PSYCHOMETRIKA
https://doi.org/10.1007/s11336-020-09700-x

A LATENT SPACE NETWORK MODEL FOR SOCIAL INFLUENCE

TRACY SWEET

UNIVERSITY OF MARYLAND

SAMRACHANA ADHIKARI

NEW YORK UNIVERSITY SCHOOL OF MEDICINE

Social network data represent interactions and relationships among groups of individuals. One aspect of social interaction is social influence, the idea that beliefs or behaviors change as a result of one's social network. The purpose of this article is to introduce a new model for social influence, the latent space model for influence, which employs latent space positions so that individuals are affected most by those who are "closest" to them in the latent space. We describe this model along with some of the contexts in which it can be used and explore the operating characteristics using a series of simulation studies. We conclude with an example of teacher advice-seeking networks to show that changes in beliefs about teaching mathematics may be attributed to network influence.

Key words: social networks, network models, latent space models, Bayesian, social influence.

1. Introduction

The term social network analysis refers to a collection of quantitative methods used when analyzing relational data, that is, individuals and their respective relationships. The types of relationships vary by context, but in the social sciences, we are often most interested in the interactions among individuals, e.g., students in a classroom, employees in a workplace, individuals in a family, and we use network data to explore these relationships. For example, we might be interested in why certain individuals interact more than others or how information travels across a network.

A field of study in social network analysis is social influence, the idea that individuals become more similar to those with whom they interact or are most closely connected. In the network literature, this phenomenon is also called contagion or diffusion and has been employed in numerous studies around the spread of disease such as HIV (Gupta et al., 1989) or the flu (Christakis and Fowler, 2010) as well as social behavior such as smoking (Steglich et al., 2012).

Unlike disease, which passes physically among individuals, social influence is more subtle. First, individuals can be influenced by those with whom they do not directly interact, and second, social influence does not necessarily occur independently. For example, a single person can be particularly influential if they are a social referent, someone with high social status. In fact, network structure often plays a role in social influence but formalizing the exact dependence structure (which may differ for each network) is quite difficult.

We would like acknowledge Jim Spillane and the Distributed Leadership Studies at Northwestern University, including the NebraskaMATH study, for the use of their data in this work. We greatly appreciate the help provided by Jim and his colleagues.

This work was support in part by the US Department of Education, Institute of Education Sciences under Grant #R305D150045. Any opinions, findings, and conclusions or recommendations expressed in this materials are those of the authors and do not necessarily reflect the views of the granting agencies.

Correspondence should be made to Tracy Sweet, Department of Human Development and Quantitative Methodology, University of Maryland, College Park, USA. Email: tsweet@umd.edu

Published online: 27 March 2020

Consider smoking as a context, student A's opinion about smoking may be influenced by the opinions of their friends but may also be influenced by the opinions of a popular student. Furthermore, student A's opinion is potentially dependent on—to varying degrees—the opinions of every other student in the network. One way to specify this dependence structure is to use a social network model to incorporate as much of the full network structure as possible into a model of influence, and social influence models include models that represent this relationship between network structure and node-level attributes in a variety of ways.

One common set of models for social influence are network autocorrelation models (NAMs; Doreian, 1989; Leenders, 2002). A less well-known but similar model is the autologistic actor attribute model (ALAAM; Robins et al., 2001; Daragonova and Robins, 2013). Both classes of models also account for the interdependence of network ties. Network autocorrelation models incorporate a weight matrix often based on network ties. These and related models have been used in a number of settings such as hospital adoption of electronic health record systems (Zheng et al., 2010), teacher technology use (Frank et al., 2004), telework behavior (Scott et al., 2012), and adolescent drinking behavior (Fujimoto et al., 2013). ALAAMs use an autologistic framework and borrow from the configurations often used in exponential random graph models (Anderson et al., 1999; Robins et al., 2007) to accommodate the network dependence structure. Although less commonly used, these models have been used in social category learning (Kashima et al., 2013).

Most current applications of NAMs assume both that social influence occurs through direct contact between two nodes and that nodes are equally influenced by those with whom they share a tie (Frank et al., 2004; Fujimoto et al., 2013). In reality, however, people are often differentially influenced by those with whom they have connections, and they can be influenced by those with whom they do not have direct ties. In fact, Paluck and Shepherd (2012) and Shepherd and Paluck (2015) have found that socially referent students are particularly influential with respect to community norms and behavior. Further, the study of social referents is not limited to education and is found across disciplines such as marketing (Watts and Dodds, 2007) and the workplace (Eckhardt et al., 2009).

Therefore, we propose the latent space model (LSM) for influence, as an alternative to existing methods for social influence and as a contribution to a growing literature on latent variable network models (e.g., Rastelli et al. (2016)). First, our model is the first latent variable network model for social influence and incorporates latent variables to accommodate network dependence. Also, our model is a fully Bayesian specification of network influence; it is a generative model that illuminates the data generating process. Furthermore, because our model employs the latent space positions from a latent space model (LSM; Hoff et al., 2002) to represent relative influence, our model accommodates differential social influence in multiple ways. The LSM for Influence also allows influence among those who do not share a tie; in particular, this model assumes that of the nodes to which one is not connected, nodes connected by a third party and social referents are the most influential. In addition, each node's level of influence can vary based on whom they are influencing. Our LSM for influence model can also be easily extended to other existing extensions of the LSM such as the hierarchical latent space model (HLSM) that utilizes multiple network data an extension as the HLSM for influence. Finally, our model utilizes outcome data measured at two time points, but allows for cross-sectional network data that are common in schools or other workplaces.

The remainder of our paper is organized as follows: we first introduce social network terminology and review some existing models for social influence for network data. We then review the LSM and formally introduce the LSM for Influence and HLSM for Influence. We explore the feasibility and utility of our proposed models through a series of simulation studies and then illustrate how our model could be used in practice with a real-world data application involving teacher networks. We conclude with some final thoughts and extensions.

2. Modeling Social Influence

Social network models represent relationship(s) between the network and node covariates. Some network models treat the network as the outcome variable and estimate the effects of node covariates on the probability or value of a network tie; these models are sometimes called social selection models. Social influence models, however, treat a node-level attribute as the dependent variable and estimate the effect of the network ties on this outcome.

As mentioned earlier, one issue with incorporating social networks into statistical models—either as outcome variables or predictor variables—is that the standard assumption of independent observations is violated. For example, if social influence does exist, then a given behavior of one individual is not independent of the other individuals. Individual i is likely influenced by the behaviors of their friend j. Furthermore, this dependency between i and j is in large part due to the dependencies among all of the connections in the network. The relationship between i and j does not exist in isolation; whether i and j have friends in common also affects the likelihood that they are friends. Another complication is that this dependence structure in a network is quite complex, varies by context, and therefore is very difficult to formally define. Therefore, typical GLMs are not applicable for modeling social influence.

We have mentioned two classes of social influence autocorrelation models, NAMs and ALAAMs. Autocorrelation models such as the original specification of NAMs and ALAAMs require a strong assumption that the network evolution process has reached a state of social equilibrium. Further, ALAAMS currently only accommodate binary node attributes.

Thus, a social influence model that relaxes these assumptions is called an exposure model (Valente, 2005); it is also referred to as a temporal version of the NAM (Leenders, 2002). In particular, the use of temporal outcome data allows a weaker assumption regarding social influence and does not assume any type of network stationarity. Further, Frank and Xu (2017) state that using longitudinal outcome data is an important step in estimating causal social influence with network data.[1] We present a temporal attribute version of a NAM as

$$Y^t = \beta_2 W Y^{t-1} + X\beta + \epsilon \tag{1}$$

where Y is the collection of nodal outcomes measured at two different times t and $t-1$, W is a weight matrix which accommodates the network structure A that affects Y, and β_2 is the estimated network effect. X is a collection of other covariates with coefficients β, and error term $\epsilon_i \sim N(0, \sigma^2)$.

Models of the form in Eq. (1) have been used to capture social contagion in numerous studies such as physician technology use (Zheng et al., 2010), market research (Iyengar et al., 2011), in studies of adolescent smoking status (Alexander et al., 2001) and employee telework decisions (Scott et al., 2012). In education, similar versions of Eq. (1) have been used to show the effects of peers on teaching practices. Frank et al. (2014) describe a multilevel version of Eq. (1) that has been used in a number of educational studies (Frank et al., 2004; 2013; Penuel et al., 2012; Sun et al., 2013). For example, Penuel et al. (2012) used this model to determine a positive effect of network ties on changes in writing instruction, and Spillane et al. (2018) extended Eq. (1) even further by embedding a weight matrix in a longitudinal growth model.

Depending on how W is specified, a number of theories of social influence can be accommodated. Leenders (2002) includes a number of possible weight matrices. When applied in the literature, W is often the row or column normalization of the adjacency matrix A, i.e., $W_{ij} = \frac{A_{ij}}{A_{i\cdot}}$, which is the ratio of the tie to the out-/in-degree. This specification assumes that nodes are equally

[1] Causal inference within a social network is an active area of research and is outside the scope of this manuscript. We direct interested readers to Shalizi and Thomas 2011; VanderWeele and An 2013.

influenced by those with whom they are connected, and although there are other specifications of W to estimate social contagion among nodes (include those that are not connected), other forms of W have not been studied in great detail and the row/column normalization specification for W appears to be the most common. Our proposed model can be thought as an extension to existing definitions of W.

Finally, any discussion about social influence models should include stochastic actor-oriented models (Snijders, 1996; Snijders et al., 2007, 2010; Steglich et al., 2010) which are part of a larger model fitting software called SIENA, Simulation Investigation for Empirical Network Analysis, (Snijders et al., 2008). Stochastic actor-oriented models are fundamentally different from the previously mentioned models (as well as our proposed model) in that they model the co-evolution of social selection and influence. Social selection is the phenomenon that nodes that are a priori similar are likely to form ties, whereas influence refers to the impacts of ties on node covariates.

