will require a much more complex model. Thus, the results should at this stage be viewed as descriptions; future extensions should be expected to help resolve the latter limitations.

In the next section of the paper, we demonstrate the utility of the DTA methodology for studying the dynamics of social networks. We present two case studies: the classic fraternity data of Theodore Newcomb (Newcomb) and the student data of van de Bunt (VDB) are re-analyzed. For each of the data sets we use DTA to study tie trajectories as well as trajectories of nodal position. For the VDB data, we also relate the characteristics of dyads with membership in trajectory groups for the dyadic tie trajectories.

4. Case studies

In this section, we present two case studies that apply the Developmental Trajectory Analysis methodology to multi-wave social network data. For each case study, we begin with a description of the data, followed by the application of DTA to uncovering tie trajectories, and then trajectories of nodal position. We also attempt to relate, when possible, characteristics of the observations to membership in each trajectory group. Since both sets of network data have been analyzed before, we also compare our results to prior results when appropriate. We estimated our models using the PROCTRAJ procedure implemented in SAS (Jones et al., 2001; Jones and Nagin, 2007).

4.1. *Newcomb's second fraternity*

In our first case study, we apply DTA to the second series of 'fraternity' data collected by Theodore Newcomb (Newcomb, 1961). The data are a series of fifteen network observations consisting of mutual rankings by seventeen male students living together at the University of Michigan. For each time period, denoted t, each of the seventeen participants, denoted i, rank the remaining sixteen. A ranking of '1' indicates that ego considers the alter 'most favorable' whereas a ranking of '16' indicates that the alter is considered 'least favorable.' There are no ties in this data. To ease interpretation, we subtracted the original ranking from 17. Thus, higher rankings imply that the alter is rated favorably; lower rankings, less favorably.

We begin the case study by estimating trajectories of nodal position. For this estimation, we create a time varying measure of nodal position by computing the average ranking that each participant i received from the other sixteen at each point in time t – denoted Y_{it}. This measure represents a simple measure of 'social status' or 'popularity', akin to in-degree for continuous data. Individuals with the highest score on this measure are rated more favorably, on average, than their peers. Similarly, individuals with low scores on this measure are viewed less favorably. We used the censored normal implementation of DTA in this example; it is most appropriate for modeling continuous data such as the average ranking (Nagin, 2005). The methodology, like other clustering approaches, requires the analyst to specify the number of trajectory groups a priori. We estimated four models ranging from '1' group to '4' groups. The two best fitting models were the three (BIC = −409.81) and four group models (BIC = −404.43). The key difference between these two models is the split in the 'middle status' group in the four group model; thus, we chose to describe the results of the three group model since it provides similar intuition with less complexity. The parameter Sigma describes the amount of variation in the data explained by the estimated trajectory model.

The results of this estimation are graphically depicted in Fig. 1 and the estimates of the trajectory model are presented in Table 1. The vertical axis represents the average ranking received by the actors (i.e. status or popularity) and the horizontal axis represents time (weeks, in this case). To ease interpretation we have labeled the three trajectory groups: *high status*, *middle status*, and *low status*. The high status group is represented by the flat line near the top of the vertical axis; the model for this trajectory has an significant intercept with $\beta = 10.576$, but the other terms are indistinguishable from zero. The high status group consists of only four individuals; these participants are ranked highly by their peers from the beginning of the study until the end. Members of this trajectory group correspond well to the 'top block' in the White et al. (1976) re-analysis; three of the four men classified in the high status trajectory are top block members.

More interesting are the trajectories of the middle status and the low status groups. The direction, magnitude, and significance of the estimation results in Table 1 suggest that quadratic polynomials sufficiently describe the shapes of these two trajectory groups. Although the high status group is differentiated from the beginning, the rest of the hierarchy is not. During the first week, members of the low status and high status groups are indistinguishable in terms of their average ranking with average rankings of 7.177 and 7.159, respectively. After the second week, a consensus begins to form and middle status and low status groups differentiate; by the sixth week a rigid status hierarchy is apparent. Membership in the low status group corresponds well to the 'bottom block' of the White et al. (1976) analysis.

