Network Science **7** (2): 160–179, 2019
doi:10.1017/nws.2019.1

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

Latent space models for network perception data

Daniel K. Sewell

Department of Biostatistics, University of Iowa, Iowa City, IA 52242, USA
Corresponding author. Email: daniel-sewell@uiowa.edu

Action Editor: Stanley Wasserman

Abstract

Social networks, wherein the edges represent nonbehavioral relations such as friendship, power, and influence, can be difficult to measure and model. A powerful tool to address this is cognitive social structures (Krackhardt, D. (1987). Cognitive social structures. *Social Networks*, *9*(2), 109–134.), where the perception of the entire network is elicited from each actor. We provide a formal statistical framework to analyze informants' perceptions of the network, implementing a latent space network model that can estimate, e.g., homophilic effects while accounting for informant error. Our model allows researchers to better understand why respondents' perceptions differ. We also describe how to construct a meaningful single aggregated network that ameliorates potential respondent error. The proposed method provides a visualization method, an estimate of the informants' biases and variances, and we describe a method for sidestepping forced-choice designs.

Keywords: cognitive social structures, latent space network model, network estimation, social network analysis, visualization

Author ORCID. Daniel K. Sewell, https://orcid.org/0000-0002-9238-4026

1. Introduction

Researchers studying social networks have typically fallen along a spectrum between two camps described by Koskinen (2004) as relativist and essentialist. The former treats a social network as a cognitive construct existing only in the minds of the network actors, whereas the latter assumes there is some interpersonal reality acting as a stimulus for the informants. Researchers of both camps have used cognitive social structures (CSS) to better answer a wide range of research questions. CSS, introduced in Krackhardt (1987), are obtained by asking each network actor to enumerate all the relationships in the network, thereby providing their perceptions of the network. Thus, while in typical social network analyses one deals with a single $n \times n$ sociomatrix, CSS provides n such sociomatrices, where n is the number of actors in the network. Analyzing such data can be difficult, as these involve $\mathcal{O}(n^3)$ observations. Since the introduction of CSS, many researchers have used ad-hoc approaches or performed subsequent network analyses on aggregation techniques introduced by Krackhardt. However, there have certainly been some important statistical methodological developments in this area worth noting.

Leaning toward the relativist side of the spectrum is work done by Bond Jr. *et al.* (1997). These authors developed an extension of the social relations model of Kenny (1994) and Warner *et al.* (1979), using an ANOVA-like decomposition to explain how perceptions vary. Swartz *et al.* (2015) continued this work, putting it within a Bayesian framework and providing researchers more sophisticated analytic tools to compare actors' self-perceptions in relation to how others see them.

Much more work has been done from the essentialist perspective. Batchelder & Romney (1988, 1989) developed their cultural consensus analysis methodology; although this work was not designed specifically for CSS data, it was nonetheless quite applicable and was tuned to CSS by work developed in Batchelder *et al.* (1997). The authors provided methodology to estimate an underlying true network by focusing on informants' "hit" and "false alarm" rates, i.e., true and false positive rates in reporting edges between individuals. In a series of unpublished papers, Koskinen furthered these models. Koskinen (2002*b*) placed a slightly tweaked version of Batchelder *et al.* (1997) within a Bayesian framework, Koskinen (2002*a*) incorporated covariates into this model, and Koskinen (2004) described how to perform model selection in this context via a Markov chain Monte Carlo (MCMC) approximation of the marginal likelihood. The hierarchical network accuracy model (HNAM) of Butts (2003) also used a Bayesian approach which, while distinct, has a similar flavor as these others. Butts (2003) assumed an underlying but unknown network, and provided estimation by accounting for false positive and false negative probabilities for each individual. Almquist (2012) provided a similar approach for egocentric network data. Additionally, Siciliano *et al.* (2012) developed a method of estimating a network by asking a subset of the network actors to enumerate the entire network. This last method, however, assumes that each actor reports with perfect accuracy his or her own ties—an assumption repeatedly shown to be faulty (Bernard *et al.*, 1982; Brewer, 2000; Brewer & Webster, 2000).

Others have studied multiple replications of networks outside of the context of CSS. This type of problem occurs when researchers are looking at network-valued samples from a population (e.g., Durante *et al.*, 2017), or performing meta-analyses (e.g., Butts, 2011). As these settings often deal with different types of data and focus on very different research questions, these works are somewhat less relevant to our present purpose.

Our proposed method provides an inferential framework on network perception data (i.e., CSS's) using a latent space model, a class of models first introduced by Hoff *et al.* (2002). Previous studies have shown "forgetting is a pervasive, non-trivial phenomenon in the recall-based elicitation of personal and social networks pertaining to a broad variety of social relations" (Brewer, 2000, p.29). By performing inference on the data generating process using our proposed model, relativists may be able to account for this potential informant error, or test–retest variability, by focusing on the probabilities of their responses, thereby obtaining a better understanding of the actors' perceptions of the network. The latent space framework can capture the complex dependence structure in the data, thereby obtaining better estimates of these underlying response probabilities. Additionally, our proposed latent space model captures specific features allowing the relativist to better understand *why* the network perceptions may differ. Similarly, essentialists who define the ground truth in terms of commonly used aggregation structures may have an alternative which is more robust to inherent variability in the individuals' responses. In addition, by implementing a latent space approach, essentialists have a very natural interpretation of the latent positions, namely that of a low-dimensional representation of the underlying social reality.

Sosa & Rodriguez (2018+) also proposed another latent space model for network perception (LSMNP) data, relying on a bilinear operationalization and assuming respondent-specific sets of latent positions. Explicitly modeled is whether or not a respondent has the same network perception as that held by the majority of the actors in the network. The proposed work contains some ideas similar to this, but holds several distinctions. First, the parameter space of the proposed method is dramatically more parsimonious than that of Sosa & Rodriguez which grows quadratically with n. In addition, our proposed approach captures each actor's overall level of confidence in his/her perceptions, the variability within each respondent's uncertainties about the $\binom{n}{2}$ pairs of actors, each actor's overall tendency to under or overreport edges, and we also describe a sensible mechanism for estimating the effect of covariates on respondents' network perceptions.

This paper makes the following contributions. First, we propose a formal statistical framework for understanding CSS data. Rather than first estimating the underlying "actual" network and then performing subsequent analyses while ignoring any possible estimation error in the first stage, our approach provides a unifying framework in which one may implement the commonly used latent space model directly on the network perception, or CSS, data, thus estimating, e.g., homophilic effects without introducing unaccounted estimation error. Second, our modeling

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

framework allows researchers to directly understand multiple salient features of informants' responses. Specifically, our model decomposes the deviations in the informants' responses into two meaningful ways: bias and variance. The bias corresponds to an overall tendency toward reporting a more or less dense network. The variance corresponds to the respondents' confidence in their network perceptions, and importantly this variance is allowed to be a function of distance in the network space between the informant and the actors being reported on. Covariates can be used to explain why actors perceive the network in the way that they do. Third, we describe how to take the output of the analysis and construct a single meaningful aggregated network reflecting specific researcher questions in the form of what we refer to as locally aggregated probabilities (LAPs). Fourth, we discuss how to expand the proposed approach in order to relax the commonly used forced-choice design in which each actor is forced to enumerate the entire network. This is important for future studies, particularly those involving large networks, in that the actors of the network are no longer forced to give a response for each pair of actors. Particularly in the case of large networks (e.g., the social network of students on a large college campus), it is unreasonable to assume that each respondent holds the same level of knowledge about each other pair of actors, and hence a forced-choice design is in effect forcing the respondents to add noise to the data. Finally, the proposed method provides a visualization method for the network perception data, condensing an $n \times n \times n$ array into a single intuitive and statistically meaningful figure.

This paper is organized as follows. Section 2 provides a short background on CSS. In Section 3, we describe our statistical model of CSS data, our method of estimation, how to incorporate covariate information into the model, how to use the model to construct networks based on varying definitions, and how our methods might be applied to larger networks for which CSS is currently infeasible. Section 4 describes a simulation study. Section 5 describes an analysis of the advice-seeking relationship data from Krackhardt (1987). We end with a brief discussion in Section 6.

2. Background

Freeman *et al.* (1987) (p.310) stated "survey methods are at the core of most sociological research." While technological advances have in more recent times changed the collection of network data whether it be online data, mobile-phone data, or data collected via proximity sensors, it is still true that many network data sets are being collected via surveys or interviews (e.g., Lorant *et al.*, 2015; Shakya *et al.*, 2015; Perry & Pescosolido, 2015; Shoham *et al.*, 2016). Freeman *et al.* go on to describe how informant inaccuracy with respect to an observable truth is a serious problem that should not be ignored. This problem was first seriously brought to light through several papers by Bernard, Killworth and Sailer (BKS) (Killworth & Bernard, 1976, 1977, 1979; Bernard *et al.*, 1979, 1982), in which a very pessimistic picture was put forth on the usefulness of social network data collected via informants, making statements such as "we have been unable to show …that cognition is related to behavior in any meaningful way whatever" (Bernard *et al.*, 1979, p.209). This picture put forth by BKS is fairly extreme, and while informant accuracy is an important issue that has been remarked upon by many others (e.g., Hammer, 1985; An & Schramski, 2015), it should be stated that other work has shown the value of self-reported network data (e.g., Romney & Faust, 1982). For example, Freeman & Romney (1987) and Freeman *et al.* (1987) have shown that self-reports may be more trustworthy when treated as representations of a "stable pattern of repeated events" (Freeman *et al.*, 1987, p.310) Additionally, in other instances, reliability has been less egregious (e.g., Bell *et al.*, 2000, noted a higher degree of reliability of drug sharing partnerships).