While this model is often seen as the gold standard for estimating social influence, this model is not applicable to all contexts. First, SIENA requires longitudinal network data (at least two time points), and there are many contexts in which measuring networks repeatedly is not feasible. Instead, network data that represent ties over a period of tie are captured at a single time point. Cross-sectional network data are common in organizations where relatively fast turnover coupled with the slow evolution of workplace relationships hinders the collection of longitudinal network data. One such example is teacher networks whose membership can change quite drastically year to year but where collegial relationships among teachers such as advice-seeking or collaboration are slow to change once formed.

3. The LSM for Social Influence

We propose a new fully Bayesian model for social influence, the latent space model (LSM) for social influence, that accommodates the fact that individuals are influenced more by some nodes than by others. This model also allows individuals to be influenced by those with whom they do not share a tie, allowing references more social influence. Before we can provide explicit details and examples illustrating how our model captures influence, one must understand the relationship between the observed network and latent space positions.

3.1. Notation

A social network among n individuals or nodes can be represented by an $n \times n$ adjacency or tie matrix,

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}, \tag{2}$$

where A_{ij} is the value of the tie or edge from actor i to actor j. A_{ij} can theoretically have any scale of measurement but for the purposes of this work, we assume tie values are binary; ties are present or absent.

Social networks are often visually depicted using a sociogram or network plot (e.g., Fig. 1), where each individual i is represented by a node and each nonzero A_{ij} is represented by an edge between nodes i and j. Directed edges (asymmetric ties) are often depicted as arrows from the nominator to the nominee and undirected edges (symmetric ties) are shown as segments.

FIGURE 1.
Two social networks, depicting asymmetric advice-seeking behavior among teachers in two schools. Vertices represent individual teachers and arrows, or edges, point from advice-seeking teachers to advice-providing teachers

3.2. Latent Space Models

Latent space models (LSMs; Hoff et al., 2002) are social network models that predict network ties. LSMs are considered social selection models; they can incorporate covariates to predict network ties. To accommodate the fact that network ties are not independent, LSMs incorporate latent variables so that conditional on these latent variables network ties can be modeled as independent. Thus, LSMs are sometimes called latent variable network models (Rastelli et al., 2016) or conditionally independent models (Dabbs et al., 2019).

More specifically, the latent variables in LSMs are positions of each node in a low dimensional latent social space. The probability that any two individuals will share a tie is inversely related to the distance between the latent positions in this latent space. And then conditional on these latent positions, which account for the existing network structure, network ties are assumed to be independent.

One generally writes the latent space model for binary cross-sectional network data as

$$
\begin{aligned}
P(A|\alpha_0, Z) &= \prod_{i \neq j} P(A_{ij}|Z_i, Z_j, \alpha_0) \\
\text{logit}\, P[A_{ij} = 1] &= \alpha_0 - ||Z_i - Z_j|| \\
Z_i\ , Z_j &\overset{iid}{\sim} MVN_d(\vec{0}, \tau I) \\
\alpha_0 &\sim N(0, \sigma_\alpha^2)
\end{aligned} \tag{3}
$$

where A is the adjacency matrix, Z_i is the position of individual i in the latent space such that these positions exist in a low d-dimensional space and are sampled from a multivariate normal distribution with variance τI where I is the d-dimensional identity matrix. Note that α_0 is an intercept term to represent a baseline probability of a tie for nodes in the same position and can be estimated or fixed a priori.

We also note that we use Euclidean distance between the two latent positions for interpretability, but other distance metrics can certainly be used. In fact, Hoff (2005) and Hoff (2009) both discuss other LSM specifications.

Latent space models have been used to estimate covariate or selection effects, the association between actor attributes and the presence of a tie. After incorporating meaningful covariates, the latent space positions function more as a way to account for unobserved network tie dependencies than as an interpretable parameters.

Our proposed model, however, uses the estimated latent space positions and the pairwise distances between nodes as a representation of network structure as well as social reference. Therefore, to capture as much network structure as possible, we will not use any covariates in our LSMs.

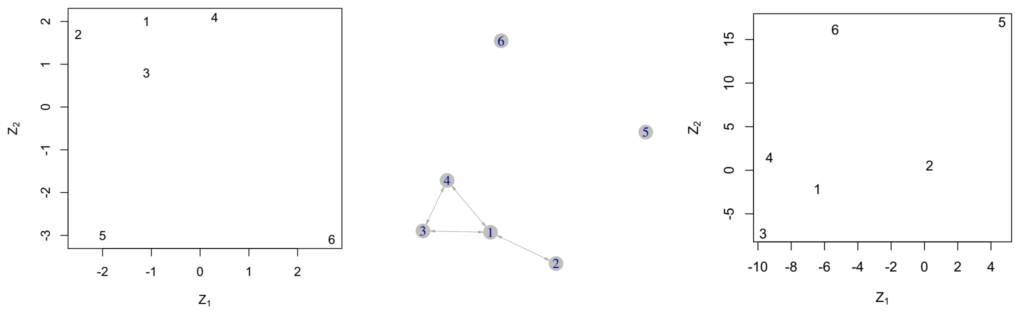

FIGURE 2.
Latent space positions (left) used to generate a social network (center). The LSM fitted with these network data estimates latent positions (right) such that nodes sharing ties tend to be closer than nodes that do not share ties

To better illustrate the relationship between latent space positions and the observed network, consider the following simple example with simulated data shown in Fig. 2. Given the known latent space positions shown on the left, one network generated from these positions is shown in the center plot. Note that this is merely one of many possible networks that can be generated from these positions. We then fit a LSM to this generated network, and the estimated latent positions are plotted on the right. Note that nodes that are close in the true latent space tend to have ties in the observed network; similarly, nodes that share a tie are estimated as being close in the latent space in the fitted model. Thus, we can think of the pairwise distances between latent space positions as a representation of the observed network structure.

Furthermore, central actors in the network tend to be central in the latent space. Node 1 in the network shown in Fig. 2 is a central node and although node 1 is not in the center of the oracle latent space, the fact that it is a central node in the network places it in the center of the estimated latent space positions. This is because the easiest way to position several nodes within a short distance to a specific node in a low dimensional space is to have that node be in the middle of the space. Similarly, nodes 5 and 6 are both isolated nodes and thus are estimated to be far from the other nodes and each other in the latent space.

3.3. The LSM for Influence

Therefore, the estimated latent positions from a LSM can be used to both represent the observed network, relatively small pairwise distances imply network ties, as well as which nodes are most central and likely to be influential to the other nodes in the network. In our simple example, nodes 5 and 6 do not have any direct connections to node 1, our most central node, but they may be influenced by such a strong actor in the network anyway. We note that the level of influence that node 1 exerts on 5 and 6 is probably less than nodes 2, 3, and, 4 who are directly connected to node 1.

We propose to use these latent space positions and the relative pairwise distances among each pair of nodes as a way of modeling temporal social influence; that is we propose a model similar to Eq. (1) in which W is a function of the latent space pairwise distances. Nodes that are close in the latent space are relatively more influential to node i than nodes that are far apart. At the same time, nodes that are particularly central in the network are likely to be placed in the center of the latent space and therefore are more influential than other nodes in that they are relatively close to many nodes.

The LSM for Influence can be thought of as the combination of a latent space model relating the observed network with latent space positions and a linear model relating the node-level outcomes with the distances among the latent space positions. Consider an adjacency matrix A for a binary network and node outcome variable Y such that Y_i is some attribute for node i. Then the LSM for Influence is defined as,

$$\begin{aligned}
A_{ij} &\sim \text{Bernoulli}\ (p_{ij}) \\
p_{ij} &= \frac{\exp(\alpha_0 - ||Z_i - Z_j||)}{1 + \exp(\alpha_0 - ||Z_i - Z_j||)} \\
Y^t &\sim N(X\beta + \beta_2 N(Z) Y^{t-1}, \sigma_Y^2) \\
Z_i &\overset{iid}{\sim} MVN_d(\vec{0}, \tau I) \\
\beta &\sim MVN(\mu_\beta, \Sigma_\beta) \\
\beta_2 &\sim N(\mu_2, \sigma_2^2)\ ,
\end{aligned} \tag{4}$$

where Z is the set of latent space positions whose pairwise distances are weighted so that $N(Z)$ is a matrix of weights as defined by the latent positions; $N(Z)_{ij}$ represents the relative weight of node j on i where individuals whose latent space positions are close have greater weight than individuals whose positions are far. Thus, $N(Z)_i Y^{t-1}$ is the weighted average of some outcome measured at time $t-1$ of $i's$ neighbors.

The regression coefficient β_2 is the overall effect of the network on Y at time t and our estimate of social influence. Note that in Eq. (4), X is a design matrix with coefficient β such that μ_β is a vector and Σ_β is a matrix. Typically X also includes some kind of baseline measure of the outcome variable Y^{t-1}. For simplicity, in future examples, we assume $X_i = (1, Y_i^{t-1})$ such that $\beta = (\beta_0, \beta_1)$.