We also estimated several other node trajectory models with a larger number of groups. The results are similar: trajectories stabilize at about the 5th or 6th week and the status hierarchy becomes stable. Several aspects of these results are important to consider. Although there is variability across individuals in their actual trajectory – there are basically three general trajectory patterns. First, a core of high status actors were present since the beginning and they maintained this status throughout the observation period. Second, there was no consensus about the membership in the middle and lower parts of the status distribution in the first week; a consensus formed relatively soon thereafter and then hierarchy remained stable. Finally, the absence of an upward sloping trajectory suggests that significant upward mobility was not common in this setting. Although data limitations prevent us from answering questions about *why* such patterns exist in the Newcomb Fraternity, several important questions do arise. For instance, what predicts membership in a trajectory group and are the associated predictors characteristics of the individuals, exogenous structural characteristics, or endogenous structural features of the network?

In addition to estimating nodal trajectory models, we also conducted a second analysis to identify dyadic tie trajectories in the Newcomb data. We estimated several models with varying number of trajectory groups. The best fitting model was a six group model (BIC = −9580.47). The results of this estimation are presented as the model in the right column of 1 and graphically depicted in Fig. 2. The horizontal axis in the figure represents time and the vertical axis represents the ranking that an actor j received from actor i. Of the six trajectories that were identified, two were relatively stable throughout the observation period. The 'low stable' and 'high high' trajectories consist of ratings where individuals were either disfavored or looked favorably upon throughout. Two trajectories, the 'mid-declining' and 'low high', begin at similar positions with high initial ratings, however one remains high and the other tends to decline. The path of these two trajectories may indicate similar first impressions, with the 'low high' trajectory representing the impressions remaining favorable and the 'mid-declining' being the impressions that were modified after further interaction. A similar, though inverted pattern can be seen with the 'low increasing' and 'low declining' trajectories. Although they begin at similar positions in the first week, they subsequently diverge. The 'low increasing' trajectory with a low initial first impression increases while the other decreases significantly. In Table 1 we can see that the coefficient magnitudes and significance levels suggest that the estimated trajectory model fits the data well.

Application of DTA to the Newcomb data uncovered stable temporal patterns or trajectories. The results suggest that both positions within the network as well as dyadic relations may be classified into 'types' that describe their sequence of temporal states. The presence of these relatively well defined trajectories suggests possibilities for studying dynamic networks. First, the presence of trajectories suggest that individual decisions to alter the state of dyadic ties aggregate into stable dynamic patterns of nodal position as well as stable temporal relationship patterns. Whether these trajectories are mere artifacts of myopic decision making on the part of actors or are viewed as 'social structures' themselves is an important question.

4.2. Gerhard van de Bunt's students

For our second case study, we apply DTA to the dynamic network data collected by Van de Bunt et al. (1999). The data consist of 7 waves of network data for 32 university freshmen. This data has some amount of missing observations, a common problem with longitudinal network data. However, because we do not have a model for the missing data, we have decided to exclude the missing data from our estimations. Ignoring the missing data often results in less error in estimates than using simple imputation methods (Huisman, 2009). The first four temporal observations were collected three weeks apart and the subsequent three observations were gathered seven weeks apart. In each period, the participants were asked to indicate the nature of their relationship with the other students. Six types of responses were possible for each dyad: best friendship, friendship, friendly relationship, neutral relationship, unknown person, and troubled relationship. We recoded the data so that best friendship, friendship and friendly relationship were coded as '1' and all other relations were coded as '0'.

We begin the case study using the VDB data by estimating trajectories of nodal position. Using the re-coded sociomatrix described above we compute in-degree for all actors i for each of the time periods t as the sum of the incoming 'friendship ties'; this time varying measure of in-degree is denoted as Y_{it}. The measure can be interpreted as social status or popularity; individuals with higher in-degree are more popular. It is important to note that unlike the Newcomb data where the increase in status for one actor necessarily implies a decrease in status for another, this is not the case for the VDB data since participants are not forced to rank alters without the possibility of ties.

The results of this estimation are graphically depicted in Fig. 3 and the estimates of the trajectory model are presented in Table 2. We estimated zero-inflated Poisson models (Nagin, 2005) with varying number of groups ranging from one to four. The estimation procedure did not converge for the four group model and produced