Hildum (1986) responded to the apparent disconnect between self-reports and observed behaviors by relating this sociological issue to those commonly dealt with by linguists. Hildum describes how just as an individual may paraphrase a statement previously made by himself/herself by using language, i.e., a vocabulary and set of syntactic rules that exists in the individual's head, so too may a network actor report on network interactions about which he or she is uncertain by using an abstracted or idealized network that is a cognitive construct of that individual. Hildum (p.85) states the extreme relativist perspective: "The *behavior* is true in the precise and limited sense that it is observable and subject to observer agreement. But the *network* exists and is 'true' only in the participant's mind." The next year, Krackhardt (1987) introduced CSS, describing how the actors' network perceptions may be compared with one another and how focusing on the cognitive reconstructions themselves allows researchers to test social theories such as Heiderian structural balance (Wasserman & Faust, 1994). These and other works helped broaden the study of informant responses to their relationship with cognition rather than just behavior. Indeed, it could easily be argued that behaviors ought to be modeled in a very different framework, such as the relational event model (Butts, 2008).

Many researchers, taking the essentialist stance, have assumed that there is in fact some underlying truth acting as a stimulus for social cognition. Krackhardt himself published papers assuming an "actual" network structure defined by whether or not two actors both agree that a relation exists between them (e.g., Krackhardt, 1990; Kilduff *et al.*, 2008). This type of aggregated network (more on this in Section 3.4) has been used as a ground truth in a large number of studies (e.g., Casciaro, 1998; Casciaro *et al.*, 1999; Johnson & Orbach, 2002; Aarstad *et al.*, 2011; Neal *et al.*, 2014, 2016) to study informant accuracy or network topologies. As mentioned in the introduction, others such as Batchelder *et al.* (1997), Koskinen (2002*b*), and Butts (2003) have developed statistical methods of inferring an underlying true sociomatrix. Still others have constructed a low-dimensional representation of the social reality rather than estimating an $n \times n$ matrix; for example, Kumbasar *et al.* (1994),Romney *et al.* (1996), and Batchelder (2002) used correspondence analysis to obtain a low-dimensional structure to visualize and analyze.

As the intention of this paper is to advance methodological developments rather than perform a systematic review of CSS, what has been presented here is necessarily brief. For more details, we suggest looking at Pattison (1994) or Brands (2013).

3. Methodology

3.1 Model

We assume that we have collected network perceptions on a random subset $\mathcal{S} \subseteq \{1, 2, \dots, n\}$ of K $(= |\mathcal{S}|)$ individuals, where n is the total number of actors in the network (note that in historical CSS data $\mathcal{S}$ often equals the full set of actors). For each individual in $\mathcal{S}$ we collect a report of the network which may be represented as an $n \times n$ adjacency matrix; we will denote these as $\{A_k\}_{k \in \mathcal{S}}$. The ith row jth column entry $A_{k,ij}$ represents individual k's reporting of whether or not an edge exists from actor i to actor j. The notation we will use assumes directed edges, but it should be trivial to adapt the methods of this section to the undirected case.

A key component of our approach is based on latent space models for network data. While conceptually researchers had embedded network data into low-dimensional spaces (e.g., Nakao & Romney, 1993; Kumbasar *et al.*, 1994; Romney *et al.*, 1996; Batchelder, 2002), this approach was formalized by Hoff *et al.* (2002) and Hoff (2005), and has since been further developed in a plethora of works (e.g., Handcock *et al.*, 2007; Hoff, 2009; Durante & Dunson, 2014; Sewell & Chen, 2016). Our parameterization of the likelihood will be based in large part on Krivitsky *et al.* (2009). These latent positions will be denoted by the $n \times p$ matrix $Z = (\mathbf{Z}_1, \dots, \mathbf{Z}_n)'$, as well as the $n \times 1$ vectors of actor-specific sender effects $\mathbf{s}$ and receiver effects $\mathbf{r}$, corresponding to social activity and popularity, respectively.

Latent space models introduce nonlinear random effects into the model to account for sophisticated dependence structures in the data. For example, using a latent space approach can capture third-order dependencies (for details, see Hoff, 2005), and could, should it be desired, explicitly capture clustering effects (Handcock *et al.*, 2007; Sewell & Chen, 2017). In our work, these

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

random effects correspond to the positions of the actors in some underlying (latent) network space in conjunction with the actors' social activity and popularity. Following the tradition of the latent space literature, we assume that the inherently complex dependencies are fully induced by the latent variables; thus, conditional on these latent variables, the dyads are independent.

Before giving the model formulation, we first highlight some key concepts we wish to incorporate into the model. We will then give the mathematical representation of the model and subsequently tie the components of the model back to these concepts. First, we wish to capture some notion of a *commonality*, or the facets of the network that are perceived in a common way across informants. These components will allow us to leverage information across networks in order to better estimate each informant's true response probabilities. For the essentialist, the commonality may also be interpreted as a low-dimensional representation of the underlying "true" social reality that is fixed but unknown.

Second, we wish to capture two ways in which individual perceptions deviate from one another. The first way is that an individual can exhibit *bias*, by which we mean the tendency for an individual to report either a higher or lower number of edges overall in the network than the average informant. The second way is for an informant to hold varying degrees of *variability*, or equivalently precision, in his or her perceptions. This can be thought of as akin to the confidence an informant has in his or her knowledge of the network; if an informant is very confident, then that informant is unlikely to give different answers when asked at multiple times, whereas if an informant has no confidence in their knowledge of an edge, then he or she may just end up guessing, leading to much more variability if the informant were to be asked multiple times about the existence of the edge (low test–retest reliability).

Lastly, and highly related to this last concept, one's confidence or knowledge about an edge ought to be local in nature. That is, there ought to be *local variability* that depends on some form of distance between the respondent and the two actors about whom the respondent is reporting. This is reflective of previous findings of associations between informants' responses and degree centrality (Romney & Faust, 1982), social distance (Krackhardt & Kilduff, 1999), path length (Adams & Moody, 2007), and Brewer (2000) found that people tended to forget weak ties more often than strong ties.

Below is our proposed LSMNP. The likelihood is given as follows:

$$\pi\left(\{A_k\}_{k\in\mathcal{S}}|Z,\mathbf{s},\mathbf{r},\{\alpha_k,\beta_{0k}\}_{k\in\mathcal{S}},\beta_1\right)=\prod_{k\in\mathcal{S}}\prod_{i\neq j}\pi\left(A_{k,ij}|\mathbf{Z}_i,\mathbf{Z}_j,\mathbf{s}_i,\mathbf{r}_j,\alpha_k,\beta_{0k},\beta_1\right), \tag{1}$$

where α_k, β_{0k}, and β_1 are unknown parameters that will be described shortly. We utilize a probit model, which dictates that the probability of k reporting an edge between i and j is given by

$$\begin{aligned}\pi(A_{k,ij}=1|\mathbf{Z}_i,\mathbf{Z}_j,\mathbf{s}_i,\mathbf{r}_j,\alpha_k,\beta_{0k},\beta_1)&=\Phi\left(\tau_{k,ij}^{1/2}\left(\alpha_k+\mathbf{s}_i+\mathbf{r}_j-d\left(\mathbf{Z}_i,\mathbf{Z}_j\right)\right)\right)\\&=\pi\left(A^*_{k,ij}>0|\mathbf{Z}_i,\mathbf{Z}_j,\mathbf{s}_i,\mathbf{r}_j,\alpha_k,\beta_{0k},\beta_1\right),\end{aligned} \tag{2}$$

where for some normally distributed auxiliary random variable $A^*_{k,ij}$

$$A^*_{k,ij}|\mathbf{Z}_i,\mathbf{Z}_j,\alpha_k,\beta_{0k},\beta_1\overset{ind}{\sim}N\left(\alpha_k+s_i+r_j-d(\mathbf{Z}_i,\mathbf{Z}_j),\tau_{k,ij}^{-1}\right), \tag{3}$$

$$\alpha_k\overset{iid}{\sim}N(\alpha,\sigma^2), \tag{4}$$

$$(s_i,r_i)'\overset{iid}{\sim}N\left(\mathbf{0},\Sigma\right), \tag{5}$$

$$\tau_{k,ij}=\beta_{0k}\cdot\exp\left\{-\beta_1\left(d(\mathbf{Z}_k,\mathbf{Z}_i)+d(\mathbf{Z}_k,\mathbf{Z}_j)\right)\right\}, \tag{6}$$

$$\beta_{0k}\sim Ga(\theta/2,\theta/2), \tag{7}$$

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

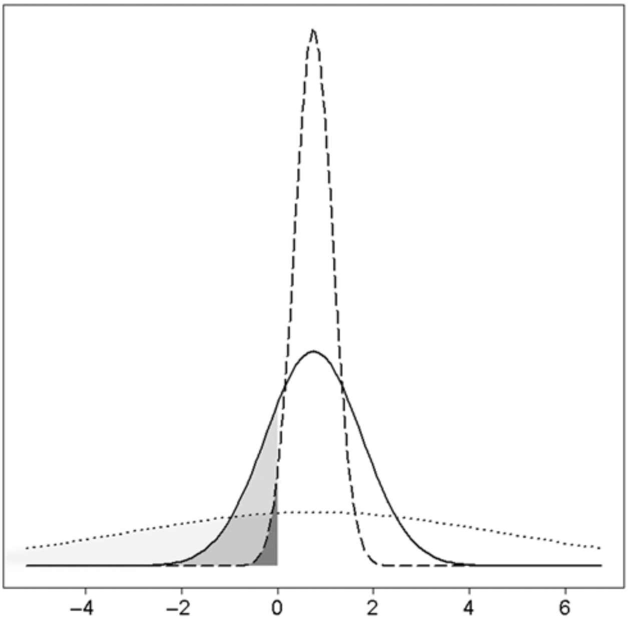

Figure 1. Illustration of variability. The curves correspond to the distribution of $A^*_{k,ij}$. The solid line corresponds to average variance, the dotted line corresponds to a high level of variance, i.e., nearly a 50/50 guess, and the dashed line corresponds to a high level of precision in reporting. The areas of the shaded regions give the probability of $A_{k,ij} = 0$.