We can also describe the LSM for Influence given in Eq. (4) by describing the generative model. Consider an example in which the outcome variable for an individual i, Y_i^t, is modeled as a function of that variable measured previously Y_i^{t-1}, and the network that existed during that time. Suppose that Z_i is the latent position for node i. Then for each node and a value of G selected a priori, we determine the G nearest neighbors. Based on the pairwise distance between i and each of its G nearest neighbors, node i is inversely influenced by each of its neighbors proportional to that neighbor's distance from i. We could also assume that ${Y_i}^t$ is predicted by the value of that covariate at a prior time, which is included in the design matrix Y^{t-1}.[2]

The generative model for this process for a single node i is given as

$$\begin{aligned}
{Y_i}^t &= \beta_0 + \beta_1 Y_i^{t-1} + \beta_2 \sum_{g \in \mathcal{G}_i} (\omega_g {Y_g}^{t-1}) + \epsilon_i \\
\omega_g &= \frac{(||Z_g - Z_i||)^{-1}}{\sum_{g \in \mathcal{G}_i} (||Z_g - Z_i||)^{-1}} \\
\text{logit}\, P(A_{ij} = 1) &= \alpha_0 - ||Z_i - Z_j||\ ,
\end{aligned} \tag{5}$$

where $\mathcal{G}_\rangle$ is the set of G nearest neighbors for i and ω_g is the weight of neighbor g. Note that β_2 is our parameter of interest, being the influence of the network on the outcome Y. Note that β_0 is the intercept of the regression and β_1 represents the effect of node i outcome at the previous time on the outcome at the current time.

To provide a context, consider our real-world example (Sect. 7). Let Y_i^t be a measure of a belief about teaching mathematics for teacher i. Then we determine the set of nearest neighbors

[2] Note: that it is possible for a node to be far from all neighbors and not influenced by others. We discuss modeling these phenomena in Sect. 7.

to node i which are likely those with whom they are connected or shares many connections, one of whom is an instructional coach; node i is influenced by them based on their relative distances to their. Then we expect that teacher i's beliefs about mathematics are a function of her previous beliefs Y_i^{t-1} as well as a weighted mean of the beliefs of her nearest neighbors, weights inversely based on their distances to her. This is true for every node in the network; they are influenced both based on their closest neighbors but also influenced differentially based on their distances to these neighbors.

Let us discuss how to select G. The value G can be any number up to $n - 1$, where n is the number of nodes in the network. How one selects G should be based on the context of the network data and is user-specified. For example, given adolescents in a school social media network, they are likely to be influenced by many more individuals than teachers in an advice-seeking network. Teachers may be influenced by a greater number of their colleagues regarding classroom management practices than they are regarding their beliefs about how students learn. Similarly, adolescent opinions about music or fashion may be influenced by more individuals than their beliefs about politics or religion.

We also note that in Eq. (4), we define the distance between latent positions as Euclidean distance. Despite our use of a symmetric distance, it is not true that any pair of nodes influences each other equally. Symmetric influence for a pair of nodes is only true if that pair of nodes are also equal distances to all other G nodes in the network, which is rare. Because each node is influenced based on the distances of their nearest G nodes, social influence varies by node and for each node. However, we acknowledge that some nodes may differ in terms of their personal receptiveness to influence, even after conditioning on these distances, and for these more extreme contexts we recommend using different distance metrics so that the latent space distance estimations adequately capture asymmetries in social influence.

To fit this model, we use a Markov Chain Monte Carlo (MCMC; Gelman et al., 2013) algorithm in which the latent positions (from a LSM) and the influence parameters are estimated simultaneously. Thus, along with all prior distributions, the full Bayesian specification of the LSM for Influence is given as

$$\begin{aligned} A_{ij} &\sim \text{Bernoulli}\ (p_{ij}) \\ \text{logit}\,(p_{ij}) &= \alpha_0 - ||Z_i - Z_j|| \\ Z_i\ , Z_j &\overset{iid}{\sim} MVN_d(\vec{0}, \tau I) \\ Y^t &\sim N(\beta_0 + \beta_1 Y^{t-1} + \beta_2 N(Z) Y^{t-1}, \sigma^2) \\ \beta_i &\sim N(\mu_0, \sigma_0^2)\ i = 0, 1, 2 \\ \sigma^2 &\sim Inv - Gamma(a, b) \\ \tau_Z &\sim Inv - Gamma(c, d)\ , \end{aligned} \tag{6}$$

where hyperparameters are chosen by the user (based on prior information).

The joint posterior distribution can be written as

$$\begin{aligned} & p(A, Z, \beta_0, \beta_1, \beta_2, Y^t, Y^{t-1}) \\ & \quad = p(A|Z, \tau) p(Y^t|Y^{t-1}, Z, \beta_0, \beta_1, \beta_2, \sigma^2) p(Z|\tau) p(p(\beta_0) p(\beta_1) p(\beta_2) p(\sigma^2)\ . \end{aligned}$$

Thus, we update β_0, β_1, β_2, and each of the latent positions Z using Metropolis-Hastings with random walk proposal distributions. We update σ^2 and τ using the standard Gibbs update for a linear regression model. Note that we do not include α_0 from Eq. (6) in our MCMC algorithm because α_0 is not estimated. There is an identifiability issue between α_0 and the latent space position variance τ; larger values of α_0 require larger pairwise distances between latent positions (i.e., larger values of τ). Thus, we fix α_0 when we fit this model and estimate τ; this also enables comparability across networks regarding latent space distances as well. We use $\alpha_0 = 2$ based on the size of our networks and the corresponding network density, but other values may be needed for networks that are much larger or that vary widely in density.

4. A Simple Example

To illustrate how the LSM for Influence models social influence, we present a simple example using simulated data. The purpose of this section is to show how the estimated latent positions naturally orient themselves around those who are most socially referent. More rigorous simulation studies follow in the next section.

We first simulate latent space positions Z for 20 nodes from a 2-dimensional multivariate normal distribution $MVN_2(0, 10I)$. Although this produces rather unrealistic networks in that they lack any kind of cluster structure, we will explore the impacts of cluster structure in Sect. 6. Once the latent space positions are generated, we use the weighted pairwise distance to estimate the effects of the network on Y^t. The rest of the data generating model is given as

$$\begin{aligned}
Y_i^{t-1} &\sim N(0,1) \\
Y_i{}^t &\sim N(\beta_1 Y_i{}^{t-1} + \beta_2 \sum_{g\in\mathcal{G}_i} (\omega_g Y_g{}^{t-1}), 0.3^2) \\
\omega_g &= \frac{(||Z_g - Z_i||)^{-1}}{\sum_{g\in\mathcal{G}_i} (||Z_g - Z_i||)^{-1}} \\
\text{logit}\,(p_{ij}) &= 1 - ||Z_i - Z_j|| \\
A_{ij} &\sim \text{Bernoulli}\,(p_{ij})\,,
\end{aligned} \tag{7}$$

where $\beta_1 = 2$ and $\beta_2 = 1$, and $\mathcal{G}_i$ is the set of $n - 1 = 19$ neighbors.

Equation 7 generates dependent variable Y^t, independent variable Y^{t-1} along with a single network A. We then fit a LSM for Influence given in Eq. 6. For model fitting, we use relatively weak prior distributions: $\beta_i \sim N(0, 100)$, $\sigma^2 \sim InvGamma(10, 20)$, and $\tau_Z \sim InvGamma(2, 20)$. We repeat this data generation and model fitting process 10 times.

Part of estimating latent positions is optimizing where nodes can exist in the low dimensional latent space such that pairwise distances appropriately reflect the presence or absence of ties. Thus, estimated latent positions generally place nodes with many ties in the center of the entire space. Figure 3 shows the estimated latent positions, as determined by the posterior mode, and the nodes shown as green triangles and red squares are those nodes with the highest out-degree and in-degree, respectively. Note that in general, the most central nodes appear in the center of the latent space.

Based on how the LS positions are situated, it is unsurprising that the most central nodes are also closest to many other nodes in the latent space. The LSM for Influence aims to utilize this positioning to model how nodes influence one another.

5. A Hierarchical Latent Space Model for Influence

In social science contexts, independent but similar networks are often collected at a single point in time. Networks of students or teachers across multiple schools (e.g., Spillane and Hop-

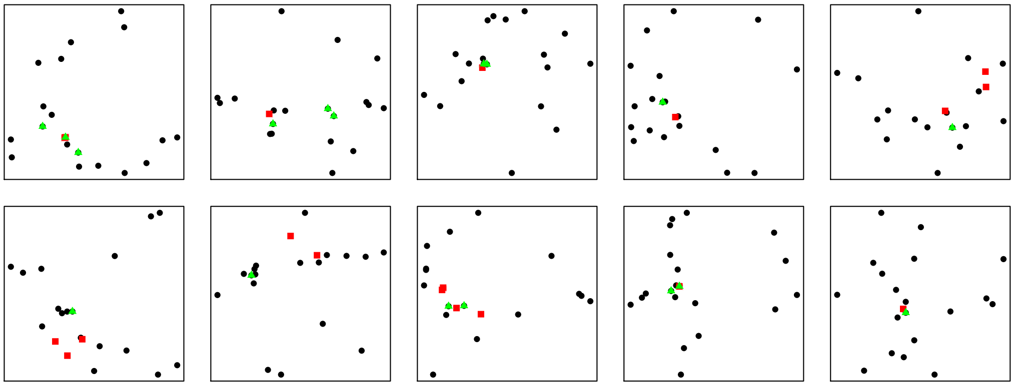

FIGURE 3.
Estimated latent space positions from LSM fit to simulated data shown in Eq. (7). Nodes with highest in- and out-degree are shown in red squares and green triangles, respectively, which suggests that socially referent nodes are generally estimated to be near the center of the clusters in the latent space (Color figure online)

kins (2013)) or networks of sharing across multiple communities (e.g., Koster (2018)) provide additional information about the variability that may exist across the same types of networks. Increasing the number of nodes is one way to decrease parameter variance, but increasing the number of networks is another. A hierarchical latent space model (HLSM; Sweet et al., 2013) accommodates these types of multiple networks. Thus, a HLSM for Influence can be written as

$$\begin{aligned}
A_{ijk} &\sim \text{Bernoulli}\,(p_{ijk}) \\
\text{logit}\,(p_{ijk}) &= \alpha_0 - ||Z_{ik} - Z_{jk}|| \\
Y_k^t &\sim N(\beta_{0k} + \beta_{1k} Y_k^{t-1} + \beta_{2k} N(Z_k) Y_k^{t-1}, \sigma^2) \\
Z_{ik}\ , Z_{jk} &\overset{iid}{\sim} MVN_d(\vec{0}, \tau I) \\
\beta_{ik} &\sim N(\mu_{\beta_i}, \sigma^2_{\beta_i})\ , i = 0, 1, 2 \\
\sigma^2 &\sim Inv - Gamma(c, d) \\
\tau_Z &\sim Inv - Gamma(e, f)\ ,
\end{aligned} \tag{8}$$

where β_{2k} represents the network effect for each network k. Similar to other multilevel models, the user can specify these models with fixed or random effects. We will use the HLSM for Influence along with a single network LSM for Influence to explore operating characteristics in Sect. 6 and use a HLSM for Influence with random effects in Sect. 7 with real-world data.

