Network analysis provides a way to represent and study "relational data" with characteristics extending beyond those of the individual. Data structures that extend beyond the country level are quite simply the norm in international relations. The dominant paradigm in international relations for dealing with such data structures, however, is not a network approach but rather a dyadic design, in which an interaction between a pair of countries is considered independent of interactions between any other pair in the system.

The implication of this assumption is that when, for example, Vietnam and the United States decide to form a trade agreement, they make this decision independently of what they have done with other countries and what other countries in the international system have done among themselves.[1] An even stronger assumption is that Japan declaring war against the United States is independent of the decision of the United States to go to war against Japan. A common defense of the dyad-only approach is that many events are only bilateral (Diehl and Wright, 2016 in press), thus alleviating the need for an approach that incorporates interdependencies between observations. This is clearly wrong. The network perspective asserts that even bilateral events and processes take place within a broader system. What takes place in one part of the system may be dependent upon events in another. At a minimum, we don't know whether independence of events and processes characterizes what we observe. We should at least examine this assertion.

The potential for interdependence among observations poses a challenge to theoretical as well as statistical modeling since the assumption made by standard approaches used across the social sciences is that observations are, at least, conditionally independent (Snijders, 2011). The consequence of ignoring this assumption has been frequently noted within the political science literature already.[2] Just as relevant is the fact that a wealth of research from other disciplines suggests that carrying the independence assumption into a study with relational data is misguided and most often leads to biased inferences.

Despite the hesitation among some in the discipline to adopt network analytic approaches, in recent years there has been a greater level of interest in understanding these approaches. For instance, in the past year a special issue focused on the application of a variety of network approaches has come out in the *Journal of Peace Research*. Particularly notable is a recent overview and comparison of a handful of network based inferential models by Cranmer et al. (2016). Specifically, they focus on the exponential random graph model (ERGM), the multiple regression quadratic assignment procedure (MRQAP), and a latent distance approach developed by Hoff et al. (2002). However, their discussion overlooks more than a decade worth of developments.[3] The principal latent

[1] There has been plenty of work done on treaty formation that would challenge this claim, e.g., see Manger et al. (2012); Kinne (2013).

[2] For example, see Beck et al. (1998); Signorino (1999); Hoff and Ward (2004); Erikson et al. (2014).

[3] Indeed, in so far as we can tell, very few in political science have actually employed the Euclidean

variable approach used in political science has been the general bilinear mixed-effects (GBME) model developed by Hoff (2005). Examples of political science applications of the GBME model include Ward et al. (2007); Cao (2012); Breunig et al. (2012); Metternich et al. (2015); Greenhill (2015).We are aware of only one political science application using the latent distance approach (Kirkland, 2012). As Hoff (2008) shows both empirically and mathematically, the distinction between the latent distance and latent factor models, such as the GBME model, is consequential when accounting for higher-order interdependencies, a point overlooked by Cranmer et al. (2016).

In this article, we introduce the additive and multiplicative effects model (AME). To highlight the benefits of this approach, we estimate this model using data from the application presented in Cranmer et al. (2016) and compare it to the other models presented in that article. By doing so we are able to show that AME provides a superior goodness of fit to the data than alternative approaches.[4] Further, through the AME approach we can estimate many different types of cross-sectional and longitudinal relational data (e.g., binomial, gaussian, and ordinal edges) in a straightforward way. The AME modeling framework provides a flexible framework to study relational data. It addresses the issue of interdependence while still allowing scholars to examine theories that may only be relevant in the monadic or dyadic level. It accounts for both nodal and dyadic dependence patterns, and can include higher-order dependencies.

1. Addressing Dependencies in Dyadic Data

Relational, or dyadic, data provide measurements of how pairs of actors relate to one another. The easiest way to organize such data is the directed dyadic design in which the unit of analysis is some set of n actors that have been paired together to form a dataset of z directed dyads. A tabular design such as this for a set of n actors, $\{i, j, k, l\}$ results in $n \times (n-1)$ observations, as shown in Table 1.

1.1. Limitations of the Standard Framework

When modeling relational data, scholars typically employ a generalized linear model (GLM). This type of model is typically expressed via a stochastic and systematic component. The stochastic component reflects assumptions about the probability distribution from which the data are generated: $y_{ij} \sim P(Y|\theta_{ij})$, with a probability density or mass function such as the normal, binomial, or Poisson. Each dyad in the sample is independently drawn from a particular distribution, given θ_{ij}. The systematic component characterizes the model for the parameters of that distribution and describes how θ_{ij} varies as a function of a set of nodal and dyadic covariates, $\mathbf{X}_{ij}$: $\theta_{ij} = \boldsymbol{\beta}^T \mathbf{X}_{ij}$. A fundamental assumption we make when applying this modeling technique is that given $\mathbf{X}_{ij}$

approach they summarize.

[4]The AME approach has been developed into an R package named **amen** and is available on CRAN (Hoff et al., 2015). Hoff (2015) provides a vignette for this package as well.

Sender	Receiver	Event
i	j	y_{ij}
	k	y_{ik}
$\vdots$	l	y_{il}
j	i	y_{ji}
	k	y_{jk}
$\vdots$	l	y_{jl}
k	i	y_{ki}
	j	y_{kj}
$\vdots$	l	y_{kl}
l	i	y_{li}
	j	y_{lj}
$\vdots$	k	y_{lk}

Table 1: Structure of datasets used in canonical design.

$\longrightarrow$

	i	j	k	l
i	NA	y_{ij}	y_{ik}	y_{il}
j	y_{ji}	NA	y_{jk}	y_{jl}
k	y_{ki}	y_{kj}	NA	y_{kl}
l	y_{li}	y_{lj}	y_{lk}	NA

Table 2: Adjacency matrix representation of data in Table 1. Senders are represented by the rows and receivers by the columns.

and the parameters of the distribution, each of the dyadic observations is conditionally independent.

The importance of this assumption becomes clearer in the process of estimating a GLM via maximum likelihood. After having chosen a set of covariates and specifying a distribution, the joint density function over all dyads using the observations from Table 1 is given:

$$P(y_{ij}, y_{ik}, \ldots, y_{lk}|\theta_{ij}, \theta_{ik}, \ldots, \theta_{lk}) = P(y_{ij}|\theta_{ij}) \times P(y_{ik}|\theta_{ik}) \times \ldots \times P(y_{lk}|\theta_{lk})$$
$$P(\mathbf{Y} \mid \boldsymbol{\theta}) = \prod_{\alpha=1}^{n\times(n-1)} P(y_\alpha|\theta_\alpha) \tag{1}$$

The joint probability is converted into a likelihood: $\mathcal{L}(\boldsymbol{\theta}|\mathbf{Y}) = \prod_{\alpha=1}^{n\times(n-1)} P(y_\alpha|\theta_\alpha)$.