and where $\Phi(\cdot)$ is the cumulative distribution function for a standard normal random variable, $d(\cdot, \cdot)$ is a dissimilarity measure such that $d : \Re^p \times \Re^p \mapsto \Re^+ \cup \{0\}$, $N(\boldsymbol{\mu}, \Sigma)$ is the normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix Σ, and $Ga(a, b)$ is the gamma distribution with shape a and rate b. Note that we have ensured model identifiability by constraining the mean of $\beta_{k,0}$ to equal 1 and the mean of **s** and **r** to equal **0**. For the remainder of this paper, we will use the Euclidean distance as the measure of dissimilarity between $\mathbf{Z}_i$ and $\mathbf{Z}_j$, although other choices could be employed such as the negative of the bilinear term $\mathbf{Z}_i'\mathbf{Z}_j$ or for directed networks $\mathbf{Z}_{i,1}'\mathbf{Z}_{j,2}$ where $\mathbf{Z}_i = (\mathbf{Z}_{i,1}', \mathbf{Z}_{i,2}')'$.

Now we will tie the parameters of the LSMNP to the concepts described previously. First, the commonality is captured in Z, **s**, and **r**. These components capture the shared perception of how close actors are in the network, and how certain actors tend to send or receive more or fewer edges. Note that an individual's sender and receiver effects are assumed to be correlated through Σ. The notion of bias is captured by α_k. For two actors k and ℓ such that $\alpha_k < \alpha_\ell$, when looking at a particular dyad in the network, actor k will, all other things being equal, be less likely to report an edge associated with the dyad than will ℓ. Variability is captured through the precision parameter $\tau_{k,ij}$. As this value decreases to zero, the probability that k will claim a link between actors i and j converges to a coinflip, i.e., $\lim_{\tau_{k,ij}\to 0} \mathbb{P}(A_{k,ij} = 1|\cdot) = 1/2$. As this value increases to infinity, the probability converges to either 0 or 1, depending on the sign of $\mathbb{E}(A^*_{k,ij})$. This is illustrated in Figure 1. An individual's overall variability is described by the parameter β_{0k}, where a larger value indicates a higher level of precision for all dyadic pairs. The idea of local variability is reflected through the second term on the right hand side of (6) (we shall constrain β_1 to be nonnegative). Variance thus increases as the respondent reports on actors farther away in the network space. If the researcher has reason to believe that respondents tend to bias their result as uncertainty grows rather than simply guessing, then instead of putting the exponential term of (6) in the precision parameter one may instead parameterize the mean of $A^*_{k,ij}$ to include $1 - \exp\left\{\beta_1\left(\|\mathbf{Z}_k - \mathbf{Z}_i\| + \|\mathbf{Z}_k - \mathbf{Z}_j\|\right)\right\}$ (or some other monotonically decreasing function in $\|\mathbf{Z}_k - \mathbf{Z}_i\|$ and $\|\mathbf{Z}_k - \mathbf{Z}_j\|$) for a bias toward zero, and the negative of this for a bias toward one. That is, if when i knows nothing about the relationship between j and k, i will be more likely to report, say,

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

a zero, then it may be more reasonable to have a more sophisticated model of the mean structure of $A^*_{k,ij}$ rather than a more sophisticated model of the precision.

Note also that as all α_k's tend toward α and as all β_{0k}'s tend toward ∞, the responses of all informants will match with probability converging to one. Hence, an individual's *accuracy*, a commonly studied measure (e.g., Bondonio, 1998; Casciaro, 1998; Aarstad *et al.*, 2011), can be defined by these two parameters, and moreover can be broken down by bias and variability.

In summary, the commonality is provided by the actor's positions within the network space, given by Z, and the social activity and popularity of the actors, defined by $\mathbf{s}$ and $\mathbf{r}$, respectively. An informant can deviate from the commonality in two ways, bias and variance. The notion of bias is captured by α_k. The notion of variance/precision is modeled as an individual's overall level of precision, captured by β_{0k}, and the degree to which this variability is local is captured by β_1.

3.2 Estimation

Estimation of the model parameters is done within a Bayesian framework. The goal is to obtain samples from the posterior distribution for all unknowns. We first place the following priors on the unknowns:

$$\alpha|\sigma^2 \sim N(\mu_\alpha, \nu_\alpha\sigma^2), \qquad \beta_1 \sim \ell N(\boldsymbol{\mu}_\beta, \nu_\beta),$$
$$\sigma^2 \sim IG(\gamma_\sigma/2, \eta_\sigma/2), \qquad \mathbf{Z}_i \stackrel{iid}{\sim} N(\mathbf{0}, \sigma_z^2 I_p),$$
$$\Sigma \sim IW(\gamma_\Sigma, \Gamma_\Sigma), \qquad \sigma_z^2 \sim IG(\gamma_z/2, \eta_z/2),$$
$$\theta \sim \ell N(\mu_\theta, \nu_\theta),$$

where $IG(a, b)$ is the inverse gamma distribution with shape a and scale b, $IW(a, B)$ is the inverse Wishart distribution with degrees of freedom a and scale matrix B, and $\ell N(a, b)$ is the log normal distribution with log-mean a and log-variance b.

As the sample sizes associated with this type of problem are small, the computation time associated with a random scan Metropolis–Hastings-within-Gibbs MCMC algorithm, while nontrivial, will typically not be prohibitive. The code we used in our analyses, which is on the author's website as an R package, took 70 s to run 50,000 iterations on the data of Section 5 ($n = 21$) on a machine with a 3.6-GHz processor and 32 GB of RAM.[1] Most of the full conditional distributions can be found in closed form, but for Z, θ, and β_1 we must implement Metropolis–Hastings (MH) steps in order to obtain posterior draws. These full conditionals of the algorithm used in Sections 4 and 5 are given in the Supplementary Material along with a few other details.

Also included in the Supplementary Material are the details on how we initialized the MCMC algorithm in the analyses presented in Sections 4 and 5. We also analyzed the data of Section 5 after initializing the MCMC algorithm with the Maximum a Posteriori estimators of the parameters and latent positions. This was done via an Iterated Conditional Modes algorithm, using Lagrange multipliers to constrain the means of $\mathbf{s}$ and $\mathbf{r}$ to be zero and of the β_{0k}'s to be one. There was no discernible effect on the posterior distribution of the parameters based on the initialization scheme. A figure comparing the two posterior distributions is given in the Supplementary Material. Hence, while further work would be required to show just how poor one's initialization must be before the MCMC algorithm is affected, in this analysis it appeared that a reasonable ad-hoc initialization performed just as well as a very good initialization.

An important note in any estimation scheme is that Z is not identifiable, as the likelihood is invariant to rotations, reflections, and translations. This is not an issue in computing the posterior distribution of the distance matrix associated with Z, but there are problems with computing the posterior distribution of Z itself. This has previously been addressed by implementing the Procrustes transformation on each MCMC draw (e.g., Hoff *et al.*, 2002; Sewell & Chen, 2015). Specifically, after a burn in period of the MCMC algorithm, we choose the first draw as a target

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

matrix, Z^*. Each subsequent draw of Z in the chain is then subjected to a Procrustes transformation, which first finds the rotation matrix R such that the Frobenius norm of $ZR - Z^*$ is minimized, and then sets Z to equal ZR.

3.3 Adding covariates

Frequently the researcher's objective is to investigate the effect of independent variables on dyadic relations. However, when considering actors' perceptions of the network, it is not obvious how best to account for this. In the past, covariates have been used descriptively (Batchelder, 2002) and to estimate the accuracy of one's perceptions (Bondonio, 1998) (though arguably using statistically inappropriate techniques). Koskinen (2002*a*) incorporated covariates within the formal statistical model built off of the previous work by Batchelder *et al.* (1997) and Koskinen (2002*b*). In Koskinen's work, covariates were included in a linear predictor term for the informants' response probabilities. As we note at the end of this section, this same strategy may be taken here, but we favor a different approach.

While we believe it reasonable to think that each actor has an understanding of the network space (i.e., the commonality), it would not be reasonable in general to expect each actor to know all actor-level covariates. It might be reasonable to assume that each informant has some understanding of how close the friendships are in the network but may not know, for example, who has and has not been raised in a single-parent home. Therefore, the stimulus of each informant's responses ought not to be the covariates themselves, but rather that which acts as the stimulus ought to be affected by the covariates. To address this, one may augment the proposed methods in Section 3.1 by incorporating ideas described by Austin *et al.* (2013). Austin *et al.* proposed using covariates to predict the latent positions as well as the sender and receiver effects. It seems arguable as to whether or not this approach is most appropriate in the context of nonnoisy network data, e.g., mobile phone network data or social media network data, where we could incorporate the covariates directly into the likelihood function. In the context of observing network perceptions, however, these ideas seem ideally suited. Importantly, this approach also allows us to estimate homophilic effects on the low-dimensional representation of the underlying social reality, Z, as well as the covariates' effects on the social activity $\mathbf{s}$ and popularity $\mathbf{r}$, i.e., the propensity for informants to attribute more or fewer edges to an actor.