6. Simulation Studies to Further Explore the LSM for Influence

We now present a series of simulation studies to explore how the LSM for Influence can be used in practice. Specifically, we will use simulated data to determine the effects of the number of networks, number of nodes in the network (network size), network cluster structure and density, the number of nearest neighbors and the overall neighbor effect (β_2) on the posterior distribution of β_2 and resulting inference.

We will explore parameter coverage as measured by 95% equal-tailed credible intervals and will examine the proportion of replications in which these credible intervals do not include zero.

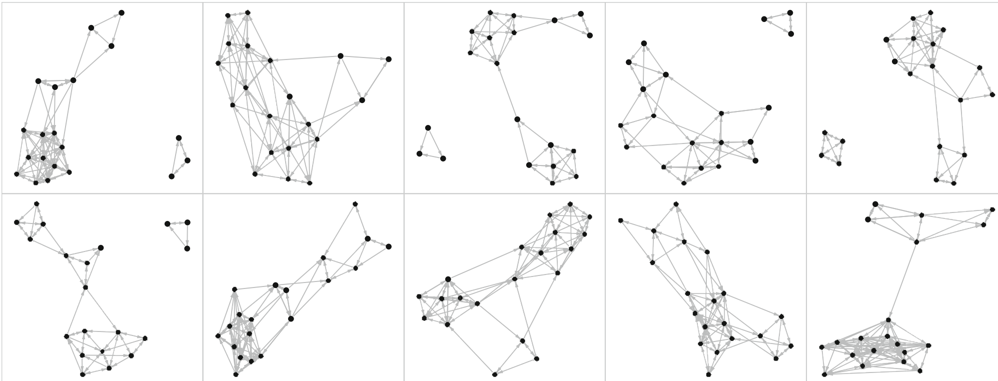

FIGURE 4.
An example of 10 networks generated from a data generating model with weak cluster structure to resemble the real-world networks introduced in Sect. 7

We will also examine the bias of the posterior means. Our goal is to determine under which conditions this model is most useful.

Further, our aim is for these simulations to be as realistic as possible, so we purposely generate our networks from a model with weak cluster structure. Keeping with the models presented so far, we continue to generate a network measured at one time point A, outcome Y^t and baseline measure Y^{t-1}. Assuming nodes belong to one of 6 clusters, the generating model for a single network is given as

$$\begin{aligned}
\mu_c &\sim MVN\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}10 & 0\\0 & 10\end{pmatrix}\right),\ c = 1,\ldots,6\\
Z_i &\sim MVN\left(\sum_i \mu_c I_{[i\in c]}, \begin{pmatrix}0.3 & 0\\0 & 0.3\end{pmatrix}\right)\\
Y_i^{t-1} &\sim N(0,1)\\
{Y_i}^t &\sim N(\beta_1 {Y_i}^{t-1} + \beta_2 \sum_{g\in\mathcal{G}_i}(\omega_g {Y_g}^{t-1}), 0.3^2)\\
\omega_g &= \frac{(||Z_g - Z_i||)^{-1}}{\sum_{g\in\mathcal{G}_i}(||Z_g - Z_i||)^{-1}}\\
\text{logit}\,(p_{ij}) &= 1 - ||Z_i - Z_j||\\
A_{ij} &\sim \text{Bernoulli}\,(p_{ij})\,,
\end{aligned} \tag{9}$$

where $\beta_1 = 2$ and $\beta_2 = 1$ for all replications except in Sect. 6.4 where β_2 varies across conditions. We also assume $G = n - 1$ for all simulations except those in 6.3 where we explore G. Further, the networks that result from this model tend to have some cluster structure. Figure 4 shows 10 networks generated from this model. Networks with similar structure are commonly seen among teacher networks.

For each simulation study condition presented below, we generated 100 replicated datasets and fit a LSM for Influence to each dataset using three MCMC chains of 10,000 samples each. We tossed the first 1000 draws and thinned the subsequent draws by 30. Although most replications reaching adequate mixing sooner, we selected 30 for ease of conducting the simulation. As a

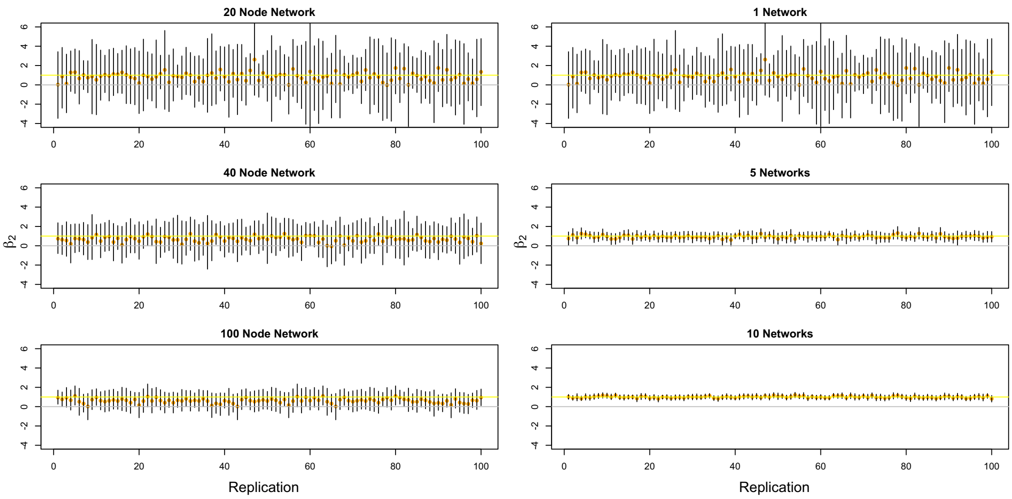

FIGURE 5.
Effect of network size (left) and the number of networks (right) on estimating the network influence effect; data from various numbers of networks were generated from the LSM for Influence and posterior means and 95% credible intervals are plotted

result, each posterior sample included 903 draws. Convergence was assessed visually for several replications and using the potential scale reduction factor (PSRF; Gelman and Rubin, 1992).[3] Processing times for model fitting for all three chains with the 20-node, 40-node and 100-node networks on a standard desktop computer were approximately 6 min, 18 min, and 2 h, respectively, for a single replication; the 5 network- and 10-network replications (with 20 nodes per network) each took 27 min and 57 min for a single replication.

6.1. Varying Network Size and the Number of Networks

It is well understood for multilevel models that standard errors decrease and power increases as both cluster size and the number of clusters increase (Snijders and Bosker, 2012), so we expect our posterior variance of β_2 to decrease as the network size increases and as the number of networks increase. In this simulation study, we explore the relative impacts of increasing both network size and the number of networks.

We first consider the size of the network: we simulated a network and influence data from Eq. (9) such that the number of individuals in the network varies from 20 to 40 to 100. Figure 5 (left) shows the posterior mean and 95% equal-tailed credible interval for β_2 for each replication and illustrates how increasing network size decreases the posterior variance. Parameter recovery, as defined as the proportion of replications such that the true value of β_2 is contained in the 95% equal-tailed credible interval, is 1 for all three network sizes.

The estimated posterior means (EAP) and modes (MAP) were so similar that we discuss EAPs moving forward. As the network size increases, the ability to precisely estimate the posterior mean becomes more difficult since the number of neighbors and thus the number of parameters to estimate also increases. We find that the EAP is generally less than the data generating value for β_2, and this negative bias of the EAP increases as network size increases; the mean bias for the EAP was -0.15, -0.31, and -0.39 for networks of size 20, 40, and 100, respectively.

We now consider the impact of 1, 5, and 10 network datasets. Let k index the network. We then generate 1, 5, and 10 networks with 20 nodes each from a multilevel LSM for Influence

[3]For example, the mean PSRF across all simulations in Sect. 6.1 was 1.002.

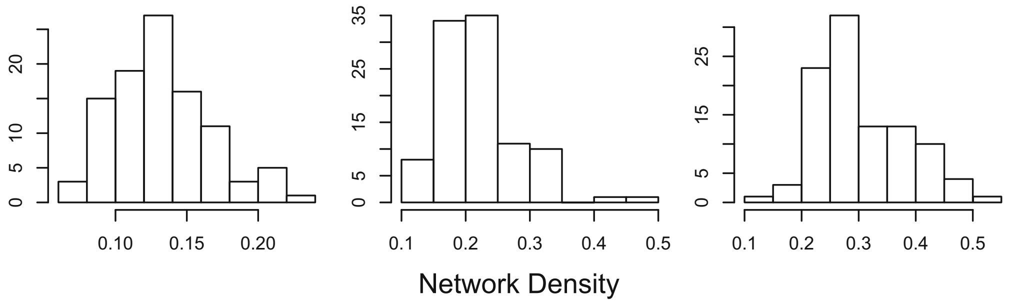

FIGURE 6.
Empirical distributions of network density for 100 networks generated from data generating models with different variances in latent space position clusters and nodes. The resulting networks are less dense (left) and more dense (right) than the original model used in the first simulation study (center)

model with the same parameters given in Eq. (9). The multiple network data are generated as isolated and independent replications; that is, nodes do not share ties across networks.