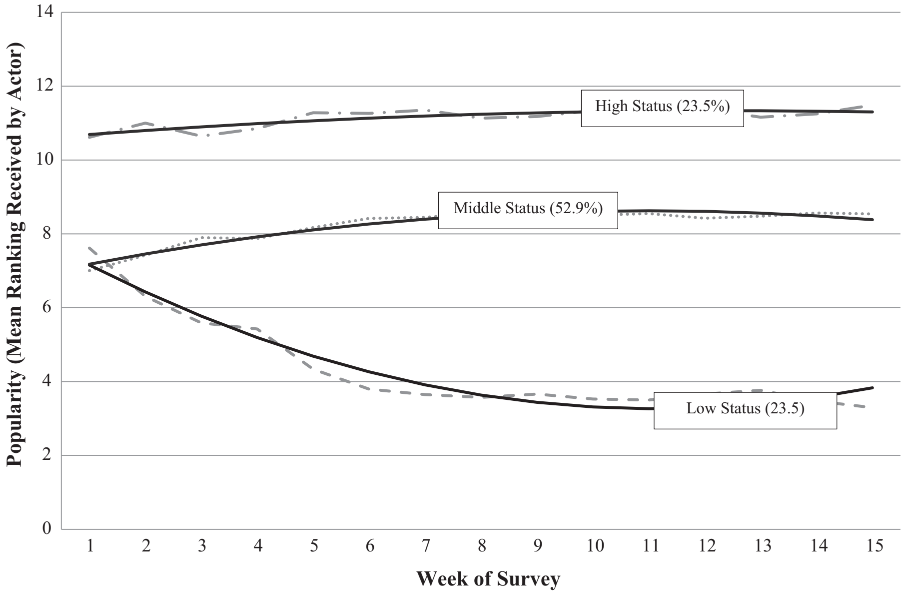


Fig. 1. Trajectories of nodal position for the Newcomb data.

a group with 0% of the observations for the three group model; thus, we describe a two group model (BIC = −464.69). The parameters Alpha0 and Alpha1 are coefficients from a logistic regression modeling the zero-inflation for both trajectory groups. The estimates in Table 2 suggest that the low in-degree trajectory begins somewhere near zero, but there is a significant upward trend. On the other hand, the high in-degree trajectory estimation indicates that these individuals begin with a low non-zero number of friends but also experience an increase in in-degree. The group proportions indicate that individuals who are members of the high in-degree trajectory constitute the majority of individuals (72%) who begin with a few friends but end up with somewhere between six to eight friends by the end of the observation period. For the second trajectory group, moderate in-degree (28%), we find a lower starting point and a lower end point; in the final observation period the in-degree of this group is approximately four.

We also estimated dyadic tie trajectories for the VDB data. For these estimations, the dependent variable was the absence or presence of a friendship tie; thus, we use the logistic regression implementation of DTA. Although we estimated models with varying number of trajectory groups, the three group model yielded the most appropriate estimation when co-variates were included in the estimation; the BIC for this model was BIC = −1632.11. These results are presented in Table 2 and the graphical representation

Table 1
Trajectory estimates for Newcomb data.

Node position trajectories				Dyadic tie trajectories			
Group	Parameter	Estimate	Standard error	Group	Parameter	Estimate	Standard error
Low status	Intercept	7.965	0.480**	High/high	Intercept	5.279	0.203**
	Linear	−0.844	0.138**		Linear	−0.248	0.024**
	Quadratic	0.038	0.008**	Low/stable	Intercept	3.499	0.746**
Middle status	Intercept	6.871	0.320**		Linear	0.349	0.175*
	Linear	0.321	0.092**		Quadratic	−0.011	0.009
	Quadratic	−0.015	0.006*	Low/increasing	Intercept	5.412	0.524**
High status	Intercept	10.576	0.480**		Linear	0.884	0.121**
	Linear	0.122	0.138		Quadratic	−0.038	0.007**
	Quadratic	−0.005	0.008	Mid/declining	Intercept	13.036	0.468**
					Linear	−1.042	0.160**
					Quadratic	0.042	0.010**
				Low/high	Intercept	10.989	0.310**
					Linear	0.408	0.091**
					Quadratic	−0.019	0.005**
				Low/declining	Intercept	14.794	0.599**
					Linear	0.562	0.169**
					Quadratic	−0.034	0.010**
	Sigma	1.098	0.050			2.483	0.029
Low status		24%		High/high		19%	
Middle status		53%		Low/stable		18%	
High status		23%		Low/increasing		17%	
				Mid/declining		15%	
				Low/high		23%	
				Low/declining		8%	
	BIC	−409.81			BIC	−9617.12	
	N	17			*N*	272	

* $p < .05$.
** $p < .01$.