The intuition behind what we are now proposing is that each actor may know which actors are close to each other, but they may or may not know why; similarly, each actor may know which actors are popular/socially active and which are not, but they may or may not know why. Thus, as before, each actor's perception is determined by both the commonality as well as their own deviations from the other actors' perceptions. Now, however, we assume that the commonality itself is associated with the covariates. That is, actors may perceive that i and j are close or that k is popular, but the closeness itself and the popularity is, unbeknown to the actors, associated with some set of salient variables. The stimulus for the informants' responses is based primarily on the commonality which is impacted by the actors' characteristics, as opposed to the stimulus for the informants' responses being the actors' characteristics themselves.

In this framework, Equations (1)–(7) are unchanged, with the exception of equation (5). For an $n \times q$ matrix of covariates X, we now assume that

$$\begin{pmatrix} \mathbf{s} \\ \mathbf{r} \end{pmatrix} \overset{iid}{\sim} N\left(\begin{pmatrix} X\boldsymbol{B}_s \\ X\boldsymbol{B}_r \end{pmatrix}, \Sigma \otimes I_n \right), \tag{8}$$

and

$$Z \sim N_{n\times p}\left(XB_z, (\sigma_z^2 I_p) \otimes I_n\right), \tag{9}$$

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

where $\boldsymbol{B}_s$ and $\boldsymbol{B}_r$ are $q \times 1$ vectors of unknown parameters, B_z is a $q \times p$ matrix of unknown parameters, and $X \sim N_{n\times p}(M, \Sigma \otimes \Psi)$ indicates X follows the matrix normal distribution (i.e., $\text{vec}(X) \sim N_{np}(\text{vec}(M), \Sigma \otimes \Psi)$). The model specification used by Austin *et al.* assumed that Z was a deterministic linear function of the covariates. We allow more flexibility by setting this linear function as the mean of a random variable.

To perform estimation, we need to first assign priors to $\boldsymbol{B}_s$, $\boldsymbol{B}_r$, and $\boldsymbol{B}_z$. The priors used in the analyses in this paper were flat multivariate normal distributions:

$$\begin{pmatrix} \boldsymbol{B}_s \\ \boldsymbol{B}_r \end{pmatrix} \sim N\left(\boldsymbol{0}, \sigma^2_{B_{sr}} I_{2q}\right), \tag{10}$$

$$B_z \sim N_{q\times p}\left(0, I_p \otimes (\sigma^2_{B_z} I_q)\right). \tag{11}$$

Several of the full conditional distributions for the other parameters will be altered by the inclusion of covariates into the model. See the Supplementary Material for details.

Model identifiability again becomes an issue when we incorporate covariates. To maintain identifiability, we constrain the columns of X to be centered (mean zero). Note that this inherently removes any intercept from the design matrix X. This implies that $\sum_i \mathbb{E}(\mathbf{s}_i) = \sum_i \mathbb{E}(\mathbf{r}_i) = \sum_i \mathbb{E}(Z_{i\ell}) = 0$ for all $\ell \in \{1, 2, \ldots, p\}$. Finally, we again need to address the likelihood's invariance to rotations. We must also consider B_z since for some rotation matrix R we have

$$\pi\ (Z, B_z, \cdot|\{A_k\}_{k\in\mathcal{S}}) \overset{\mathcal{D}}{=} \pi\ (ZR, B_zR, \cdot|\{A_k\}_{k\in\mathcal{S}}).$$

We need to only slightly modify the procrustes procedure described in Section 3.2 to eliminate identifiability issues with B_z. Specifically, if we rotate a new draw of the latent positions by some rotation matrix R which minimizes $\|ZR - Z^*\|_F$ for some target matrix Z, then we necessarily need to rotate B_z by the same rotation matrix. Hence if we reset Z to be ZR, we must also reset the current value of B_z to be B_zR.

3.4 Locally aggregated probabilities

With CSS data, we are sometimes required to aggregate data to obtain useful results. For example, if we assume an underlying network (from the essentialist perspective) we often times wish to better understand the topology of the network (e.g., Johnson & Orbach, 2002; Kilduff *et al.*, 2008; Grippa & Gloor, 2009, aggregate CSS data to investigate network features). Yet how do we do this? Do we construct an edge between actors i and j only if i claims a relationship toward j, or perhaps only if j claims a relationship from i? Do we construct such an edge if either i or j claim an edge from i to j, or perhaps only if both i and j claim that the directed edge exists? The primary question one must answer is "what is it we are trying to measure?" The answer to this should provide the answers to the former questions.

Krackhardt (1987) introduced the notion of "locally aggregated structures" (LASs), which are reductions of CSS that make inference on the relationship (relationship s in the context of digraphs) between two actors i and j based only on the responses of i and j. Here "locally" is used in the sense that the structure is constructed from the responses of the two most local actors of the network, i and j themselves. LASs have been widely used to define a reference network, often described as the "true," "actual," or "benchmark" network (Krackhardt, 1990; Bondonio, 1998; Casciaro *et al.*, 1999, and references therein) and hence are important to be able to estimate accurately.

Here we point out two ways in which our context and methods differ from that of Krackhardt's. First, we do not assume we have data on all members of the network, and so we cannot always completely reconstruct a network using LASs. Second, due to the noisiness of survey network

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

data (Brewer, 2000), we are interested in the underlying probabilities of the LASs in the hope of obtaining representations more robust to low test–retest reliability. We therefore present LAPs which closely mirror the original LASs of Krackhardt.

- Row-dominated LAP: $\mathbb{P}(A_{i,ij} = 1)$,
- Column-dominated LAP: $\mathbb{P}(A_{j,ij} = 1)$,
- LAP intersection rule: $\mathbb{P}(A_{i,ij} \cap A_{j,ij})$,
- LAP union rule: $\mathbb{P}(A_{i,ij} \cup A_{j,ij})$.

To quote Krackhardt (p.116), row-dominated LAP corresponds to the question "who are you related to in this way?", and the column-dominated LAP corresponds to the question "who is related to you in this way?" The intersection and union rules correspond, respectively, to the questions "Do both individuals agree that they are related in this way?", or "Does at least one of the individuals claim such a relation?" LASs thus generalize data collected in traditional sociometry and enable researchers to answer more precise questions based on how one defines the network.

It is important to note that like the LASs, LAPs are focused on how a pair of actors would respond, but unlike LASs, LAPs leverage information from everyone in the network since these probabilities are in large part constructed from the commonality (they also, of course, depend on the individual deviations of i and j, thereby accounting for the bias and variance of these two actors). Using the LAPs allows one to make either hard or soft predictions of the responses of the actors, and in so doing construct a network according one of the above notions. Note that the soft predictions may be used as a weighted network to be further analyzed.

The LAPs can be estimated in the usual ways, such as the posterior mean. For example, the posterior mean of the LAP intersection rule for the pair (i, j) is estimated by

$$\frac{1}{M}\sum_{\ell=1}^{M}\left[\Phi\left(\sqrt{\tau_{i,ij}^{(\ell)}}\left(\alpha_i^{(\ell)} + \mathbf{s}_i^{(\ell)} + \mathbf{r}_j^{(\ell)} - d(\mathbf{Z}_i^{(\ell)}, \mathbf{Z}_j^{(\ell)})\right)\right)\right] \cdot \left[\Phi\left(\sqrt{\tau_{j,ij}^{(\ell)}}\left(\alpha_j^{(\ell)} + \mathbf{s}_i^{(\ell)} + \mathbf{r}_j^{(\ell)} - d(\mathbf{Z}_i^{(\ell)}, \mathbf{Z}_j^{(\ell)})\right)\right)\right],$$

where M is the number of MCMC draws and the superscript $^{(\ell)}$ indicates the ℓ^{th} draw.

3.5 Forced-choice

In many network settings, particularly as the size of the network grows, it can be expected that each actor will have only a partial knowledge of the network. Indeed, we have tried to somewhat account for this already by letting the precision decrease with distance as seen in (6). At some point, the respondent is clearly only guessing. To quote Butts (2003) (p.136), "one might also argue that [standard practice in network research] occasionally prompts us to collect too *much* [data on ties between actors]." The respondents will, for some edges, not feel knowledgeable, yet, in the forced-choice framework, they *must* give some answer.

It would be reasonable, then, to include the option for the respondent to claim ignorance regarding a pair of actors. In our current framework, it is obvious (we hope) what would lead a respondent k to declare such ignorance: the two actors in question are simply too far away in the network space for k to have any common knowledge about them (recall that the commonality is primarily reflected in Z). In other words, there is a *perception radius* about which k has a *local perception* of the network, but beyond which the respondent knows very little.

A simple way to incorporate this into the methods described so far could be the following: let $Y_{k,ij}$ equal one if k claims knowledge about an edge between i and j, and zero if k claims ignorance. Then $\mathbb{P}(A_{k,ij} = 1 | Y_{k,ij} = 1, \cdot)$ would be of the same form as (2), and $\mathbb{P}(A_{k,ij} = 1 | Y_{k,ij} = 0) = p_k$

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

(e.g., $p_k = 1/2$). Of course, if $Y_{k,ij} = 0$ then $A_{k,ij}$ is missing, but since we do not care to perform inference on the p_k's, this does not matter. A reasonable form for the distribution of $Y_{k,ij}$ that follows the same logic as the proposed model would be

$$\mathbb{P}(Y_{k,ij} = 1 | \cdot) = \Phi\left(\delta_k - \frac{1}{2}\xi\left(\|\mathbf{Z}_k - \mathbf{Z}_i\| + \|\mathbf{Z}_k - \mathbf{Z}_j\|\right)\right), \tag{12}$$

where $\xi(>0)$ is a scale parameter and $\delta_k(\geq 0)$ can be thought of as k's radius which marks the boundary of how far the (scaled) distances can be before there is a less than 50% chance of k claiming knowledge about the link between i and j.