Figure 5 (right) shows the posterior mean estimates for β_2 along with the 95% credible intervals, and parameter recovery is 1 for all three simulation cells. In addition, increasing the number of networks has a large impact on the posterior variance, such that the posterior probability that $\beta_2 > 0$ also increases. We also found that increasing the number of networks decreases the bias of the EAP estimate as well; mean bias of the EAP was -0.38, -0.06, and -0.004 for the 1 network, 5 network, and 10 network conditions, respectively.

6.2. Varying Network Density and Cluster Structure

In our data generating model Eq. (9), we purposely created networks with densities similar to those seen in real-world teacher advice networks (see Sect. 7); in addition, these networks sometimes have some cluster structure so we also included clusters of latent space positions in our data generating model. In this study, we explore the effects of density and cluster structure.

For these simulations, we use 5 networks for our simulated data so that any differences in posterior variance across conditions are more obvious. The original data generating model in Eq. (9) resulted in a network with mean density estimated to be approximately 0.2. To simulate networks that are less dense, we increased the variance of the 6 clusters to be $30I_2$ as well as the variance of nodes around each cluster $0.5I_2$. To simulate networks that are more dense, we decreased these variances to be $5I_2$ and $0.2I_2$, respectively. Figure 6 shows the empirical distributions of network density for each of these conditions.

For networks of different density, parameter recovery rates remained at 1. Similarly, the posterior variance was also similar across conditions. In terms of inference, we inferred positive network influence effects approximately the same proportion of times across conditions as well. The only difference was in the bias of the EAP which becomes increasingly negative as density increased. The mean bias of the EAP was 0.003, -0.06, and -0.10 for the three density conditions (Table 1). This is unsurprising as the ability to estimate the posterior latent distances between nodes decreases as the density of the network increases. Further, increasing the number of ties also increases the number of nodes that have non-trivial influence on their neighbors.

We now turn to changing cluster structure in our network. Note that we did change cluster structure to increase and decrease density but in this simulation we focus solely on cluster structure while keeping network density to be as constant as possible (near 0.2). We continue to use a dataset of 5 networks each with 20 nodes, and we compare networks with 4 dispersed cluster centers with nodes tightly distributed around each center, our standard network of 6 cluster centers and

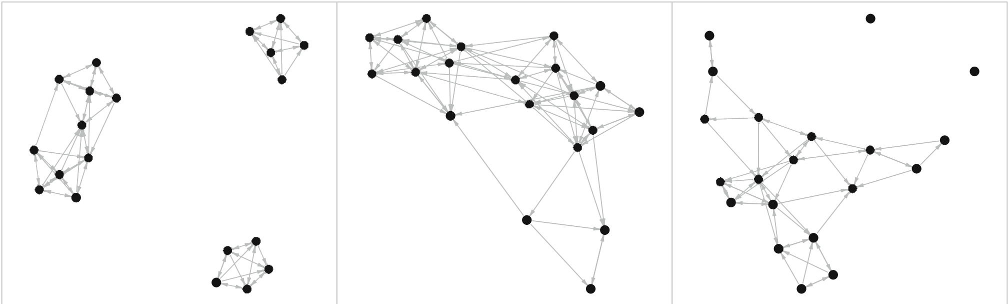

FIGURE 7.
Examples of networks generated with various cluster structure: cluster structure gets weaker from left to right as the variance (dispersion) of cluster centers decreases and the variance of LS positions around those centers increase

TABLE 1.
Mean bias of EAP across 100 replications of varying density and cluster structure

Mean density	Bias EAP	Cluster structure	Bias EAP
0.15	0.003	Tight clusters	0.02
0.23	– 0.06	Moderate clusters	– 0.06
0.32	– 0.10	Minimal clusters	– 0.17

nodes loosely distributed around them, and a network with 2 cluster centers nearby and nodes very loosely distributed around them. Formally, the data generating model for Z includes cluster center variances of $30I_2$, $10I_2$, and $5I_2$ and node variances of $0.05I_2$, $0.3I_2$, and $10I_2$ for the three conditions, respectively. Examples of the networks generated in each condition are given in Fig. 7. Note that the number of cluster centers does not map to the number of clusters, but the combination of cluster center variance and variance around each center together results in varying cluster structure.

With respect to parameter coverage, rates remained at 1. As with density, changing cluster structure had little impact on posterior variance. Cluster structure does impact the bias of the EAP (Table 1); it is more difficult to precisely estimate network influence effects in networks with minimal cluster structure. This is not surprising because in networks with tight cluster structure, nodes have ties to the same sets of nodes. We can estimate the relative weights of neighbors more precisely when each node is a part of a small group of neighbors that all have approximately equal weight. Estimating social influence is therefore more difficult with networks without cluster structure.

6.3. The Number of Nearest Neighbors

The purpose of this simulation is to examine the effects of the number of nearest neighbors G while also investigating one form of model misspecification. Recall G is the number of people influencing each node, and the number of influential nodes varies depending on the context; for example, an adolescent's opinions about movies are influenced by a wide variety of friends, whereas a teacher's pedagogical beliefs are likely influenced by only a handful of teachers.

We examine the effects of both the number of influential nodes used to generate data as well as the number of influential nodes used to estimate the LSM for Influence. Thus, this section also

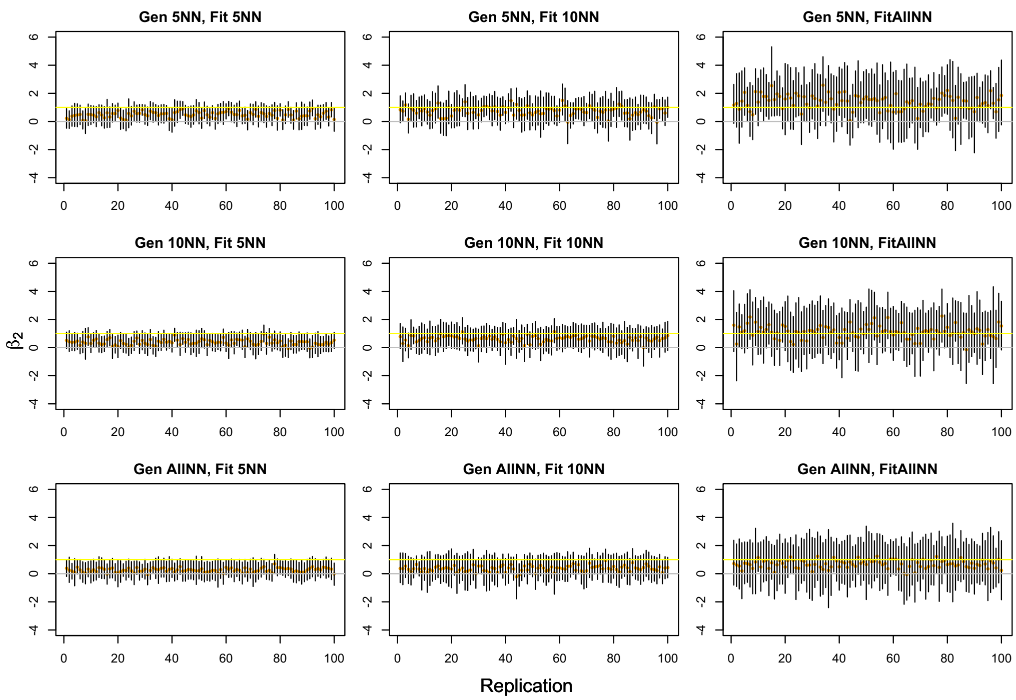

FIGURE 8.
Parameter estimates of the network influence effect when the model is misspecified and correctly specified based on the number of influential neighbors (NN). Note that networks were generated with 40 nodes so that All NN denotes 39 neighbors

explores model misspecification since we are fitting models that are purposely different from the models used to generate data.

We generate 1 network with 40 nodes from Eq. (9) using the same parameter values as in Sect. 6.1, but the data vary in terms of the number of nearest neighbors G used to generate $\sum_{g \in G}(\omega_g Y_g^{t-1})$. We let $G = (5, 10, 39)$ and simulate 100 replications for each dataset. We then fit each model with $G = (5, 10, 39)$.

Figure 8 shows the posterior distributions for β_2 for the correctly specified and two misspecified models for $G = 5, 10, 39$. Most of these findings are expected. We see that the posterior means of β_2 decreases as the number of neighbors specified is increasingly misspecified (Fig. 8, first column). Further, the posterior variance increases as the number of neighbors in the fitted models increases and that the best fitting models appear to be those that are correctly specified. What is most surprising is that posterior means appear to be overestimated when the model specified 39 neighbors and the data were generated with 5 (see Fig. 8 top right); that is biased EAP estimation is likely when G is too large. Thus, one should consider the context of the relationship so as to not grossly overestimate the number of influential nodes in a network; for example, adolescent behavior is likely influenced by a greater number of their peers than teacher pedagogical practices by other teachers.