The parameters are estimated by maximizing the likelihood. The likelihood as defined above is only valid if y_{ij} is independent of y_{ji} and y_{ik} given the set of covariates we specified, or the values of θ_{ij}. Assuming that the dyad y_{ij} is conditionally independent of the dyad y_{ji} asserts that there is no level of reciprocity in a dataset, an assumption that in many cases would seem quite untenable. A harder problem to handle is the assumption that y_{ij} is conditionally independent of y_{ik}, the difficulty here follows from the possibility that i's relationship with k is dependent on how i relates to j and how j relates to k, or more simply put the "enemy of my enemy [may be] my friend".

The presence of these types of interdependencies in relational data complicates the *a priori* assumption of observational independence. Without this assumption the joint density function cannot be written in the way described above and a valid like-

lihood does not exist.[5] Accordingly, inferences drawn from misspecified models that ignore potential interdependencies between dyadic observations are likely to have a number of issues including biased estimates of the effect of independent variables, uncalibrated confidence intervals, and poor predictive performance.[6] By ignoring these interdependencies, we ignore a potentially important part of the data generating process behind relational data, namely, network phenomena.

1.2. Social Relations Regression Model: Additive Part of AME

The dependencies that tend to develop in relational data can be more easily understood when we move away from stacking dyads on top of one another and turn instead to adjacency matrices as shown in Table 2. Operationally, this type of data structure is represented as a $n \times n$ matrix, $\mathbf{Y}$, where the diagonals in the matrix are typically undefined. The ij^{th} entry defines the relationship between i and j and can be continuous or discrete. For example, in undirected data an event cannot be attributed to a specific sender or receiver rather it is just an indication of something that happened between a pair of countries or a relationship they share (e.g., two countries might have mutually agreed to form an alliance). If the relationship is undirected, the ji^{th} entry will equal the ij^{th} entry. Sociomatrices of directed relations are not symmetric, there is a specific sender and receiver, as in the case of bilateral or multilateral aid.

A common type of structural interdependency that arises in relational data is "preferential attachment" (Barabási and Réka, 1999). This is typically categorized as a first-order, or nodal, dependency and represents the fact that we typically find significant heterogeneity in activity levels across nodes. The implication of this across-node heterogeneity is within-node heterogeneity of ties, meaning that values across a row, say $\{y_{ij}, y_{ik}, y_{il}\}$, will be more similar to each other than other values in the adjacency matrix because each of these values has a common sender i. This type of dependency manifests in cases where sender i tends to be more active or less active in the network than other senders. The emergence of this type of structure often occurs in relational datasets such as trade and conflict. In both of these cases, there are a set of countries that tend to be more active than others. Similarly, while some actors may be more active in sending ties to others in the network, we might also observe that others are more popular targets, this would manifest in observations down a column, $\{y_{ji}, y_{ki}, y_{li}\}$, being more similar. Last, we might also find that actors who are more likely to send ties in a network are also more likely to receive them, meaning that the row and column means of an adjacency matrix may be correlated. First-order dependencies are equally important to take into account in undirected relational structures, the only difference being that nodal heterogeneity will be equivalent across rows and columns. The pres-

[5] This problem has been noted in works such as Lai (1995); Manger et al. (2012); Kinne (2013).

[6] In cases where there is only "dyadic clustering" and no higher-order network effects such as transitivity, Aronow et al. (2015) show that a nonparametric, variance estimator can be used to deal with the statistical issues that arise when working with dyadic data.

ence of this type of heterogeneity in directed and undirected relational data leads to a violation of the conditional independence assumption underlying the models in our standard tool-kit, but can be easily accommodated in the GLM framework with the inclusion of additive sender and receiver random effects.

Another ubiquitous type of structural interdependency is reciprocity. This is a second-order, or dyadic, dependency relevant only to directed datasets, and asserts that values of y_{ij} and y_{ji} may be statistically dependent. In studies of social and economic behavior, direct reciprocity–the notion that actors learn to "respond in kind" to one another– is argued to be an essential component of behavior. This concept has deep roots in political science (Richardson, 1960). The clearest example of the relevance of this dependency comes from the conflict literature, as we would expect that if, for instance, Iran behaved aggressively towards Saudi Arabia that this would induce Saudi Arabia to behave aggressively in return. The prevalence of these types of potential interactions within directed dyadic data also complicates the basic assumption of observational independence.

The relevance of modeling first- and second-order dependencies has long been recognized within some social sciences particularly in psychology. Warner et al. (1979) developed the social relational model (SRM), a type of ANOVA decomposition technique, that facilitates this undertaking. The SRM is of particular note as it provides the error structure for the additive effects component of the AME framework that we introduce here. The goal of the SRM is to decompose the variance of observations in an adjacency matrix in terms of heterogeneity across row means (out-degree), heterogeneity along column means (in-degree), correlation between row and column means, and correlations within dyads. Wong (1982) and Li and Loken (2002) provide a random effects representation of the SRM:

$$
\begin{aligned}
y_{ij} &= \mu + e_{ij} \\
e_{ij} &= a_i + b_j + \epsilon_{ij} \\
\{(a_1, b_1), \ldots, (a_n, b_n)\} &\overset{\text{iid}}{\sim} N(0, \Sigma_{ab}) \\
\{(\epsilon_{ij}, \epsilon_{ji}) : \; i \neq j\} &\overset{\text{iid}}{\sim} N(0, \Sigma_\epsilon), \text{ where} \\
\Sigma_{ab} = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix} \quad \Sigma_\epsilon &= \sigma_\epsilon^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}
\end{aligned}
\tag{2}
$$

The basic idea here is simple, μ provides a baseline measure of the density or sparsity of a network, and e_{ij} represents residual variation. The residual variation decomposes into parts: a row/sender effect (a_i), a column/receiver effect (b_j), and a within-dyad effect (ϵ_{ij}). The row and column effects are modeled jointly to account for correlation in how active an actor is in sending and receiving ties. Heterogeneity in the row and column means is captured by σ_a^2 and σ_b^2, respectively, and σ_{ab} describes the linear relationship between these two effects (i.e., whether actors who send [receive] a lot of

ties also receive [send] a lot of ties). Beyond these first-order dependencies, second-order dependencies are described by σ_ϵ^2 and a within dyad correlation, or reciprocity, parameter ρ.