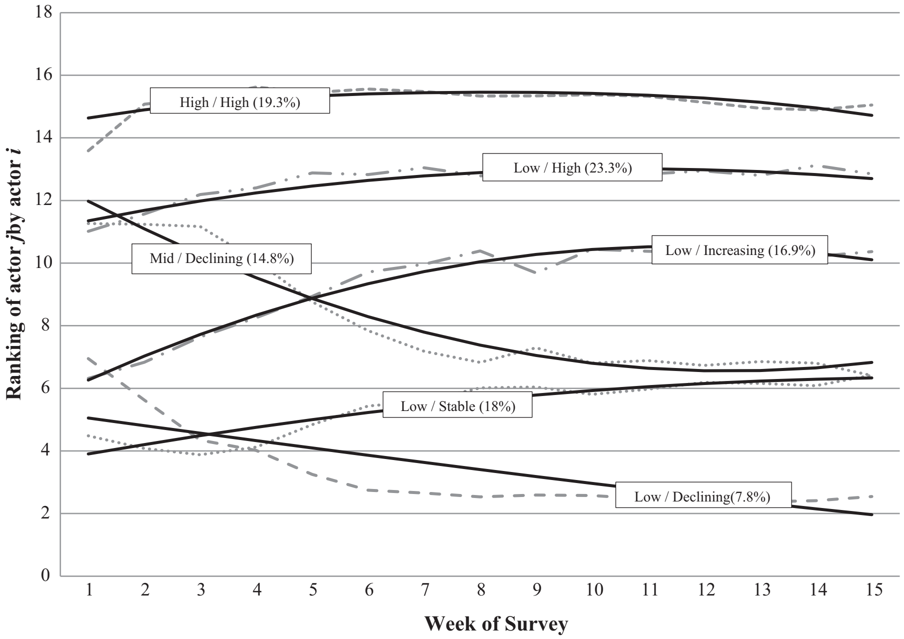


Fig. 2. Tie trajectories for the Newcomb data.

is presented in Fig. 4. The horizontal axis represents time and the vertical axis represents the probability of an $i \rightarrow j$ tie at time period t. The direction and significance of the trajectory parameter estimates are appropriate for each of the trajectory group polynomials. The largest trajectory group is the 'not friends' trajectory, making up approximately 75% of all directed dyad sequences in the data. The next two trajectory groups are the 'slow friendship' and 'fast friends' groups, making up 14% and 11% of all dyads, respectively.

The dyadic tie trajectory estimation for the VDB data also linked co-variates to the probability of group membership. The results for this estimation are presented in Table 3. We note that this analysis is exploratory. Thus, our results should be considered preliminary and descriptive at this stage. With these limitations in mind, we find two suggestive effects of exogenous covariates on the likelihood a dyad will follow a certain trajectory. First, we find that dyads where both individuals are male are more likely to be members of the fast friends trajectory ($p \leq .05$). Furthermore, there is a small effect suggesting that dyads where both members are enrolled in program 4 are less likely to be in the "not friends" trajectory.

We compared these results to the analysis conducted in Snijders (2005); their finding suggested that for "female students the value of a friendship with a male or a female other student is about the same, while male students have a clear preference for friendships to other males." Our results do not contradict this statement – rather the differential likelihood of a homophilous male dyad to follow the fast friends trajectory suggests a possible mechanism producing