We now examine parameter recovery and EAP bias more rigorously in Table 2. As expected, parameter recovery decreases as the misspecification of the model increases. EAP bias is also impacted by model misspecification and the EAP estimates are further from the truth the more misspecified the model is. Again, we note that the EAPs tend to be overestimated when the number

TABLE 2.
Parameter recovery of β_2 (left) and the mean bias of the posterior mean (right) for various combinations of nearest neighbors in the generative and fitted models

	Parameter recovery				Mean bias		
	Fitted model				Fitted model		
	5	10	39		5	10	39
Gen. model				Gen. model			
5	0.80	1	1	5	− 0.51	− 0.26	0.39
10	0.54	1	1	10	− 0.63	− 0.38	0.11
39	0.41	0.94	1	39	− 0.74	− 0.63	− 0.31

TABLE 3.
As expected, larger values of β_2 result in simulated data whose fitted models recover with decreasing rates but have increasing power estimates

β_2	P (95% CI includes 0)[a]	Mean bias of the posterior mean
0.2	0	− 0.02
0.6	0.83	− 0.03
1	0.98	− 0.06
1.4	1.00	− 0.02

[a] A 95% CI not including 0 is evidence of a positive influence effect

of neighbors in the fitted model is larger than the true number of influential neighbors and the bias increases as the inaccuracy in G increases.

6.4. *Varying the Effect of Network Influence*

Finally, we explore how the value of β_2 affects parameter recovery and the posterior distribution of β_2. Using Eq. (9) as our generating model, each dataset consisted of 5 networks of 20 nodes, and the generative model assumed all nodes were included in the set of nearest neighbors. We generated 100 datasets using the various values for $\beta_2 = (0.2, 0.6, 1, 1.4)$ where $\beta_1 = 2$ and all other parameters are the same as in Sect. 6.1.

Parameter recovery rates were 1 for all conditions, and unsurprisingly, as the value of β_2 increased, the proportion of times we would infer a positive network influence also increased. Table 3 shows the proportion of times one would conclude a positive influence effect as well as the mean bias (across replications) of the posterior mean. Bias of the posterior mean is near zero for all conditions, suggesting both little bias and little impact of the value of β_2 on EAP bias.

6.5. *A Comparison of NAM and LSM for Influence*

Our final simulation study compares a NAM and LSM for Influence. There are four conditions in this study: data generated from a NAM and fit with a NAM, data generated from a NAM and fit with a LSM for influence, data generated from LSM for Influence and fit with a NAM and data generated from LSM for Influence and fit with a NAM.

We generated data from the temporal NAM given in Eq. (1) with 5 networks each with 20 nodes such that $\beta_2 = 1$. We assume nodes were equally influenced by every node to whom they sent a tie; thus, W_{ij} in Eq. (1) is $\frac{A_{ij}}{\sum(A_i)}$. The data generating model for the LSM for Influence data

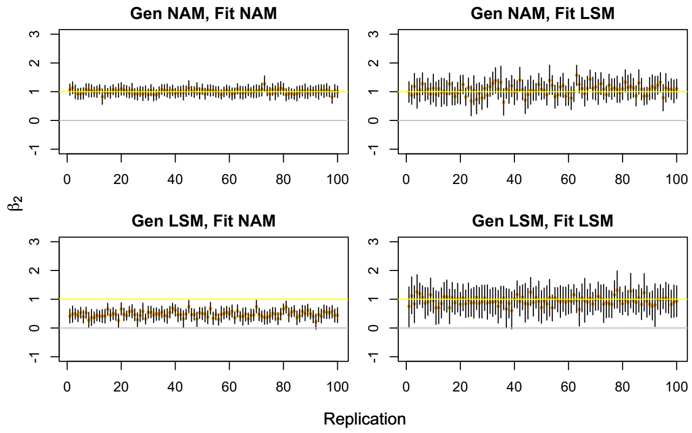

FIGURE 9.
Posterior distributions of LSM and NAM fit to data generated from a NAM and LSM suggest that the ability to detect influence effects decreases when the model is misspecified

generating model was the same as that used in Sect. 6.1, Eq. (9). The fitted model for the NAM used the same definition of W as the data generating model.

Figure 9 shows the posterior distributions for the network effect (β_2) for each of the four conditions. Parameter recover rates were 1 when the data generating model and fitted model were the same. In addition, we examined the bias of the EAP estimate and found that the mean bias of the EAP estimate was 0.004 and -0.06 for the correctly specified NAM and LSM models, respectively (Fig. 9 top left and bottom right). When the models were misspecified, both models resulted in the same inference, positive effects of network influence. The difference is that the mean EAP bias was much larger for data generating from a LSM and fit with a NAM (-0.52) than data generated from a NAM and fit with a LSM (0.06).

The results of this simulation study suggest that the LSM for Influence appears to recover network influence effects fairly well regardless of the true data generating model. That is, there appears to be greater error in fitting a NAM to the data if the data did not actually come from a NAM. This suggests that the LSM for Influence could be more robust to model misspecification.

6.6. *Summary of Simulations*

We conducted five small simulation studies to explore some of the operating characteristics of the LSM for Influence. We examined the effects of the number of networks and network size, the density and cluster structure of the networks, the number of nearest (influential) neighbors, the level of network influence, and model misspecification.

Parameter recovery as measured by 95% credible interval coverage was 100% in all well-specified models, and the posterior probability of a positive influence effect increased as the number of networks increased and the true influence effect increased. We also found that increasing the size of the network (and nearest neighbors) also increased the posterior variance for the influence effect. This was also true in networks with more ties (higher density) as well as networks with distances more difficult to estimate (no cluster structure).

With respect to misspecified models, we found that misspecifying the model also impacted the EAP estimates and to a small extent parameter recovery. EAPs tended to be underestimated most when the data generating model was a LSM for Influence and the fitted model was NAM. EAPs were largely similar when the fitted model was a LSM for Influence, regardless of the true model. Thus, the LSM for Influence appears to be more robust to misspecification of the true model.

Similarly, misspecifying the number of neighbors also impacted parameter recovery and EAP estimation. When the true number of influence neighbors was underestimated, the influence parameter was underestimated (more so than usual), and when the true number of influential neighbors was overestimated, there is positive bias (on average) of the point estimate for the network effect. This is particularly important moving forward as researchers will need to give thought as to how many individuals are influential in their networks.

Further, some readers may be concerned with the conditions that resulted in positive bias of the posterior mean and whether that translates to incorrectly detecting effects that are not present. Therefore, we examined the situations in which our posterior mean had the largest bias, data generated with $G = 5$ were fitted with $G = 39$ (see Sect. 6.3). We conducted a small simulation study to determine whether effects are incorrectly detected under these conditions when the true influence effect is null. Data generation and model fitting were identical to that in Sect. 6.3 and over 20 replications, we found that the mean EAP bias was 0.03. This suggests that we are unlikely to detect network influence when none exists, even when the number of nearest neighbors is grossly overestimated.

7. An Application to Advice-Seeking Networks

To illustrate an example using real network data, we present an analysis in which we use the LSM for Influence to estimate the impact of teacher advice-seeking networks on beliefs about mathematics. Education researchers are often interested in the mechanisms through which teaching quality improves, and changes in personal beliefs about instruction are often the precursors to improving instruction.

We use data collected among elementary school staff in a mid-sized Midwestern school district that we call Auburn Park. It is a suburban school district serving approximately 5800 students in 14 elementary schools. Instructional and administrative staff members were surveyed in Spring 2012 and Spring 2013. This survey included questions about teacher beliefs regarding mathematics and also asked respondents to list staff members to whom they turn for advice and information regarding curriculum and instruction. For each person selected, respondents also included the subject surrounding that relationship; that is, respondents selected whether the tie was for literacy, mathematics, science, other, or any combination of the above.

For our example, we focus on mathematics beliefs in 2012 and 2013 and mathematics advice-seeking ties collected in 2013. We hypothesize that teacher beliefs about mathematics in 2013 are a function of their beliefs in 2012 as well as other teacher beliefs in their networks. In addition, we use the 2013 advice-seeking networks as a measure of the interactions that took place between 2012–2013 since teachers were asked to nominate others based on their interactions over the past year.

We measured beliefs about mathematics using responses from 11 survey items; see Table 4 which includes two beliefs constructs. One construct focuses on how students should learn math and the other is how teachers should facilitate mathematics instruction. For example, teachers who agree with the statement, *Recall of number facts should precede the development of an understanding of the related operation*, are likely to feel strongly about the order of other procedural skills over principled knowledge. Spillane et al. (2018) call this construct reform-oriented beliefs, but without their theoretical foundation, we will refer to this construct as *principled beliefs*.

The second construct is perhaps more subtle and includes statements such as *Teachers should allow students to find their own solutions to math problems even if they are inefficient*. This belief

TABLE 4.
Teachers in Auburn Park were surveyed about their beliefs about teaching mathematics. They were asked to what extent they agreed with the statements below, and their data suggest two constructs: how students should learn math (*principled beliefs*) and how teacher should facilitate mathematics instruction (*student-centered beliefs*)

Item	Statement
1	Recall of number facts should precede the development of an understanding of the related operation
2	Students should master computational procedures before they are expected to understand how those procedures work
3	Time should be spent practicing computational procedures before students are expected to understand the procedures
4	Students should not solve simple word problems until they have mastered some number facts
5	Time should be spent practicing computational procedures before students spend much time solving problems
6	Students will not understand an operation until they have mastered some of the relevant number facts
7	Teachers should encourage students to find their own solutions to math problems even if they are inefficient
8	Teachers should allow students to figure out their own ways to solve simple word problems
9	The goals of instruction in mathematics are best achieved when students find their own methods for solving problems
10	Most students can figure out ways to solve many mathematics problems without any adult help
11	Mathematics should be presented to children in such a way that they can discover relationships for themselves

generally involves teachers asking students guiding questions instead of simply providing information, so that the students construct their own knowledge through exploration and experimentation. We denote this construct as *student-centered beliefs*.