The SRM covariance structure described in Equation 2 can be incorporated into the systematic component of a GLM framework to produce the social relations regression model (SRRM): $\boldsymbol{\beta}^T\mathbf{X}_{ij}+a_i+b_j+\epsilon_{ij}$, where $\boldsymbol{\beta}^T\mathbf{X}_{ij}$ accommodates the inclusion of dyadic, sender, and receiver covariates. This approach (Hoff, 2005) incorporates row, column, and within-dyad dependence in way that is widely used and understood by applied researchers: a regression framework and additive random effects to accommodate variances and covariances often seen in relational data. Furthermore, this handles a diversity of outcome distributions (e.g., binomial, ordinal, etc.). In the case of binary data this can be done by utilizing a latent variable representation of a probit regression model. This approach can also easily incorporate ordinal and rank-ordered data.

1.3. Latent Factor Model: Multiplicative Part of AME

Missing from the framework provided by the SRM is an accounting of third-order dependence patterns that can arise in relational data. The ubiquity of third-order effects in relational datasets arises from the presence of some set of shared attributes between nodes that affects their probability of interacting with one another. For example, one finding from the gravity model of trade is that neighboring countries are more likely to trade with one another; in this case, the shared attribute is simply geographic proximity. A finding common in the political economy literature is that democracies are more likely to form trade agreements with one another, and the shared attribute here is a country's political system. Both geographic proximity and a country's political system are examples of homophily, which captures the idea that the relationships between actors with similar characteristics in a network are likely to be stronger than nodes with different characteristics.

A binary network where actors tend to form ties with others based on some set of shared characteristics often leads to a network graph with a high number of "transitive triads" in which sets of actors $\{i, j, k\}$ are each linked to each another. The left-most plot in Figure 1 provides a representation of a network that exhibits this type of pattern. Such structures develop when the interactions between actors result from some set of shared attributes those actors may possess. The relevant implication of this when it comes to conducting statistical inference is that–unless we are able to specify the list of exogenous variable that may explain this prevalence of triads–the probability of j and k forming a tie is not independent of the ties that already exist between those actors and i.

Another third-order dependence pattern that cannot be accounted for in the additive effects framework is stochastic equivalence. A pair of actors ij are stochastically equivalent if the probability of i relating to, and being related to, by every other actor is the same as the probability for j. This refers to the idea that there will be groups of nodes in a network with similar relational patterns. The occurrence of a dependence

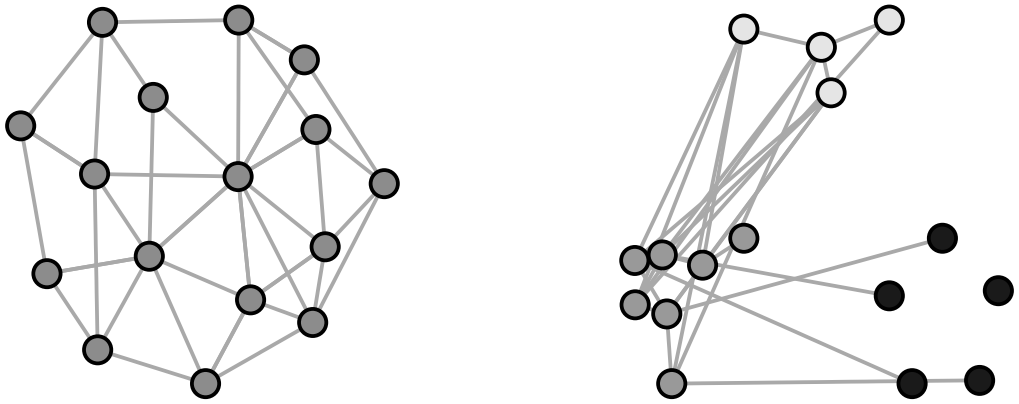

Figure 1: Graph on the left is a representation of an undirected network that exhibits a high degree of homophily, while on the right we show an undirected network that exhibits stochastic equivalence.

pattern such as this is not uncommon in the social science applications. Manger et al. (2012) posit and estimate a stochastic equivalence structure to explain the formation of preferential trade agreements (PTAs). Specifically, they suggest that PTA formation is related to differences in per capita income levels between countries. Countries falling into high, middle, and low income per capita levels will have patterns of PTA formation that are determined by the groups into which they fall. Such a structure is represented in the right-most panel of Figure 1, here the lightly shaded group of nodes at the top can represent high-income countries, nodes on the bottom-left middle-income, and the darkest shade of nodes low-income countries. The behavior of actors in a network can at times be governed by group level dynamics, and failing to account for such dynamics leaves potentially important parts of the data generating process ignored.

If we are able to explicitly model the variety of shared attributes that might cause third-order dependence patterns to develop, then the additive effects framework of the SRRM is likely enough to justify the conditional independence assumption that is central to the GLM framework. In the context of most observational research, however, the assumption that we have included all relevant explanatory variables is untenable. The implausibility of this assumption is, in spirit, the same reason why we no longer model time-series cross-sectional data without accounting for the temporal structure of the data.

1.3.1. ***Latent Variable Models***

To account for third-order dependence patterns within the context of the SRRM we turn to latent variable models, which have become a popular approach for modeling relational data in fields as diverse as biology to computer science to the social sciences. These models assumes that relationships between nodes are mediated by a small number (K) of node-specific unobserved latent variables. One reason for their increased usage is that they enable researchers to capture and visualize third-order

dependencies in a way that other approaches are not able to replicate. Additionally, the conditional independence assumption eliminates the model degeneracy issue, facilitates the testing of a variety of nodal and dyadic level theories, and provides a range of computational advantages (Hunter et al., 2012).