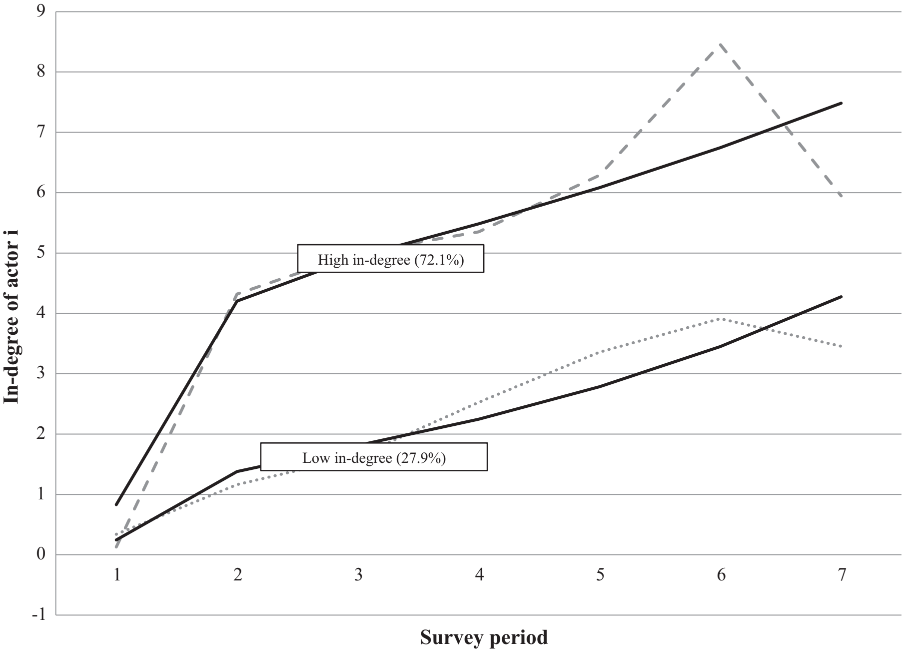


Fig. 3. Trajectories of nodal position for van de Bunt students.

Table 2
Trajectory estimates for van de Bunt data.

Node position trajectories				Dyadic tie trajectories			
Group	Parameter	Estimate	Standard error	Group	Parameter	Estimate	Standard error
Low in-degree	Intercept	−0.05	0.387	Slow friendship	Intercept	−3.791	0.311**
	Linear	0.215	0.066**		Linear	0.787	0.07**
High in-degree	Intercept	1.287	0.112**	Not friends	Intercept	−6.699	0.805**
	Linear	0.104	0.021**		Linear	1.113	0.385**
					Quadratic	−0.093	0.043*
				Fast friends	Intercept	−7.932	1.037**
					Linear	5.747	0.766**
	Alpha0	5.501	1.603**		Quadratic	−0.595	0.090**
	Alpha1	−4.156	1.354**				
Low in-degree		28%		Slow friendship		14%	
High in-degree		72%		Not friends		75%	
				Fast friends		11%	
	BIC	−464.49			BIC	−1632.11	
	N	32			*N*	992	

* $p < .05$.
** $p < .01$.

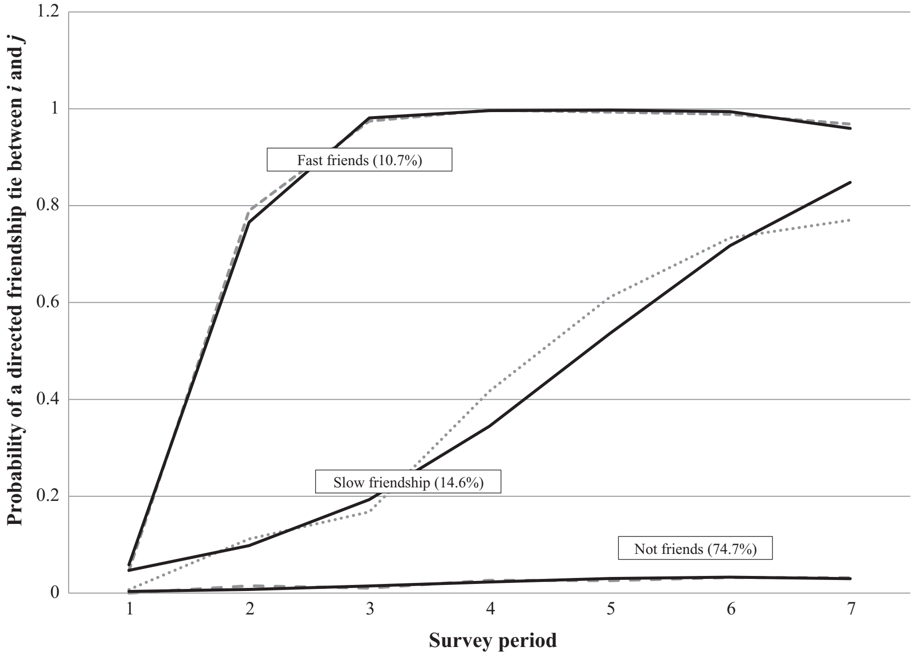


Fig. 4. Tie trajectories for van de Bunt students.

Table 3
Multinomial logistic regression predicting membership in trajectory group for VDB dyadic trajectories.