This leads to the following likelihood that is proportional to the parameters of interest:

$$\prod_{k\in\mathcal{S}}\prod_{i\neq j}\left[\Phi\left(\delta_k - \xi\left(\|\mathbf{Z}_k - \mathbf{Z}_i\| + \|\mathbf{Z}_k - \mathbf{Z}_j\|\right)\right)\right]^{Y_{k,ij}}\left[1 - \Phi\left(\delta_k - \xi\left(\|\mathbf{Z}_k - \mathbf{Z}_i\| + \|\mathbf{Z}_k - \mathbf{Z}_j\|\right)\right)\right]^{1-Y_{k,ij}}$$
$$\cdot\left[\Phi\left(\tau_{k,ij}^{1/2}(\alpha_k + \mathbf{s}_i + \mathbf{r}_j - \|\mathbf{Z}_i - \mathbf{Z}_j\|)\right)\right]^{Y_{k,ij}A_{k,ij}}$$
$$\times\left[1 - \Phi\left(\tau_{k,ij}^{1/2}(\alpha_k + \mathbf{s}_i + \mathbf{r}_j - \|\mathbf{Z}_i - \mathbf{Z}_j\|)\right)\right]^{Y_{k,ij}(1-A_{k,ij})}. \tag{13}$$

Estimation can be performed much as before, though we do not go into further detail here. To the author's knowledge, there is no publicly available data set collected in this way, though with this methodology one may be able to collect CSS-like data on medium to large networks.

4. Simulation studies

4.1 Setting the hyperparameters

Very often we wish for the prior distribution to have minimal impact on the inference, and hence choose flat priors. In the next simulation study (Section 4.2) and the real data analysis of Section 5, we chose priors to be flat. Specifically, we set $\mu_\alpha = 0$, $\nu_\alpha = 100$, $\gamma_\sigma = 0.001$, $\eta_\sigma = 0.001$, $\gamma_\Sigma = 2$, $\Gamma_\Sigma = 0.001I_2$, $\mu_\theta = 0$, $\nu_\theta = 3$, $\mu_\beta = 0$, $\nu_\beta = 100$, $\gamma_z = 0.001$, $\eta_z = 0.001$, $\sigma^2_{B_{sr}} = 100$, and $\sigma^2_{B_z} = 100$.

To determine how sensitive the posterior mean estimates were to the choice of hyperparameters, we analyzed the data of Section 5 100 times, where each time we randomly simulated a set of hyperparameters. These hyperparameters were drawn according to the distributions given in Table 1. For example, for each of the 100 analyses we draw μ_α from a normal distribution centered at zero with standard deviation equal to 2. We then computed the correlation between the posterior means and the hyperparameters over the 100 analyses. These correlations are also given in Table 1. From this table, we see that in almost all cases the effect of the hyperparameter has a small relationship with the parameter estimate. The exceptions to this are θ and β_1. However, neither of these had very large variation in the posterior means (coefficient of variation equal to 0.62 and 0.19, respectively), rendering these stronger correlations less worrisome.

4.2 Choosing the dimension of the latent space

An important facet of fitting any latent space network model is choosing the dimension of the latent space p. We performed a simulation study to evaluate the Deviance Information Criterion (DIC) (Spiegelhalter *et al.*, 2002) as a method for dimension selection. We simulated 100 data sets in the following manner. First, we fit the data described in Section 5. We then used the posterior mean of the model parameters for $p = 3$ to generate data according to the generative model given by Equations (2)–(7). We then ran the random scan MH-within-Gibbs MCMC algorithm for each of the 100 data sets for p equal to $1, 2, \ldots, 10$ (for a total of 100 data sets and 1,000 model fits). For each data set, we computed the DIC for these 10 fits and chose the $\hat{p}$ with the smallest DIC.

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

Table 1. Results from a simulation study testing the effect of hyperparameters on the posterior mean estimate of the parameters. One hundred analyses of the same data were performed where for each analysis the hyperparameters were drawn from the distribution given in the third column. The fourth column provides the correlation between the hyperparameters and the parameter estimates

Parameter	Hyperparameter	Distribution	Correlation
α	μ_α	$N(0, 4)$	0.10
	ν_α	$Ga(1, 0.01)$	−0.17
σ^2	γ_σ	$Ga(1, 100)$	0.04
	η_σ	$Ga(1, 100)$	−0.02
θ	μ_θ	$N(0, 4)$	0.55
	ν_θ	$Ga(1, 0.5)$	0.27
$\log(\beta_1)$	μ_β	$N(0, 4)$	0.47
	ν_β	$Ga(1, 0.5)$	−0.58
σ_z^2	γ_z	$Ga(1, 100)$	0.10
	η_z	$Ga(1, 100)$	0.12
Age (sender)			−0.15
Tenure (sender)			−0.09
Manager (sender)			0.02
Dept2 (sender)			0.09
Dept3 (sender)			0.10
Dept4 (sender)			−0.07
Age (receiver)	$\sigma^2_{B_{sr}}$	$Ga(1, 0.01)$	−0.18
Tenure (receiver)			−0.05
Manager (receiver)			−0.02
Dept2 (receiver)			≈ 0
Dept3 (receiver)			0.07
Dept4 (receiver)			−0.08

The true dimension therefore was 3 and we would hope that DIC would choose $\hat{p} = 3$ more often than any other value. Indeed, this was the case, as 3 was chosen correctly 77% of the time. The dimension was chosen by DIC to be either 3 or 4 97% of the time, and never was the dimension chosen to be greater than 5. In 95% of the cases $p = 3$ was in the smallest two DIC values, 99% in the smallest three, and always within the smallest four. Figure 2 shows the rankings over the 100 simulations for each $p \in \{1, 2, \ldots, 10\}$. Each line corresponds to a value of p (labeled at its highest point), the horizontal axis corresponds to the ranking (lowest to highest DIC), and the vertical axis corresponds to the proportion of simulations that each p achieved that particular ranking. Interestingly, from this figure we can see that $p = 1$ and $p = 2$ were almost always the largest and second largest DIC, respectively, despite it being closer to the truth than $p \in \{6, \ldots, 10\}$, implying that while DIC does a good job at selecting either the correct dimension or one dimension too large, it is highly unlikely that DIC will lead one to underfit the data.

5. Advice-seeking network

As an example data analysis, we now apply the proposed methods to an advice-seeking CSS data set (Krackhardt, 1987). These data come from a manufacturing firm consisting of around 100 employees that produces high-tech machinery. Twenty-one of these employees are part of management, and these individuals were each given a survey consisting of a series of questions about who goes to whom for help and advice at work. Hence, we have a 21 × 21 × 21 array of data, where

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

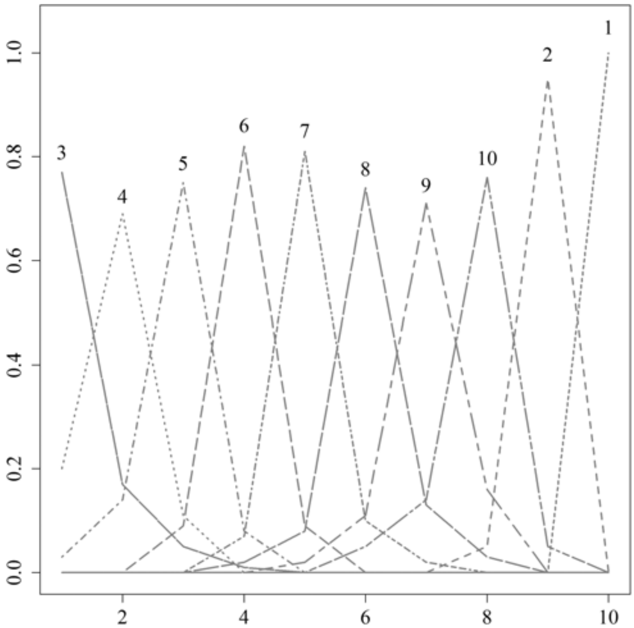

Figure 2. Results from a simulation study evaluating the efficacy of DIC to choose the correct dimension of the latent space. The true dimension is 3. Each line corresponds to the DIC when fitting a particular p, $p \in \{1, 2, \ldots, 10\}$ (labeled at its highest point); the horizontal axis corresponds to the ranking (lowest to highest DIC); and the vertical axis corresponds to the proportion of simulations that each p achieved that particular ranking.

the kth slice is the network perception of the kth employee. We also have covariate information on these individuals. We thus model the mean of the sender and receiver effects and of the latent positions as linear combinations of age, tenure (i.e., length of employment at the firm), position (either a vice president or manager), and to which of the four departments the individuals belong. One individual was the CEO, not belonging to a department. To account for this individual, we added another (nuisance) factor level variable that was set to 1 for the CEO and 0 for all others.

We fit the LSMNP to the full data set, letting the MCMC run for 3 million iterations, thinning by keeping every 1,000th iteration, and using a burn in period of 2 million iterations. We fit the data with $p = 1, 2, 3$. We were unable to obtain convergence within a reasonable number of iterations for larger values of p. The DIC values for $p = 1, 2, 3$ were, respectively, 7662, 7271, and 7192, implying that the dimension of the latent space we should select is $p = 3$. Trace plots, ACF plots, and CCF plots are all provided in the Supplementary Material. We performed the Geweke diagnostic (Geweke, 1992) to determine convergence of the MCMC chain. This yielded the following p-values corresponding to the null hypothesis that the samples are drawn from a stationary distribution for $\alpha, \sigma^2, \beta_1, \sigma_Z^2$, and θ, respectively: 0.59, 0.65, 0.48, 0.76, and 0.16. In order to investigate the degree to which we may be overfitting the data we also analyzed a subset of size $K = 4$ (hence we are analyzing an $n \times n \times 4$ array, $\approx 20\%$ of the data). For this smaller data set, we needed to run 10 million iterations to ensure the chain had converged. The Geweke diagnostic yielded the following p-values for $\alpha, \sigma^2, \beta_1, \sigma_Z^2$, and θ, respectively: 0.15, 0.32, 0.54, 0.67, 0.08.