A number of major latent variable approaches have been developed to represent third-order dependencies in relational data, we focus on two here: the latent distance model and the latent factor model.[7] For the sake of exposition, we consider the case where relations are symmetric to describe the differences between these approaches. Both of these approaches can be incorporated into an undirected version of the framework that we have been constructing through the inclusion of an additional term to the model for y_{ij}, $\alpha(u_i, u_j)$, that captures latent third-order characteristics of a network, where u_i and u_j are node-specific latent variables. General definitions for how $\alpha(u_i, u_j)$ is defined for these latent variable models are shown in Equations 3. One other point of note about these approaches is that researchers have to specify a value for K. In the case of the latent distance and factor models, a value of K equal to two or three is typically large enough to account for third-order dependencies in relational data.

$$
\begin{aligned}
&\text{Latent distance model} \\
&\qquad \alpha(\mathbf{u}_i, \mathbf{u}_j) = -|\mathbf{u}_i - \mathbf{u}_j| \\
&\qquad \mathbf{u}_i \in \mathbb{R}^K,\ i \in \{1, \ldots, n\} \\
&\text{Latent factor model} \\
&\qquad \alpha(\mathbf{u}_i, \mathbf{u}_j) = \mathbf{u}_i^T \Lambda \mathbf{u}_j \\
&\qquad \mathbf{u}_i \in \mathbb{R}^K,\ i \in \{1, \ldots, n\} \\
&\qquad \Lambda \text{ a } K \times K \text{ diagonal matrix}
\end{aligned}
\tag{3}
$$

The latent distance model was developed by Hoff et al. (2002) to capture homophily. In this approach, each node i has some unknown latent position in K dimensional space, $\mathbf{u}_i \in \mathbb{R}^K$, and the probability of a tie between a pair ij is a function of the negative Euclidean distance between them: $-|\mathbf{u}_i - \mathbf{u}_j|$. Hoff et al. (2002) show that because latent distances for a triple of actors obey the triangle inequality, this formulation models the tendencies toward homophily commonly found in social networks. This approach has been operationalized in the **latentnet** package developed by Krivitsky and Handcock (2015). However, this approach also comes with an important shortcoming: it confounds stochastic equivalence and homophily. Consider two nodes i and j that are proximate to one another in K dimensional Euclidean space, this suggests not only

[7] Though latent distance models have become a popular modeling tool in some disciplines (Salter-Townshend et al., 2012), we are aware of only one publication that has used this approach in political science, see Kirkland (2012). The bi-linear latent space approach, however, has been used in a variety of works in political science.

that $|\mathbf{u}_i - \mathbf{u}_j|$ is small but also that $|\mathbf{u}_i - \mathbf{u}_l| \approx |\mathbf{u}_j - \mathbf{u}_l|$, the result being that nodes i and j will by construction assumed to possess the same relational patterns with other actors such as l (i.e., that they are stochastically equivalent).Thus latent distance models confound strong ties with stochastic equivalence. This approach cannot adequately model data with many ties between nodes that have different network roles.

An early iteration of the latent factor approach was presented in Hoff (2005) and introduced to political science by Hoff and Ward (2004), but the revised approach is motivated by an eigenvalue decomposition of a network.[8] The motivation for this alternative framework stems from the fact that many real networks exhibit varying degrees of stochastic equivalence and homophily. In these situations, using either the latent distance or class model would end up representing only a part of the network structure. In the latent factor model, each actor has an unobserved vector of characteristics, $\mathbf{u}_i = \{u_{i,1}, \ldots, u_{i,K}\}$, which describe their behavior as an actor in the network. The probability of a tie from i to j depends on the extent to which $\mathbf{u}_i$ and $\mathbf{u}_j$ are "similar" (i.e., point in the same direction) and on whether the entries of Λ are greater than or less than zero.

More specifically, the similarity in the latent factors, $\mathbf{u}_i \approx \mathbf{u}_j$, corresponds to how stochastically equivalent a pair of actors are and the eigenvalue determines whether the network exhibits positive or negative homophily. For example, say that that we estimate a rank-one latent factor model (i.e., $K = 1$), in this case $\mathbf{u}_i$ is represented by a scalar $u_{i,1}$, similarly, $\mathbf{u}_j = u_{j,1}$, and Λ will have just one diagonal element λ. The average effect this will have on y_{ij} is simply $\lambda \times u_i \times u_j$, where a positive value of $\lambda > 0$ indicates homophily and $\lambda < 0$ anti-homophily. Hoff (2008) shows that such a model can represent both homophily and stochastic equivalence, and that the alternative latent variable approaches can be represented as a latent factor model but not vice versa. In the directed version of this approach, we use the singular value decomposition,[9] here actors in the network have a vector of latent characteristics to describe their behavior as a sender, denoted by $\mathbf{u}$, and as a receiver, $\mathbf{v}$: $\mathbf{u}_i, \mathbf{v}_j \in \mathbb{R}^K$ (Hoff, 2009). These again can alter the probability, or in the continuous case value, of an interaction between ij additively: $\mathbf{u}_i^T \mathbf{D} \mathbf{v}_j$, where $\mathbf{D}$ is an $n \times n$ diagonal matrix. The latent factor model is incorporated into the AME approach as a multiplicative effect to account for third-order dependencies (Hoff, 2009; Hoff et al., 2015).

Merging either of these approaches into the additive effects probit framework is possible through the addition of a term that captures third-order interdependencies. In the **latentnet** package this is done by directly incorporating $|\mathbf{u}_i - \mathbf{u}_j|$ as a fixed effect:

[8] An important difference in the earlier approaches such as the GBME compared to the model that we present here is that Λ was taken to be the identity matrix thus stochastic equivalence could not be characterized. This approach should also not be confused with the projection model introduced in Hoff et al. (2002).

[9] The singular value decomposition is a model based analogue to the eigenvalue decomposition for directed networks.