The social network data used in this example are mathematics advice-seeking networks which are both binary and directed. Plots of each of the 14 school networks are shown in Fig. 10. Not only do schools vary in size from 13 to 30, but network structure also varies. The proportion of observed ties ranges from 0.06 to 0.13 which suggests fairly sparse networks. Note that we only include ties to teachers within the same school because one assumption of the LSM for Influence is that individuals are only influenced by those individuals in their own network, which we defined as the school. A very small percentage of ties did occur between schools, but very few of these ties were between teachers.

We fit a HLSM for Influence for the 14 schools similar to the model introduced in Sect. 6.1, and to fully explore the ways in which mathematics advice-seeking ties are related to end-of-year beliefs about mathematics, we include random slopes so that network effects can vary across schools. We also allow intercepts to vary across schools.

Define Y_k^t as the math belief scores for school k in 2013. We model each score as a function of Y_k^{t-1}, the math belief in 2012 along with the math beliefs of each node's neighbors during the 2012–2013 school year. The random effects model is given as

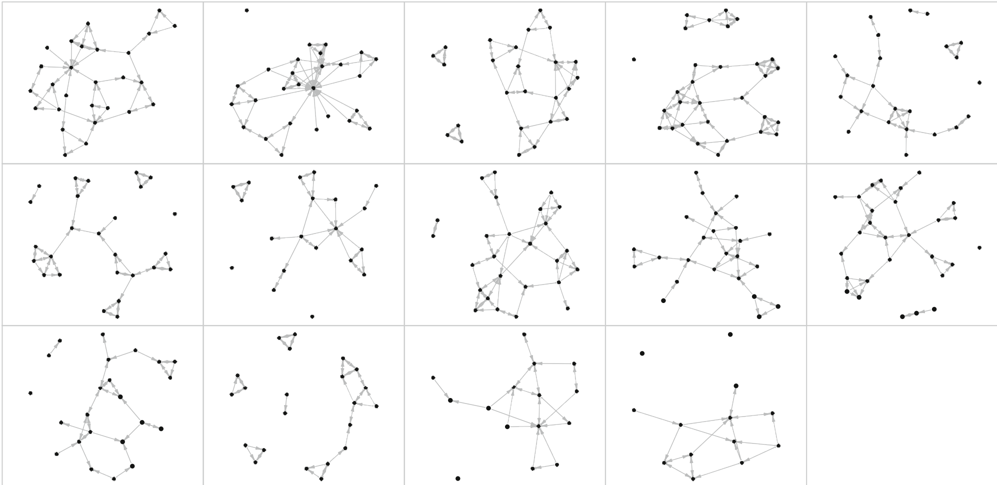

FIGURE 10.
Mathematics advice-seeking networks among teachers in 14 schools in Auburn Park illustrate the variability in network structure across schools

$$
\begin{aligned}
A_{ijk} &\sim \text{Bernoulli}\,(p_{ijk}) \\
\text{logit}\,(p_{ijk}) &= \alpha_0 - ||Z_{ik} - Z_{jk}|| \\
Z_{ik}\,,\, Z_{jk} &\overset{iid}{\sim} MVN(\vec{0}, \tau I_2) \\
Y_k^t &\sim N(\beta_{0k} + \beta_{1k} Y_k^{t-1} + \beta_{2k} N(Z_k) Y_k^{t-1} X_k, \sigma^2) \\
\beta_{0k} &\sim N(\mu_0, \sigma_0^2) \\
\beta_{1k} &\sim N(\mu_1, \sigma_1^2) \\
\beta_{2k} &\sim N(\mu_2, \sigma_2^2) \\
\mu_i &\sim N(0, 10),\;\; i = 0, 1, 2 \\
\tau &\sim Inv - Gamma(2, 20)\;, \\
\sigma_i^2 &\sim Inv - Gamma(10, 20)\;,
\end{aligned} \tag{10}
$$

In addition, Fig. 10 shows that most networks include isolated nodes and we now augment our model so that nodes that are relatively far from every other nodes are not susceptible to social influence. Our model now includes a variable $X_k = (X_{1k}, X_{2k}, \ldots, X_{nk})$ that is a binary indicator such that $X_{ik} = 0$ if the distance from node i to their closest neighbor is in the top 25% of all distances. For example, suppose node i's nearest neighbor is relatively far from them. Then their outcome at time t, Y_{ik}^t is a function of only their outcome at the previous time point, Y_{ik}^{t-1}. Finally, based on our simulation studies and the fact that teachers are likely not influenced by every other person in the building, we set the number of influential neighbors to be 10. Further, the original network survey allowed teachers to nominate up to 12 names, and none of the teachers nominated more than 10 teachers.

Because there are two belief constructs, we fit two separate HLSM for Influence models to predict 2013 *principled beliefs* and *student-centered beliefs*. The random effects model is also estimated using a MCMC algorithm similar the one presented in Sect. 3.3; the differences in estimation are simply that the parameters at the higher level in the hierarchy (e.g., μ_i, σ_i^2) are now estimated using Gibbs updates.

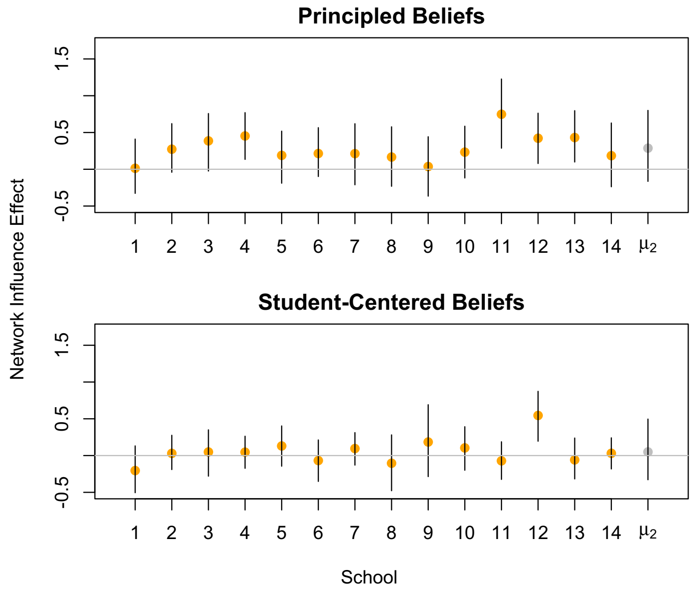

FIGURE 11.
Posterior mode and corresponding 95% equal-tailed credible intervals of β_{2k} and μ_2 from Eq. (10) indicate evidence of a positive effect of the 10 nearest neighbors (as defined by the latent space positions and the observed network) on principled math beliefs but there is little evidence of influence effects on student-centered math beliefs

We fit each model using three MCMC chains of 100k; 1500 burn-in and thinning by 500 resulted in posterior samples of 594. Summaries for the posterior distribution of the network effect β_{2k} for each of the 14 schools in each HLSM for Influence model are given in Fig. 11, which shows expected a posteriori estimates along with the 95% equal-tailed credible intervals for principled beliefs (top) and student-centered beliefs (bottom). We also include the level 2 influence estimate, μ_2.

In four schools, there is strong evidence that the network effect is positive on principled math beliefs. In two other schools, there is some evidence of a positive network effect given posterior probabilities of $\beta_2 > 0$ above 0.9. The overall mean effect (μ_2) also has a posterior distribution such that $P(\mu_2) > 0$ is 0.89, mild evidence of a network effect across all 14 schools. There is less evidence of an effect of the network on student-centered beliefs although there is a clear positive network effect in one school. Other schools appear to have null network effects, but none of the schools have negative effects.

We note one limitation of this application: the HLSM for Influence was fit to directed network data, but the model itself uses a distance metric that results in symmetric distances between pairs of nodes to construct the latent positions. It is possible that alternative specifications of the latent space model may be more appropriate for some contexts, including advice-seeking ties. This should not entirely discount using directed ties for other types of relationships that are symmetric such as friendship nominations.

8. Discussion

Social network data represent the interactions and relationships among a group of individuals. Incorporating network data into models of individual outcomes allows researchers to account for social influence since changes in behavior rarely occur in isolation. Modeling social influence not only enables researchers to understand which outcomes might be influenced but also allows researchers to estimate the extent to which such outcomes are influenced by social networks.

We introduced new social influence models, the LSM for Influence and its multilevel version the HLSM for Influence, which utilize a latent variable representation of a network to estimate how nodes influence one another. The motivation for this model is the idea that individuals are not only differentially influenced by those with whom they interact but they can also be influenced by those with whom they do not share a direct connection; these influence processes can also act simultaneously. In the same vein that people shop at certain stores based on celebrity endorsement, teachers may alter teaching practices based on changes a model teacher implements in their classroom.

To better understand some of the operating characteristics of our proposed model, we conducted a series of simulation studies to explore the feasibility and utility of our model. Parameter coverage was at or near 100% in correctly specified models and decreased as misspecification increased. The ability to detect network influence increased as expected: when the network size or the number of networks increased, when the true effect increased and when cluster structure increased. We found that model misspecification impacted our model and found that our models seem to be more robust to misspecification than autocorrelation models. Regarding misspecification of the number of nearest neighbors, we found that greatly overestimating the number of neighbors tends to increase the bias in the influence parameters; not necessarily finding effects that do not exist, but inflating effects that do exist.

Finally, we presented a real-world analysis where we fit a HLSM for Influence to estimate the effects of mathematics advice-seeking networks on changes in teacher beliefs about mathematics. We found some evidence that both beliefs constructs were affected by the network, and in particular principled beliefs about mathematics changed to be more similar to the beliefs of the individuals closest to them.