For each data set ($K = 21$ and $K = 4$), we estimated the four LAPs described in Section 3.4. We then evaluated the fit of the LAPs with the corresponding observed LASs via the area under the ROC curve (AUC) and mean absolute error (MAE). We also computed the AUC and MAE using HNAM and the methods of Swartz *et al.* (2015) (which we will refer to as SGM). LASs are extremely useful and have historically been widely used in practice; the LAPs that have a high (low) AUC (MAE) demonstrate a strong relationship with these LASs while being more robust to informant error. Finally, we compared all three methods using DIC. Table 2 provides the results from these analyses. By both AUC and MAE, SGM performs the best when the full data are fit, but performs the worst when only part of the data are fit and the measures of fit are dominated by out-of-sample predictions. This seems to imply that SGM, whose parameter space is $\mathcal{O}(n^2)$,

Table 2. AUC and MAE values for comparing LSMNP, SGM, or HNAM with the observed LASs. The values corresponding to the best performance are bolded.

		Row		Column		Intersection		Union		
		AUC	MAE	AUC	MAE	AUC	MAE	AUC	MAE	DIC
$K = n$	LSMNP	0.807	0.339	0.895	0.310	0.885	0.235	0.859	0.330	**7192**
	SGM	**0.874**	**0.295**	**0.897**	**0.301**	**0.904**	**0.223**	**0.898**	**0.279**	7286
	HNAM	0.790	0.360	0.873	0.336	0.881	0.252	0.816	0.358	7805
$K = 4$	LSMNP	0.746	**0.348**	0.775	**0.341**	0.820	**0.248**	0.790	**0.345**	**1402**
	SGM	0.725	0.401	0.768	0.394	0.816	0.285	0.740	0.369	1778
	HNAM	**0.759**	0.385	**0.802**	0.381	**0.847**	0.280	**0.802**	0.360	1518

tends toward overfitting the data. When we compare LSMNP to HNAM when $K = 4$, AUC prefers HNAM, while MAE prefers LSMNP. For both $K = 21$ and $K = 4$, DIC indicates that LSMNP ought to be preferred. Beyond the conclusions of MAE and DIC, there are several reasons to prefer the use of the LSMNP. First, this approach provides a much more nuanced understanding of why network perceptions differ, as described in detail in Section 3.1. Further, the LSMNP can incorporate covariates directly into the model, allowing for the estimation of homophilic effects without relying on a two-stage estimation approach. Finally, the model naturally outputs a statistically meaningful visualization of the array of network perceptions.

We have also considered graphical measures of goodness-of-fit as proposed by Hunter *et al.* (2008) in order to validate the LSMNP as a data generating mechanism for the high-tech advice CSS. Of course this graphical goodness-of-fit was designed with a single network in mind, and hence we only considered a consensus structure (Krackhardt, 1987) with the threshold set such that an edge existed in the consensus structure if it was reported by 25% or more of the actors. That is, the consensus structure is an $n \times n$ adjacency matrix A_{CS} defined such that $A_{CS,ij}$ equals 1 if $\sum_{k=1}^{n} A_{k,ij} >$ threshold and zero otherwise. We considered the out-degree distribution, the in-degree distribution, the geodesic distance distribution, and triad census (i.e., counts of 16 possible types of triads as described by Davis & Leinhardt (1972)). We then used the predictive posterior distribution to simulate 1,000 CSS data sets, computed the consensus structure for each data set, and then computed the in- and out-degree distributions, the geodesic distributions, and the triad census. We also did the same for SGM and HNAM. Figure 3 shows the boxplots for the 3,000 data sets (1,000 per method), with the observed in-degree distributions (out-degree looked very similar and was thus omitted for space), geodesic distance distribution, and triad census given in the solid line. The solid gray lines indicate marginal 95% credible intervals. From this figure, we see that the LSMNP appears to be a the best representation of the true data generating process, as SGM and HNAM mischaracterize the number of dyads that are directly linked as well as the number of various triad configurations, with the SGM having much smaller mischaracterizations than HNAM. This figure provides evidence that the LSMNP is a good approximation of the true data generating mechanism.

Figure 4 shows the posterior means of the sender and receiver effects for each of the 21 actors. The actor's covariates have also been indicated graphically in this figure, and from it we can gain insight into the relationship between an individual's covariates and the tendency for the actors in the network to attribute more or fewer edges to that individual (note that more formal inference comes from looking at $\boldsymbol{B}_s$ and $\boldsymbol{B}_r$). We see that there appears to be a positive relationship between an individual's sender and receiver effect, and that vice presidents (and the CEO) tend to have higher sending and receiving effects than do managers. Table 3 provides the posterior means and 95% credible intervals for the coefficients $(\boldsymbol{B}_s, \boldsymbol{B}_r)'$. From this table we can see that, as may be expected, the more tenure the more likely the individual is to be sought after for advice, and we also see that managers are less likely to both seek and provide advice than are vice presidents, as well as some differences in general advice-seeking behavior between departments.

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

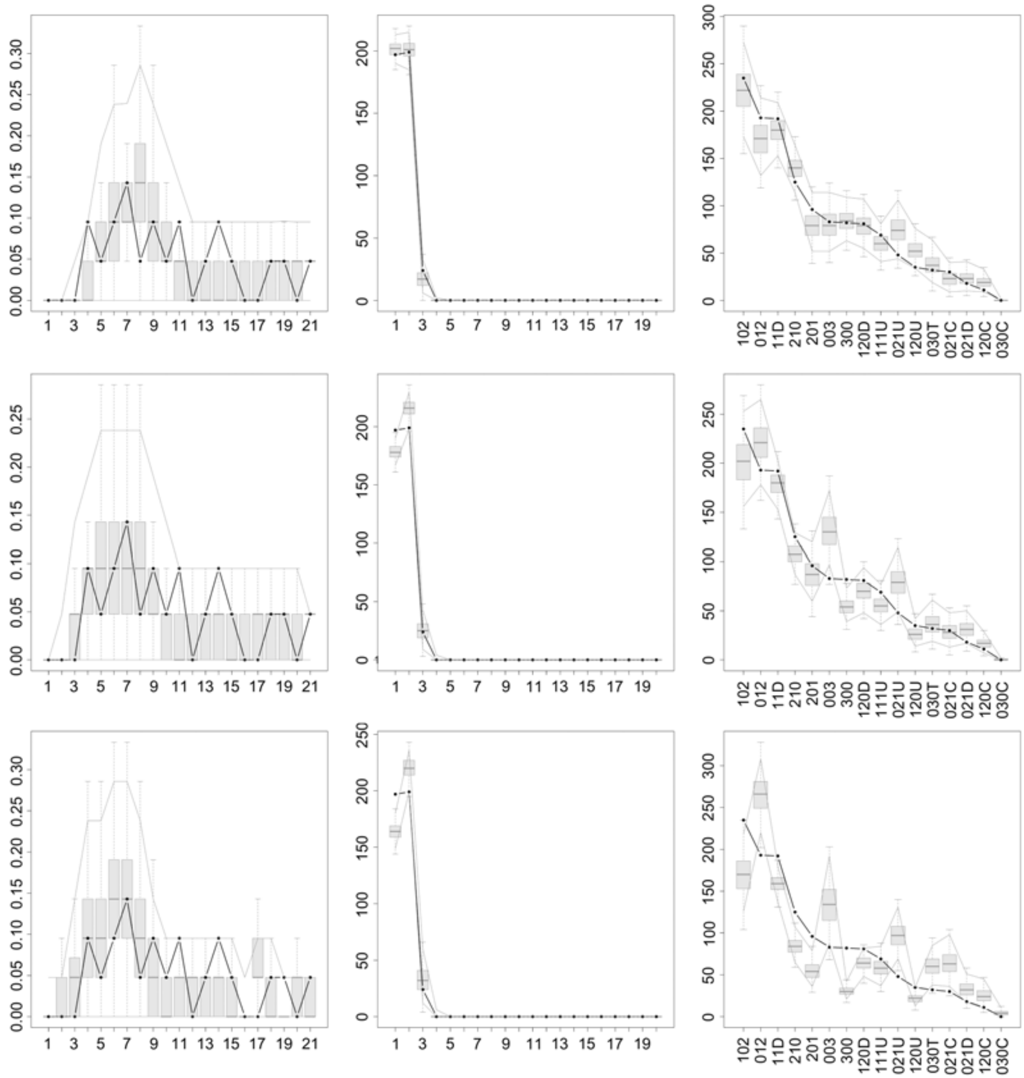

Figure 3. Graphical goodness-of-fits. Boxplots correspond to the consensus structures from 1,000 CSS data sets generated from each of the LSMNP (top row), SGM (middle row), and HNAM (bottom row) predictive posterior distribution, the solid gray lines give the 95% credible intervals, and the solid line gives the observed consensus structure. The columns correspond to in-degree distribution, geodesic distance distribution, and triad census distribution from left to right, respectively.