$\theta_{ij} = \boldsymbol{\beta}^T \mathbf{X}_{ij} - |\mathbf{u}_i - \mathbf{u}_j|$.[10] However, incorporating the term in this way can affect our estimation of the linear relationship between the exogenous nodal and dyadic covariates. This results from collinearity between that set of exogenous attributes and the nodal positions of actors in the latent space. The intuition behind why collinearity occurs is not surprising given our discussion above. The latent space is essentially used to capture dependencies that can result from shared attributes between nodes. Thus if a particular exogenous covariate is actually predictive of relations between i and j, due to homophily, this effect will be correlated with the nodal positions of actors in a K dimensional Euclidean space. Additionally, interpretation of exogenous covariates is not as straightforward because the coefficients will be modeling a "max value" between nodes, if nodes had the same latent position. In the latent factor framework, this is not an issue because each of the random effect terms used to account for interdependencies has a mean of zero. The parameter estimates for the exogenous covariates from the latent factor approach can be interpreted as the average effect they have on the dependent variable after having accounted for network dependencies. The AME approach considers the regression model shown in Equation 4:

$$\begin{aligned} y_{ij} &= g(\theta_{ij}) \\ \theta_{ij} &= \boldsymbol{\beta}^T \mathbf{X}_{ij} + e_{ij} \\ e_{ij} &= a_i + b_j + \epsilon_{ij} + \alpha(\mathbf{u}_i, \mathbf{v}_j) \text{ , where} \\ \alpha(\mathbf{u}_i, \mathbf{v}_j) &= \mathbf{u}_i^T \mathbf{D} \mathbf{v}_j = \sum_{k \in K} d_k u_{ik} v_{jk} \end{aligned} \tag{4}$$

Using this framework, we are able to model the dyadic observations as conditionally independent given $\boldsymbol{\theta}$, where $\boldsymbol{\theta}$ depends on the the unobserved random effects, $\mathbf{e}$. $\mathbf{e}$ is then modeled to account for the potential first, second, and third-order dependencies that we have discussed. As described in Equation 2, $a_i + b_j + \epsilon_{ij}$, are the additive random effects in this framework and account for sender, receiver, and within-dyad dependence. The multiplicative effects, $\mathbf{u}_i^T \mathbf{D} \mathbf{v}_j$, are used to capture higher-order dependence patterns that are left over in $\boldsymbol{\theta}$ after accounting for any known covariate information. Thus the third-order interdependencies captured in the latent factor space of AME are those that could not have been explained by the exogenous nodal and dyadic covariates that have already been included in the model, or the additive row and column random effects. A Bayesian procedure in which parameters are iteratively updated using a Gibbs sampler is available in the **amen** package to estimate this type of generalized linear mixed effects model from continuous, binary, ordinal, and other

[10] The **latentnet** package also allows for the specification of a bilinear latent space that is closely related to the projection model introduced in Hoff et al. (2002).

relational data types.[11]

Taken together, the additive effects portion of AME (described by the SRM) and the multiplicative effects (described by the latent factor model) provide a modeling framework similar to the GLMs that many scholars currently use, and has the benefit of being able to not only deal with interdependencies in relational data but also provide explicit estimates of these dependencies after having taken into account observable information. Specifically, we can obtain degree based effects for actors in the network, the level of reciprocity between actors, and also visualize the third-order interdependencies that remain in the data. This latter point is important to note as effectively using these visualizations may also help users of this approach to determine whether or not the inclusion of some other dyadic or nodal variable is necessary to accounting for patterns such as homophily or stochastic equivalence.

1.3.2. ***ERGMs***

An alternative approach to accounting for third-order dependence patterns are ERGMs. ERGM approaches are useful when researchers are interested in the role that a specific list of network statistics have in giving rise to a certain network. These network statistics could include the number of transitive triads in a network, balanced triads, reciprocal pairs and so on.[12] In the ERGM framework, a set of statistics, $S(\mathbf{Y})$, define a model. Given the chosen set of statistics, the probability of observing a particular network dataset $\mathbf{Y}$ can be expressed as:

$$\Pr(Y = y) = \frac{\exp(\boldsymbol{\beta}^T S(y))}{\sum_{z \in \mathcal{Y}} \exp(\boldsymbol{\beta}^T S(z))}, y \in \mathcal{Y} \tag{5}$$

$\boldsymbol{\beta}$ represents a vector of model coefficients for the specified network statistics, $\mathcal{Y}$ denotes the set of all obtainable networks, and the denominator is used as a normalizing factor (Hunter et al., 2008). This approach provides a way to state that the probability of observing a given network depends on the patterns that it exhibits, which are operationalized in the list of network statistics specified by the researcher. Within this approach one can test the role that a variety of network statistics play in giving rise to a particular network.

One issue that arises when conducting statistical inference with this model is in the calculation of the normalizing factor, which is what ensures that the expression above corresponds to a legitimate probability distribution. For even a trivially sized di-

[11] The set of parameters that are estimated in the model from the observed data, $\{\mathbf{Y}, \mathbf{X}\}$, are: latent Gaussian variables ($\boldsymbol{\theta}$); nodal and/or dyadic regression coefficients ($\boldsymbol{\beta}$); additive nodal random effects ($\{(a_i, b_i)\} \in \{i = 1, \ldots, n\}$); network covariance ($\Sigma_{ab}$, Σ_{ϵ}); multiplicative effects ($\mathbf{U}$, $\mathbf{V}$, and $\mathbf{D}$). Further details on this process can be found in Hoff (2005) and Hoff (2009).

[12] Snijders et al. (2006) provides a detailed list of network statistics that can be included in an ERGM model specification.

rected network that has only 20 actors, calculating the denominator means summing over $2^{20\times(20-1)} = 2^{380}$ possible networks, or, to put it another way, more than the total number of atoms in the universe. One of the first approaches to deal with this issue was a computationally fast pseudo-likelihood approach developed by Strauss and Ikeda (1990). However, this approach ignores the interdependent nature of observations in relational data, as a result, many have argued that the standard errors remain unreliable (Van Duijn et al., 2009). The pseudo-likelihood approach has became increasingly unpopular in recent years among those in the network analysis community, particularly, as simulation based techniques have developed–though it has not disappeared. One favored approach in the literature is to approximate the MLE using Markov Chain Monte Carlo techniques, also referred to as MCMC-MLE.

The MCMC-MLE approach is an advancement but notable problems remain. Chatterjee and Diaconis (2013) have shown that MCMC procedures can take an exponential time to converge for broad classes of ERGMs unless the dyadic observations are independent. This is a result of the fact that MCMC procedures visit an infinitesimally small portion of the set of possible graphs. A related issue when estimating ERGMs is that the estimated model can become degenerate even if the observed graph is not degenerate. This means that the model is placing a large amount of probability on a small subset of networks that fall in the set of obtainable networks, $\mathcal{Y}$, but share little resemblance with the observed network (Schweinberger, 2011). Some have argued that model degeneracy is simply a result of model misspecification (Goodreau et al., 2008; Handcock et al., 2008). This points to an important caveat in interpreting the implications of an often cited basis for ERGM, the Hammersley-Clifford theorem. Though this theorem ensures that any network can be represented through an ERGM, it says nothing about the complexity of the sufficient statistics ($S(y)$) required to do so. Failure to properly account for higher-order dependence structures through an appropriate specification can at best lead to model degeneracy, which provides an obvious indication that the specification needs to be altered, and at worst deliver a result that converges but does not appropriately capture the interdependencies in the network. The consequence of the latter case is a set of inferences that will continue to be biased as a result of unmeasured heterogeneity, thus defeating the major motivation for pursuing an inferential network model in the first place.