	Parameter	Estimate	Standard error
Not friends	Constant	1.749	0.204
	Both in program 2	0.874	1.064
	Both in program 3	0.361	0.483
	Both in program 4	−0.426	0.257^{+}
	Both smokers	0.164	0.343
	Both female	−0.184	0.251
	Both male	1.778	1.325
Fast friends	Constant	−0.273	0.281
	Both in program 2	0.896	1.142
	Both in program 3	−0.051	0.675
	Both in program 4	−0.367	0.371
	Both smokers	−0.554	0.54
	Both female	0.013	0.353
	Both male	2.746	1.375*

* $p < .05$.
$^{**}p < .01$.
$^{+}$ $p < .1$.

the earlier results. Although both male and female students may have equivalent initial propensities to form homophilous ties, homophilous male ties are more likely to become friends at a faster rate, and also remain friends throughout the observation period. Thus, over the entire observation period we would likely observe more same gender homophily among males than females.

5. Discussion

The intention of this article is twofold. First, we attempted to outline some problems in network theory that may benefit from taking a trajectory perspective. Second, we demonstrate the applicability of the DTA methodology for studying network dynamics. The methodology is demonstrated using two network case studies: the Newcomb Fraternity and the van de Bunt students. For each of the case studies we estimate dyadic tie trajectories, describing the dynamics of directed relations between actors in a social network, and trajectories of nodal positions, describing the dynamics of actors' network positions or centrality.

While most research examining network dynamics has focused on transitions – the change in one network state to another (for exceptions see Moody et al., 2011) – studying trajectories offers a complementary approach to understanding the temporal dimensions of social structure. The presence of certain trajectories as well as the absence of others suggests typical temporal patterns of social structure at both the dyadic and nodal level. At the level of the dyad, we find three basic patterns within the data: (1) stable trajectories, (2) monotonically increasing trajectories and (3) monotonically decreasing trajectories. Most dyadic trajectories fall within these categories. For example, all dyadic trajectories in the Newcomb case are either stable (e.g. high/high, low/high, mid/declining, and low/stable) or monotonically increasing (low/increasing) or decreasing (mid/declining and low/declining). This is not to say, however, that non-monotonic trajectories do not exist within the data. Rather, their frequency is limited in the data sets we analyze in this paper.

The application of the DTA methodology to the study of social networks helps us make progress towards answering an important question in network theory: “what is the developmental sequence of network structure over time?” (Granovetter, 1973, p. 1378). We find that the trajectory approach is useful for describing, in several ways, multi-wave network data. In the present study, the application of DTA is primarily descriptive and inductive; it can fruitfully serve as a complement to more formal model-based approaches for analyzing dynamic social networks.

At this stage, the trajectory approach is particularly useful for theory generation; although the possibilities for applying DTA to problems in social network analysis are many. The simplest application is the use of the methodology for determining the types of trajectories that are likely to be observed in different social contexts and across different types of ties. For instance, how common are non-monotonic friendship trajectories or other types of trajectories for that matter? Similarly, many interesting questions arise once trajectory groups have been determined: are certain types of homophilous ties more likely to follow one trajectory group rather than another? To this point, in the case study using the van de Bunt data, we find that male same-gender dyads are more likely to follow the fast friends trajectory. Furthermore, trajectories may capture important features of social relations beyond the absence or presence of a tie that could have differential impact on the adoption of innovations, the spread of rumors, or other processes that unfold over networks.

Although the case studies presented in this article consist of relatively small samples of students, there are opportunities to apply the methodology to other types of network data. For instance, DTA may in fact be appropriate for analyzing multi-wave ego-network data or data from online social networks. The full potential of the DTA methodology in the context of social network analysis has not been demonstrated in this paper. We expect that key extensions, modifications, and applications of the methodology will arise from the analysis of substantive network questions by scholars studying these topics.

Acknowledgement

The author would like the editor and the three anonymous referees for their very helpful and detailed comments.

References

Bendersky, C., Shah, N., 2010. The cost of status enhancement: performance effects of individuals’ status mobility in task groups. Organization Science, p. Online first.

Burt, R.S., 2006. Political disagreement: the survival of diverse opinions within communication networks. American Journal of Sociology 112 (1), 310–312.

Carley, K.M., Krackhardt, D., 1996. Cognitive inconsistencies and non-symmetric friendship. Social Networks 18 (1), 1–27.

Cartwright, D., Harary, F., 1956. Structural balance: a generalization of Heider’s theory. Psychological Review 63 (5), 277–293.

Davis, G., Carley, K., 2008. Clearing the FOG: fuzzy, overlapping groups for social networks. Social Networks 30 (3), 201–212.