While this paper has provided evidence that the LSM for Influence is both feasible and has the potential to estimate influence effects in networks with social referents as well as across multiple networks, there are several areas for future work. First, additional work on how G is selected is needed. While we encourage users to consider the context of their network to determine G, we believe some of the work on goodness of fit may be useful in helping users determine the number of nearest neighbors. For example, posterior predictive checks may be useful in determining a maximum value to use. A second area is to consider other specifications of a latent space model that may better accommodate directed ties; similarly, a model that accommodates valued or ordinal ties is a natural extension.

Other extensions to this work include how we model social influence; there are other latent variable approaches for modeling network data and these parameters could also be used to model influence. For example, a mixed membership stochastic block model (Airoldi et al., 2008) might be used to create soft clusters and influence could be specified based on relative cluster membership. Another extension includes the work proposed by He and Hoff (2017), which combines existing work on longitudinal latent space models (Sewell and Chen, 2015; Sarkar and Moore, 2005) to develop longitudinal latent space models that can model both selection and influence. Finally, another area of work is on interventions on networks; there is some literature suggesting that exogenous shocks to a system (like a social network) change the way individuals interact (Weinbaum et al., 2008; Spillane and Hopkins, 2013). Coupled with models for experimental interventions (Sweet et al., 2013), a natural next step is to incorporate influence models to model how a social network could operate as a mediator in an intervention (e.g., Sweet, 2019).

References

Airoldi, E., Blei, D., Fienberg, S., & Xing, E. (2008). Mixed membership stochastic blockmodels. *The Journal of Machine Learning Research*, *9*, 1981–2014.

Alexander, C., Piazza, M., Mekos, D., & Valente, T. (2001). Peers, schools, and adolescent cigarette smoking. *Journal of Adolescent Health*, *29*, 22–30.

Anderson, C. J., Wasserman, S., & Crouch, B. (1999). A p* primer: Logit models for social networks. *Social Networks*, *21*, 37–66.

Christakis, N., & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks. *PloS ONE*, *5*, e12948.

Dabbs, B., Adhikari, S., & Sweet, T. (2019). Conditionally independent dyads (CID) network models: An integrative approach to statistical social network analysis. **(Under review)**.

Daragonova, G., & Robins, G. (2013). Autologistic actor attribute models. In D. Lusher, J. Koskinen, & G . L. Robins (Eds.), *Exponential random graph models for social networks: Theories, methods and applications* (pp. 102–114). Cambridge, MA: Cambridge University Press.

Doreian, P. (1989). Network autocorrelation models: Problems and prospects. In D. Griffith (Ed.), *Spatial statistics: Past, present, future* (pp. 369–393). Ann Arboro, MI: Michigan Document Services.

Eckhardt, A., Laumer, S., & Weitzel, T. (2009). Who influences whom? Analyzing workplace referents' social influence on IT adoption and non-adoption. *Journal of Information Technology*, *24*, 11–24.

Frank, K. A., Lo, Y.-J., & Sun, M. (2014). Social network analysis of the influences of educational reforms on teachers practices and interactions. *Zeitschrift für Erziehungswissenschaft*, *17*, 117–134.

Frank, K. A., Muller, C., & Mueller, A. S. (2013). The embeddedness of adolescent friendship nominations: The formation of social capital in emergent network structures. *American Journal of Sociology*, *119*, 216–253.

Frank, K. A., & Xu, R. (Forthcoming). Causal inference in network analysis: Navigating dependencies and alternative explanations to create good science. In *Oxford handbook of social network analysis*.

Frank, K. A., Zhao, Y., & Borman, K. (2004). Social capital and the diffusion of innovations within organizations: The case of computer technology in schools. *Sociology of Education*, *77*, 148–171.

Fujimoto, K., Wang, P., & Valente, T. (2013). The decomposed affiliation exposure model: A network approach to segregating peer influences from crowds and organized sports. *Network Science*, *1*, 154–169.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: Taylor & Francis.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*, 457–472.

Gupta, S., Anderson, R. M., & May, R. M. (1989). Networks of sexual contacts: Implications for the pattern of spread of HIV. *Aids*, *3*, 807–818.

He, Y., & Hoff, P. D. (2017). Multiplicative coevolution regression models for longitudinal networks and nodal attributes. arXiv preprint arXiv:1712.02497.

Hoff, P. D. (2005). Bilinear mixed-effects models for dyadic data. *Journal of the American Statistical Association*, *100*, 286–295.

Hoff, P. D. (2009). Multiplicative latent factor models for description and prediction of social networks. *Computational & Mathematical Organization Theory*, *15*, 261–272.

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

Iyengar, R., Van den Bulte, C., & Valente, T. W. (2011). Opinion leadership and social contagion in new product diffusion. *Marketing Science*, *30*, 195–212.

Kashima, Y., Wilson, S., Lusher, D., Pearson, L., & Pearson, C. (2013). The acquisition of perceived descriptive norms as social category learning in social networks. *Social Networks*, *35*, 711–719.

Koster, J. (2018). Family ties: The multilevel effects of households and kinship on the networks of individuals. *Royal Society Open Science*, *5*, 172159.

Leenders, R. T. A. (2002). Modeling social influence through network autocorrelation: Constructing the weight matrix. *Social Networks*, *24*, 21–47.

Paluck, E. L., & Shepherd, H. (2012). The salience of social referents: A field experiment of collective norms and harassment behavior in a school social network. *Journal of Personality and Social Psychology*, *103*, 899–915.

Penuel, W. R., Sun, M., Frank, K. A., & Gallagher, H. A. (2012). Using social network analysis to study how collegial interactions can augment teacher learning from external professional development. *American Journal of Education*, *119*, 103–136.

Rastelli, R., Friel, N., & Raftery, A. E. (2016). Properties of latent variable network models. *Network Science*, *4*, 407–432.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, *29*, 173–191. **(Special section: Advances in exponential random graph (p*) models)**.

Robins, G. L., Pattison, P. E., & Elliott, P. (2001). Network models for social influence processes. *Psychometrika*, *66*, 161–189.

Sarkar, P., & Moore, A. W. (2005). Dynamic social network analysis using latent space models. *ACM SIGKDD Explorations Newsletter*, *7*, 31–40.

Scott, D. M., Dam, I., Páez, A., & Wilton, R. D. (2012). Investigating the effects of social influence on the choice to telework. *Environment and Planning A*, *44*, 1016–1031.

Sewell, D. K., & Chen, Y. (2015). Latent space models for dynamic networks. *Journal of the American Statistical Association*, *110*, 1646–1657.

Shalizi, C., & Thomas, A. (2011). Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research*, *40*, 211–239.

Shepherd, H., & Paluck, E. L. (2015). Stopping the drama: Gendered influence in a network field experiment. *Social Psychology Quarterly*, *78*, 173–193.

Snijders, T. (1996). Stochastic actor-oriented models for network change. *The Journal of Mathematical Sociology*, *21*, 149–172.

Snijders, T. A., & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Snijders, T. A., Steglich, C. E., Schweinberger, M., & Huisman, M. (2008). *Manual for SIENA version 3.2*. Groningen: University of Groningen: ICS, Department of Sociology.

Snijders, T. A., van de Bunt, G. G., & Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, *32*, 44–60. **(Dynamics of social networks)**.

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

Spillane, J. P., & Hopkins, M. (2013). Organizing for instruction in education systems and school organizations: How the subject matters. *Journal of Curriculum Studies*, *45*, 721–747.

Spillane, J. P., Hopkins, M., & Sweet, T. M. (2018). School district educational infrastructure and change at scale: Teacher peer interactions and their beliefs about mathematics instruction. *American Educational Research Journal*, *55*, 532–571. https://doi.org/10.3102/0002831217743928.

Steglich, C., Sinclair, P., Holliday, J., & Moore, L. (2012). Actor-based analysis of peer influence in A Stop Smoking In Schools Trial (ASSIST). *Social Networks*, *34*, 359–369.

Steglich, C., Snijders, T. A., & Pearson, M. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological Methodology*, *40*, 329–393.

Sun, M., Penuel, W. R., Frank, K. A., Gallagher, H. A., & Youngs, P. (2013). Shaping professional development to promote the diffusion of instructional expertise among teachers. *Educational Evaluation and Policy Analysis*, *35*, 344–369.

Sweet, T. M. (2019). Modeling social networks as mediators: A mixed membership stochastic blockmodel for mediation. *Journal of Educational and Behavioral Statistics*, *44*, 210–240.

Sweet, T. M., Thomas, A. C., & Junker, B. W. (2013). Hierarchical network models for education research: Hierarchical latent space models. *Journal of Educational and Behavioral Statistics*, *38*, 295–318.

Valente, T. W. (2005). Network models and methods for studying the diffusion of innovations. In P. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 98–116). New York, NY: Cambridge University Press.

Van der Weele, T. J., & An, W. (2013). Social networks and causal inference. In S. Morgan (Ed.), *Handbook of causal analysis for social research* (pp. 353–374). New York, NY: Springer.

Watts, D. J., & Dodds, P. S. (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research*, *24*, 441–458.

Weinbaum, E., Cole, R., Weiss, M., & Supovitz, J. (2008). Going with the flow: Communication and reform in high schools. In J. Supovitz & E. Weinbaum (Eds.), *The implementation gap: Understanding reform in high schools* (pp. 68–102). New York, NY: Teachers College Press.

Zheng, K., Padman, R., Krackhardt, D., Johnson, M. P., & Diamond, H. S. (2010). Social networks and physician adoption of electronic health records: Insights from an empirical study. *Journal of the American Medical Informatics Association*, *17*, 328–336.

Manuscript Received: 18 DEC 2017
Final Version Received: 7 FEB 2020