In the following section we undertake a comparison of the latent distance model, ERGM, and the AME model using an application chosen by Cranmer et al. (2016).In doing so, we are able to highlight the benefits that the AME model provides over alternatives.

2. Empirical Comparison

Cranmer et al. (2016, p. 8) note that scholars must model third-order effects and "must also specify them in a complete and correct manner" or the ERGM model will be misspecified. To avoid providing an incorrect specification when comparing ERGM, we use the specification that they stipulated as theoretically correct. Their application

utilizes a cross-sectional network measuring whether an actor indicated that they collaborated with another during the policy design of the Swiss CO_2 act (Ingold, 2008).[13] The Swiss government proposed this act in 1995 with the goal of undertaking a 10% reduction in CO_2 emissions by 2012. The act was accepted in the Swiss Parliament in 2000 and implemented in 2008. Ingold (2008), and subsequent work by Ingold and Fischer (2014), sought to determine what drives collaboration among actors trying to affect climate change policy. The set of actors included in this network are those that were identified by experts as holding an important position in Swiss climate policy.[14] In total, Ingold (2008) identifies 34 relevant actors: five state actors, eleven industry and business representatives, seven environmental NGOs and civil society organizations, five political parties, and six scientific institutions and consultants.

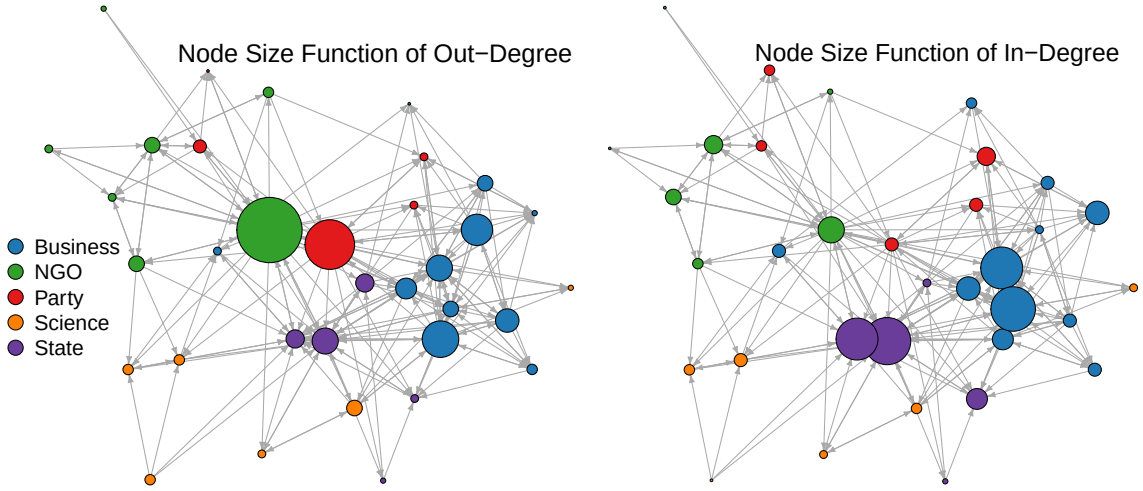

Figure 2: Network visualizations of the Swiss climate change mitigation network. Nodes are colored by type of actor, and directed edges indicate relationships between actors. The network on the left weights node size by the number of out-going ties, and on the right the number of incoming-ties.

Figure 2 provides visualizations for this directed collaboration network. Nodes are colored by the type of actor and a directed edge indicates an actor stated that they collaborated with another, and determining which actor indicated the collaboration can be ascertained by the direction of the arrow. The majority of industry and business actors are clustering together, meaning that these types of actors tend to indicate they collaborated with one another during the policy design process. Three of the state actors are pushed towards the center of the graph because they share relationships

[13] This is a directed relational matrix as an actor i can indicate that they collaborated with j but j may not have stated that they collaborated with i.

[14] For further details on the methodology utilized in choosing the set of actors see Ingold (2008); Ingold and Fischer (2014).

with many actors in the network. Most of the actors classified as scientific institutions are pushed towards the far left border of the graphs as it seems they tend to interact among themselves and just a few of the other actors.

To illustrate nodal heterogeneity in the case of the Swiss climate change mitigation networks we weight the size of nodes, in the network on the left, by the number of their outgoing ties, and on the right by their incoming ties. From the network on the left, we can see that each of the scientific institutions and consultants shown in Figure 2 indicate that they collaborate with relatively few organizations, especially, in comparison with actors from industry and business. Additionally, there is even variation within actor types as evidenced by differences amongst NGO or political party actors. Similar findings of nodal heterogeneity emerge if we turn our attention to examining nodes by their incoming ties.

To understand what factors may play a role in shaping collaboration in this relational data structure a modeling approach is necessary. Cranmer et al. (2016) follow Ingold and Fischer (2014) in developing a model specification. We do not review the specification in detail here, instead we just provide a summary of the variables to be included and the theoretical expectations of their effects in Table **??**.

2.1. Parameter Estimates

Using the specification described in Table **??** we compare five different modeling approaches. The first four approaches chosen here, as in Cranmer et al. (2016), are a logistic regression model, MRQAP, ERGM, and a latent space model (LSM) in which third-order dependencies are accounted for via a two-dimensional Euclidean distance metric.[15] Parameter estimates for these four approaches are shown in Table 3.

The fifth column shows the results from using the additive and multiplicative effects model (AME), in which we account for nodal and dyadic heterogeneity using the SRM and third-order effects represented by a latent factor approach in which we set $K = 2$.[16] Cranmer et al. (2016) provide a lengthy discussion of the differences between the first four modelling approaches that we will not repeat here. More relevant for us are how parameter estimates from AME relate to other approaches. The first point to note is that, in general, the parameter estimates returned by the AME are similar to those of MRQAP and ERGM but quite different from the LSM. For example, while the LSM returns a result for the `Opposition/alliance` variable that diverges from MRQAP and ERGM, the AME returns a result that is not only similar to those approaches but in line with the theoretical expectations of Ingold and Fischer (2014). Similar discrepancies between LSM and other approaches appear for parameters such as `Influence attribution` and `Alter's influence degree`. Each of these discrepancies are resolved when using AME. In part, this is a function of how the LSM approach as operationalized

[15] For a detailed discussion on the MRQAP see Dekker et al. (2007).