Figure 5 shows the MCMC draws of the latent positions after post-processing via the Procrustes transformation (a color version is available in the Supplementary Material, as well as a similar plot for the sender and receiver effects). To increase the clarity, we used principal components analysis (PCA) to plot the postertior means of the latent positions in a two-dimensional space in Figure 6. An edge was drawn if at least two-thirds of respondents reported it, with the thickness corresponding to the number of respondents that reported the edge. Again, the covariates are shown to help understand how the covariates are related to the social locales of the actors. From this we see that the departments are separated and that within each department vice presidents tend to be separated from the managers. Table 4 provides the posterior means and 95% credible intervals for the coefficients B_z. From this table, we confirm what we see visually in Figure 6: where an individual is located in the network space depends primarily on the department to which the

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

Table 3. Estimates and 95% credible intervals for $\boldsymbol{B}_s$ and $\boldsymbol{B}_r$. The baseline category corresponding to manager is vice president

	Sender	Receiver
Age	0.015 (-5.8×10^{-6}, 0.029)	0.00085 (−0.018, 0.018)
Tenure	−0.0096 (−0.027, 0.0066)	**0.031 (0.0080, 0.053)**
Manager	**−0.72 (−0.93, −0.50)**	**−1.8 (−2.2, −1.6)**
Dept 2	0.00069 (−0.27, 0.24)	**−0.43 (−0.81, −0.12)**
Dept 3	**−0.49 (−0.74, −0.25)**	**−0.54 (−0.86, −0.18)**
Dept 4	**0.43 (0.092, 0.81)**	**0.52 (0.12, 0.98)**

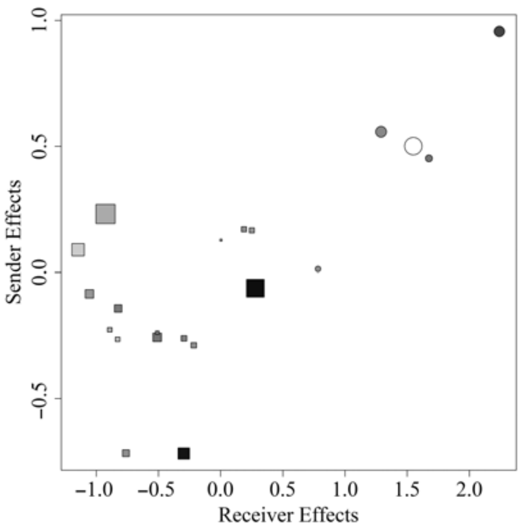

Figure 4. Sender effects (**s**) and receiver effects (**r**) corresponding to the advice-seeking network. Circles indicate vice presidents and squares indicate managers. The size corresponds to the age of the individual (larger implies older), and similarly the shading corresponds to the tenure (darker implies longer tenure). The hollow circle corresponds to the CEO.

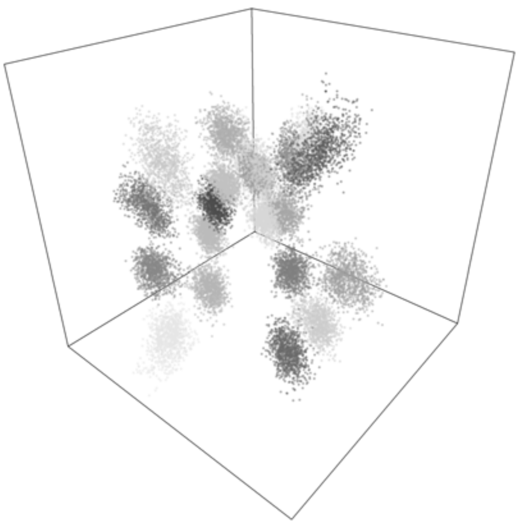

Figure 5. MCMC draws of the three-dimensional latent positions Z. (A color version is available in the Supplementary Material.)

individual belongs. There is additional evidence that the position (manager vs. vice president) also plays a role in an individual's location in the network space. That is, there is evidence of assortativity/homophily by the level of management and by department.

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

Table 4. Estimates and 95% credible intervals for B_Z. The baseline category corresponding to manager is vice president

	Dim 1	Dim 2	Dim 3
Age	0.033 (−0.010, 0.071)	−0.0035 (−0.045, 0.042)	−0.0016 (−0.037, 0.052)
Tenure	−0.031 (−0.096, 0.026)	−0.038 (−0.097, 0.018)	0.039 (−0.037, 0.089)
Manager	−0.081 (−0.96, 1.3)	**0.88 (0.22, 1.78)**	−1.0 (−1.7, 0.060)
Dept2	1.5 (−0.25, 2.25)	**1.73 (0.40, 2.4)**	**1.5 (0.93, 2.9)**
Dept3	0.51 (−0.72, 1.5)	**2.6 (1.5, 3.4)**	0.59 (−0.077, 2.1)
Dept4	−0.90 (−2.0, 0.12)	**2.5 (1.1, 3.3)**	0.58 (−0.47, 2.1)

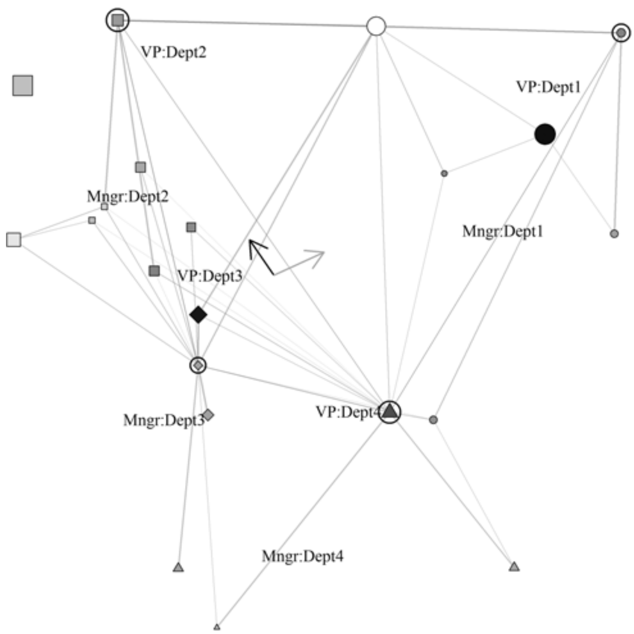

Figure 6. Latent positions Z corresponding to the advice-seeking network. The shape corresponds to the department (1 = circle, 2 = square, 3 = diamond, 4 = triangle). Vice presidents are circled, whereas managers (and the CEO) are not. The size of the point corresponds to the age of the individual (larger implies older), and similarly the shading of the point corresponds to the tenure (darker implies longer tenure). The black (gray) arrow indicates the direction an individual's latent position mean would change if their age (tenure) increased by a standard deviation. The hollow circle is the CEO.

6. Discussion

We have provided a statistical model for better understanding network perception data and how individuals' perceptions differ. CSS data are extremely rich and provide abundant information to not just the relativist who requires multiple perceptions with which to compare but also the essentialist who desires repeated observations. We hope that the proposed approach is useful under either method of interpretation. Specifically, the relativist may better compare the actors' perceptions that are subject to potential informant error, as well as quantitatively investigate how informants' responses deviate from each other through bias and variability. Essentialists may have better approximations of various aggregation structures, estimate a low-dimensional representation of the underlying social reality, estimate homophilic effects on this representation, and estimate informant accuracy.

A comment from an anonymous referee described the potential use of looking at measures of fit such as AUC as a function of the number of respondents K in informing future studies. If there exists a certain threshold K^* above which the information added to the data by obtaining $K > K^*$ is negligible, then this could save considerable time and expense. It would seem highly likely, however, that this threshold would be a function of n and the signal-to-noise ratio (e.g., if θ and σ^2 vary widely between contexts and environments). Determining this threshold (which would involve some criterion for determining what is "negligible addition of information") would

https://doi.org/10.1017/nws.2019.1 Published online by Cambridge University Press

be a valuable but nontrivial task, and one which would likely require either ample prior data or some type of adaptive sampling design to achieve.

Section 3.5 discussed relaxing the commonly used forced-choice design. Although our analyses did not revolve around these types of data, we believe this to be an important future direction for collecting network perception data. We anticipate that such a relaxed design would not only make larger studies feasible but also ought to filter out unnecessary noise that otherwise would have been introduced into the data.

Acknowledgments. I am grateful to the anonymous reviewers whose careful consideration and insightful comments and suggestions have led to considerable improvements upon the original manuscript.

Supplementary. For supplementary material for this article, please visit http://dx.doi.org/10.1017/nws.2019.1.

Conflict of interest. Daniel K. Sewell has nothing to disclose.

Note

1 Note that while for $p = 1, 2$ we did not need many iterations to ensure convergence of the MCMC chain, we did need 3 million iterations for $p = 3$, which required a run time of 56 minutes.

References

Aarstad, J., Selart, M., & Troye, S. (2011). Advice seeking network structures and the learning organization. *Problems and Perspectives in Management*, *9*(2), 63–70.

Adams, J., & Moody, J. (2007). To tell the truth: Measuring concordance in multiply reported network data. *Social Networks*, *29*(1), 44–58.

Almquist, Z. W. (2012). Random errors in egocentric networks. *Social Networks*, *34*(4), 493–505.

An, W., & Schramski, S. (2015). Analysis of contested reports in exchange networks based on actors credibility. *Social Networks*, *40*, 25–33.

Austin, A., Linkletter, C., & Wu, Z. (2013). Covariate-defined latent space random effects model. *Social Networks*, *35*(3), 338–346.

Batchelder, E. (2002). Comparing three simultaneous measurements of a sociocognitive network. *Social Networks*, *24*(3), 261–277.

Batchelder, W. H., & Romney, A. K. (1988). Test theory without an answer key. *Psychometrika*, *53*(1), 71–92.

Batchelder, W. H., & Romney, A. K. (1989). *New results in test theory without an answer key* (pp. 229–248). Berlin, Heidelberg: Springer.