Dekker, D., Krackhardt, D., Snijders, T.A.B., 2007. Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. Psychometrika 72 (4), 563–581.

Doreian, P., Kapuscinski, R., Krackhardt, D., Szczypula, J., 1996. A brief history of balance through time. Journal of Mathematical Sociology 21 (1–2), 113–131.

Dumenci, L., Windle, M., 2001. Cluster analysis as a method of recovering types of intraindividual growth trajectories: a Monte Carlo study. Multivariate Behavioral Research 36 (4), 501–522.

Elder, G., 1985. Life Course Dynamics: Trajectories and Transitions, 1968–1980. Cornell University Press.

Freeman, L., 1977. A set of measures of centrality based on betweenness. Sociometry 40 (1), 35–41.

Freeman, L., 1979. Centrality in social networks: conceptual clarification. Social Networks 1 (3), 215–239.

Granovetter, M., 1973. Strength of weak ties. American Journal of Sociology 78 (6), 1360–1380.

Hoff, P., Raftery, A., Handcock, M., 2002. Latent space approaches to social network analysis. Journal of the American Statistical Association 97 (460), 1090–1098.

Huisman, M., 2009. Imputation of missing network data: some simple procedures. Journal of Social Structure 10 (1), 1–29.

Ibarra, H., 1992. Homophily and differential returns: sex-differences in network structure and access in an advertising firm. Administrative Science Quarterly 37 (3), 422–447.

Jones, B., Nagin, D., Roeder, K., 2001. A SAS procedure based on mixture models for estimating developmental trajectories. Sociological Methods & Research 29 (3), 374.

Jones, B.L., Nagin, D.S., 2007. Advances in group-based trajectory modeling and an SAS procedure for estimating them. Sociological Methods & Research 35 (4), 542–571.

Kaufman, L., Rousseeuw, P., 1990. Finding Groups in Data: An Introduction to Cluster Analysis, vol. 39. Wiley Online Library.

McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Annual Reviews in Sociology 27 (1), 415–444.

Moody, J., Brynildsen, W.D., Osgood, D.W., Feinberg, M.E., Gest, S., 2011. Popularity trajectories and substance use in early adolescence. Social Networks 33 (2), 101–112.

Muthen, B., 2001. Second-generation structural equation modeling with a combination of categorical and continuous latent variables: new opportunities for latent class/latent curve modeling. In: New Methods of the Analysis of Change. American Psychological Association.

Nagin, D., 1999. Analyzing developmental trajectories: a semiparametric, group-based approach. Psychological Methods 4, 139–157.

Nagin, D., 2005. Group-based Modeling of Development. Harvard University Press.

Nagin, D., Tremblay, R., 1999. Trajectories of boys’ physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. Child Development 70 (5), 1181–1196.

Newcomb, T.M., 1961. The Acquaintance Process. Holt, New York.

Robins, G., Pattison, P., Kalish, Y., Lusher, D., 2007. An introduction to exponential random graph (p^*) models for social networks. Social Networks 29 (2), 173–191.

Snijders, T., 2001. The statistical evaluation of social network dynamics. Sociological Methodology 31 (1), 361–395.

Snijders, T., 2005. Models for longitudinal network data. Models and Methods in Social Network Analysis, 215–247.

Snijders, T., Van de Bunt, G., Steglich, C., 2010. Introduction to stochastic actor-based models for network dynamics. Social Networks 32 (1), 44–60.

Snijders, T.A., Doreian, P., 2010. Introduction to the special issue on network dynamics. Social Networks 32 (1), 1–3.

Van de Bunt, G., Van Duijn, M., Snijders, T., 1999. Friendship networks through time: an actor-oriented dynamic statistical network model. Computational & Mathematical Organization Theory 5 (2), 167–192.

Vermunt, J., Tran, B., Magidson, J., 2008. Latent class models in longitudinal research. In: Handbook of Longitudinal Research: Design, Measurement, and Analysis, pp. 373–385.

White, H., Boorman, S., Breiger, R., 1976. Social structure from multiple networks. I. Block models of roles and positions. American Journal of Sociology 81 (4), 730–780.

Zheng, K., Padman, R., Johnson, M., Diamond, H., 2005. Understanding technology adoption in clinical care: clinician adoption behavior of a point-of-care reminder system. International Journal of Medical Informatics 74 (7–8), 535–543.