[16] Table **??** in Section **??** of the Appendix shows that the parameter estimates presented here for the AME model remain very similar no matter the K chosen.

in the **latentnet** package can confound the effects of covariates with the latent space metric.[17]

[17] As shown in Table **??** in Section **??** of the Appendix, these differences persist even when incorporating sender and receiver random effects or when switching to a bilinear approach to handle third-order dependencies.

	Logit	MRQAP	LSM	ERGM	AME
Intercept/Edges	-4.44*	-4.24*	0.94*	-12.17*	-3.39*
	(0.34)		[0.09; 1.82]	(1.40)	[-4.38; -2.50]
Conflicting policy preferences					
Business vs. NGO	-0.86	-0.87*	-1.37*	-1.11*	-1.37*
	(0.46)		[-2.42; -0.41]	(0.51)	[-2.44; -0.47]
Opposition/alliance	1.21*	1.14*	0.00	1.22*	1.08*
	(0.20)		[-0.40; 0.39]	(0.20)	[0.72; 1.47]
Preference dissimilarity	-0.07	-0.60	-1.76*	-0.44	-0.79*
	(0.37)		[-2.62; -0.90]	(0.39)	[-1.55; -0.08]
Transaction costs					
Joint forum participation	0.88*	0.75*	1.51*	0.90*	0.92*
	(0.27)		[0.86; 2.17]	(0.28)	[0.40; 1.47]
Influence					
Influence attribution	1.20*	1.29*	0.08	1.00*	1.09*
	(0.22)		[-0.40; 0.55]	(0.21)	[0.69; 1.53]
Alter's influence indegree	0.10*	0.11*	0.01	0.21*	0.11*
	(0.02)		[-0.03; 0.04]	(0.04)	[0.07; 0.15]
Influence absolute diff.	-0.03*	-0.06*	0.04	-0.05*	-0.07*
	(0.02)		[-0.01; 0.09]	(0.01)	[-0.11; -0.03]
Alter = Government actor	0.63*	0.68	-0.46	1.04*	0.55
	(0.25)		[-1.08; 0.14]	(0.34)	[-0.07; 1.15]
Functional requirements					
Ego = Environmental NGO	0.88*	0.99	-0.60	0.79*	0.67
	(0.26)		[-1.32; 0.09]	(0.17)	[-0.38; 1.71]
Same actor type	0.74*	1.12*	1.17*	0.99*	1.04*
	(0.22)		[0.63; 1.71]	(0.23)	[0.63; 1.50]
Endogenous dependencies					
Mutuality	1.22*	1.00*		0.81*	0.39
	(0.21)			(0.25)	[-0.12; 0.96]
Outdegree popularity				0.95*	
				(0.09)	
Twopaths				-0.04*	
				(0.02)	
GWIdegree (2.0)				3.42*	
				(1.47)	
GWESP (1.0)				0.58*	
				(0.16)	
GWOdegree (0.5)				8.42*	
				(2.11)	

Table 3: * $p < 0.05$. Logistic regression and ERGM results are shown with standard errors in parentheses. MRQAP provides no standard errors. LSM and AME are shown with 95% posterior credible intervals provided in brackets.

There are also notable differences between the parameter estimates that result from the MRQAP, ERGM, and the AME. Using the AME we find evidence that `Preference dissimilarity` is associated with a reduced probability of collaboration between a pair of actors, which is in line with the theoretical expectations stated earlier. Additionally, the AME and MRQAP results differ from ERGM for the nodal effects related to whether a receiver of a collaboration is a government actor, `Alter=Government actor`, and whether the sender is an environmental NGO, `Ego=Environmental NGO`.

When it comes to estimating higher-order effects, ERGM is able to provide explicit estimates of a variety of higher-order parameters, however, this comes with the caveat that these are the "right" set of endogenous dependencies. The AME approach, as shown in Equation 4, estimates network dependencies by examining patterns left over after taking into account the observed covariates. For the sake of space, we focus on examining the third-order dependencies left over after accounting for the observed covariates and network covariance structure modeled by the SRM. A visualization of remaining third-order dependencies is shown in Figure 3.

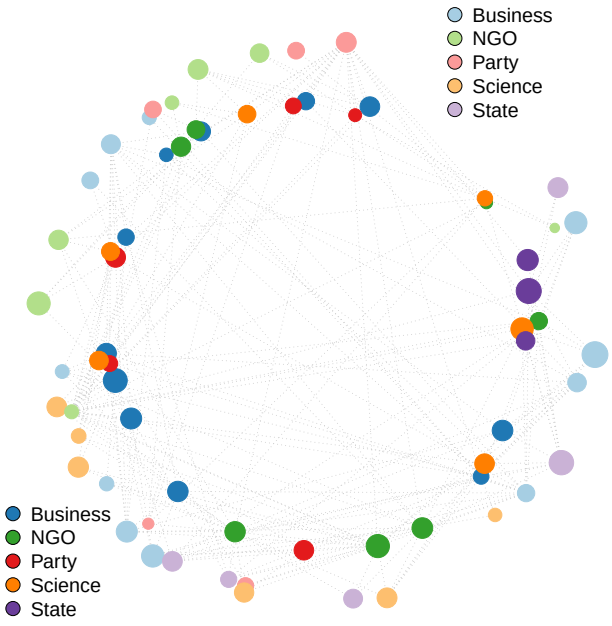

Figure 3: Circle plot of estimated latent factors.

In Figure 3, the directions of $\hat{u}_i$'s and $\hat{v}_i$'s are noted in lighter and darker shades, respectively, of an actor's type.[18] The size of actors is a function of the magnitude of the vectors, and dashed lines between actors indicate greater than expected levels of collaboration based on the regression term and additive effects. In the case of the application dataset that we are using here organization names have been anonymized

[18] For example, actors from industry and business are assigned a color of blue and the direction of $\hat{u}_i$ for these actors is shown in light blue and $\hat{v}_i$ in dark blue

and no additional covariate information is available. However, if we were to observe nodes sharing certain attributes clustering together in this circle plot that would mean such an attribute could be an important factor in helping us to understand collaborations among actors in this network. Given how actors of different types are distributed in almost a random fashion in this plot, we can at least be sure that it is unlikely other third-order patterns can be picked up by that factor.