Batchelder, W. H., Kumbasar, E., & Boyd, J. P. (1997). Consensus analysis of three-way social network data. *Journal of Mathematical Sociology*, *22*(1), 29–58.

Bell, D. C., Montoya, L. D., & Atkinson, J. S. (2000). Partner concordance in reports of joint risk behaviors. *Journal of Acquired Immune Deficiency Syndromes*, *25*, 173–181.

Bernard, H. R., Killworth, P. D., & Sailer, L. (1979). Informant accuracy in social network data IV: A comparison of clique-level structure in behavioral and cognitive network data. *Social Networks*, *2*(3), 191–218.

Bernard, H. R., Killworth, P. D., & Sailer, L. (1982). Informant accuracy in social-network data V: An experimental attempt to predict actual communication from recall data. *Social Science Research*, *11*(1), 30–66.

Bond Jr., C. F., Horn, E. M., & Kenny, D. A. (1997). A model for triadic relations. *Psychological Methods*, *2*(1), 79–94.

Bondonio, D. (1998). Predictors of accuracy in perceiving informal social networks. *Social Networks*, *20*(4), 301–330.

Brands, R. A. (2013). Cognitive social structures in social network research: A review. *Journal of Organizational Behavior*, *34*, S82–S103.

Brewer, D. D. (2000). Forgetting in the recall-based elicitation of personal and social networks. *Social Networks*, *22*(1), 29–43.

Brewer, D. D., & Webster, C. M. (2000). Forgetting of friends and its effects on measuring friendship networks. *Social Networks*, *21*(4), 361–373.

Butts, C. T. (2003). Network inference, error, and informant (in)accuracy: A Bayesian approach. *Social Networks*, *25*(2), 103–140.

Butts, C. T. (2008). A relational event framework for social action. *Sociological Methodology*, *38*, 155–200.

Butts, C. T. (2011). Bayesian meta-analysis of social network data via conditional uniform graph quantiles. *Sociological Methodology*, *41*, 257–298.

Casciaro, T. (1998). Seeing things clearly: Social structure, personality, and accuracy in social network perception. *Social Networks*, *20*(4), 331–351.

Casciaro, T., Carley, K. M., & Krackhardt, D. (1999). Positive affectivity and accuracy in social network perception. *Motivation and Emotion*, *23*(4), 285–306.

Davis, J. A., & Leinhardt, S. (1972). Sociological theories in progress. In *The structure of positive interpersonal relations in small groups* (vol. 2). Boston: Houghton Mifflin.

Durante, D., & Dunson, D. B. (2014). Nonparametric Bayes dynamic modelling of relational data. *Biometrika*, *101*(4), 883–898.

Durante, D., Dunson, D. B, & Vogelstein, J. T. (2017). Nonparametric Bayes modeling of populations of networks. *Journal of the American Statistical Association*, *112*(520), 1516–1530.

Freeman, L., & Romney, A. K. (1987). Words, deeds and social structure: A preliminary study of the reliability of informants. *Human Organization*, *46*(4), 330–334.

Freeman, L. C., Romney, A. K., & Freeman, S. C. (1987). Cognitive structure and informant accuracy. *American Anthropologist*, *89*(2), 310–325.

Geweke, J. (1992). Bayesian statistics 4. In *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*. Oxford: Clarendon Press.

Grippa, F., & Gloor, P. A. (2009). You are who remembers you. detecting leadership through accuracy of recall. *Social Networks*, *31*(4), 255–261.

Hammer, M. (1985). Implications of behavioral and cognitive reciprocity in social network data. *Social Networks*, *7*(2), 189–201.

Handcock, M. S, Raftery, A. E, & Tantrum, J. M. (2007). Model-based clustering for social networks. *Journal of the Royal Statistical Society, Series A*, *170*(2), 301–354.

Hildum, D. C. (1986). 'Competence' and 'performance' in network structure. *Social Networks*, *8*(1), 79–95.

Hoff, P. D. (2005). Bilinear mixed-effects models for dyadic data. *Journal of the American Statistical Association*, *100*(469), 286–295.

Hoff, P. D. (2009). Multiplicative latent factor models for description and prediction of social networks. *Computational and Mathematical Organization Theory*, *15*(4), 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*(460), 1090–1098.

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

Johnson, J. C., & Orbach, M. K. (2002). Perceiving the political landscape: Ego biases in cognitive political networks. *Social Networks*, *24*(3), 291–310.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: Guilford Press.

Kilduff, M., Crossland, C., Tsai, W., & Krackhardt, D. (2008). Organizational network perceptions versus reality: A small world after all? *Organizational Behavior and Human Decision Processes*, *107*(1), 15–28.

Killworth, P. D., & Bernard, H. R. (1976). Informant accuracy in social network data. *Human Organization*, *35*(3), 269–286.

Killworth, P. D., & Bernard, H. R. (1977). Informant accuracy in social network data II. *Human Communication Research*, *4*(1), 3–18.

Killworth, P. D., & Bernard, H. R. (1979). Informant accuracy in social network data III: A comparison of triadic structure in behavioral and cognitive data. *Social Networks*, *2*(1), 19–46.

Koskinen, J. H. (2002a). *Bayesian analysis of cognitive social structures with covariates* (Working Paper 2002:3). Department of Statistics.

Koskinen, J. H. (2002b). *Bayesian analysis of perceived social networks* (Working Paper 2002:3). Department of Statistics.

Koskinen, J. H. (2004). *Model selection for cognitive social structures* (Working Paper 2004:3). Department of Statistics.

Krackhardt, D. (1987). Cognitive social structures. *Social Networks*, *9*(2), 109–134.

Krackhardt, D. (1990). Assessing the political landscape: Structure, cognition, and power in organizations. *Administrative Science Quarterly*, *35*(2), 342–369.

Krackhardt, D., & Kilduff, M. (1999). Whether close or far: Social distance effects on perceived balance in friendship networks. *Journal of Personality and Social Psychology*, *76*(5), 770–782.

Krivitsky, P. N., Handcock, M. S., Raftery, A. E., & Hoff, P. D. (2009). Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models. *Social Networks*, *31*(3), 204–213.

Kumbasar, E., Rommey, A. K., & Batchelder, W. H. (1994). Systematic biases in social perception. *American Journal of Sociology*, *100*(2), 477–505.

Lorant, V., Soto, V. E., Alves, J., Federico, B., Kinnunen, J., Kuipers, M., ... Kunst, A. (2015). Smoking in school-aged adolescents: Design of a social network survey in six european countries. *BMC Research Notes*, *8*, 91.

Nakao, K., & Romney, A. K. (1993). Longitudinal approach to subgroup formation: Re-analysis of newcomb's fraternity data. *Social Networks*, *15*(2), 109–131.

Neal, J. W., Neal, Z. P., & Cappella, E. (2014). I know who my friends are, but do you? Predictors of self-reported and peer-inferred relationships. *Child Development*, *85*(4), 1366–1372.

Neal, J. W., Neal, Z. P., & Cappella, E. (2016). Seeing and being seen: Predictors of accurate perceptions about classmates relationships. *Social Networks*, *44*, 1 – 8.

Pattison, P. (1994). Social cognition in context: Some applications of social network analysis. (Chap. 4, pp. 79–109). *Advances in Social Network Analysis*. Thousand Oaks, USA: Sage.
Perry, B. L., & Pescosolido, B. A. (2015). Social network activation: The role of health discussion partners in recovery from mental illness. *Social Science & Medicine*, *125*, 116–128.
Romney, A. K., & Faust, K. (1982). Predicting the structure of a communications network from recalled data. *Social Networks*, *4*, 285–304.
Romney, A. K., Brewer, D. D., & Batchelder, W. H. (1996). The relation between typicality and semantic similarity structure. *Journal of Quantitative Anthropology*, *6*(1–2), 1–14.
Sewell, D. K, & Chen, Y. (2015). Latent space models for dynamic networks. *Journal of the American Statistical Association*, *110*(512), 1646–1657.
Sewell, D. K, & Chen, Y. (2016). Latent space models for dynamic networks with weighted edges. *Social Networks*, *44*, 105–116.
Sewell, D. K, & Chen, Y. (2017). Latent space approaches to community detection in dynamic networks. *Bayesian Analysis*, *12*(2), 351–377.
Shakya, H. B., Christakis, N. A., & Fowler, J. H. (2015). Social network predictors of latrine ownership. *Social Science & Medicine*, *125*, 129–138.
Shoham, D. A., Harris, J. K., Mundt, M., & McGaghie, W. (2016). A network model of communication in an interprofessional team of healthcare professionals: A cross-sectional study of a burn unit. *Journal of Interprofessional Care*, *30*(5), 661–667.
Siciliano, M. D, Yenigun, D., & Ertan, G. (2012). Estimating network structure via random sampling: Cognitive social structures and the adaptive threshold method. *Social Networks*, *34*(4), 585–600.
Sosa, J., & Rodriguez, A. (2018+). A latent space model for cognitive social structures data. *arXiv:1711.03662*.
Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B*, *64*(4), 583–639.
Swartz, T. B, Gill, P. S, & Muthukumarana, S. (2015). A Bayesian approach for the analysis of triadic data in cognitive social structures. *Journal of the Royal Statistical Society, Series C*, *64*(4), 593–610.
Warner, R. M, Kenny, D. A, & Stoto, M. (1979). A new round robin analysis of variance for social interaction data. *Journal of Personality and Social Psychology*, *37*(10), 1742–1757.
Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.

Cite this article: Sewell D. K. (2019). Latent space models for network perception data. *Network Science* **7**, 160–179. https://doi.org/10.1017/nws.2019.1