2.2. Tie Formation Prediction

How do these approaches fit the data out-of-sample? We utilize a cross-validation procedure to assess the out-of-sample performance for each of the models presented in Table 3 as follows:

- Randomly divide the $n \times (n-1)$ data points into S sets of roughly equal size, letting s_{ij} be the set to which pair $\{ij\}$ is assigned.
- For each $s \in \{1, \ldots, S\}$:
  - Obtain estimates of the model parameters conditional on $\{y_{ij} : s_{ij} \neq s\}$, the data on pairs not in set s.
  - For pairs $\{kl\}$ in set s, let $\hat{y}_{kl} = E[y_{kl}|\{y_{ij} : s_{ij} \neq s\}]$, the predicted value of y_{kl} obtained using data not in set s.

The procedure summarized in the steps above generates a sociomatrix of out-of-sample predictions of the observed data. Each entry $\hat{y}_{ij}$ is a predicted value obtained from using a subset of the data that does not include y_{ij}. In this application we set S to 45 which corresponds to randomly excluding approximately 2% of the data from the estimation. Such a low number of observations were excluded in every fold because excluding any more observations would cause the ERGM specification to result in a degenerate model that empirically can not be fit. This highlights the computational difficulties associated with ERGMs in the presence of even small levels of missingness. Not only do you have to have the exactly correct model specification, you also have to have virtually all of the data on the true network.

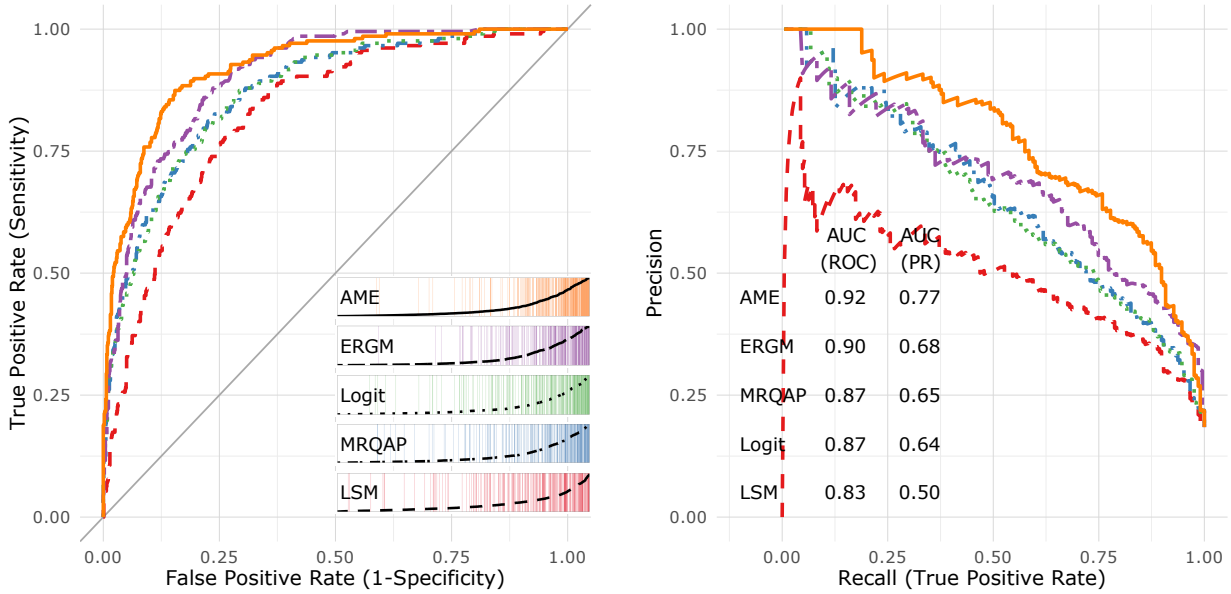

Figure 4: Assessments of out-of-sample predictive performance using ROC curves, separation plots, and PR curves. AUC statistics are provided as well for both curves.

Using the set of out-of-sample predictions we generate from the cross-validation procedure, we provide a series of tests to assess model fit. First, is a diagnostic that is common in the political science literature. The left-most plot in Figure 4 compares the five approaches in terms of their ability to predict the out-of-sample occurrence of collaboration based on Receiver Operating Characteristic (ROC) curves. ROC curves provide a comparison of the trade-off between the True Positive Rate (TPR), sensitivity, False Positive Rate (FPR), 1-specificity, for each model. On this diagnostic, the AME model performs best closely followed by ERGM. The MRQAP and Logit approaches perform similarly, and the LSM approach lags notably behind the other specifications.[19]

A more intuitive visualization of the differences between these modeling approaches can be gleaned through examining the separation plots included on the right-bottom edge of the ROC plot. This visualization tool plots each of the observations, in this case actor pairs, in the dataset according to their predicted value from left (low values) to right (high values). Models with a good fit should have all network links, here these are colored by the modeling approach, towards the right of the plot. Using this type of

[19] Figure **??** in the Appendix provides additional comparisons between our AME approach and various parameterizations of the LSM, in each case we find that the AME approach provides far superior results in terms of out-of-sample predictive performance. However, the LSM approach does begin to perform notably better when incorporating sender and receiver random effects. We also compare performance when using varying values of K for the AME model, we find that increasing K to 3 or 4 does not improve out-of-sample model fit. Results are shown in Figure **??**. Typically setting $K = 2$ works well for most applied cases.

visualization we can again see that the AME and ERGM models performs better than the alternatives.

The last diagnostic we highlight to assess predictive performance are precision-recall (PR) curves. In both ROC and PR space we utilize the TPR, also referred to as recall–though in the former it is plotted on the y-axis and the latter the x-axis. The difference, however, is that in ROC space we utilize the FPR, while in PR space we use precision. FPR measures the fraction of negative examples that are misclassified as positive, while precision measures the fraction of examples classified as positive that are truly positive. PR curves are useful in situations where correctly predicting events is more interesting than simply predicting non-events (Davis and Goadrich, 2006). This is especially relevant in the context of studying many relational datasets in political science such as conflict, because events in such data are extremely sparse and it is relatively easy to correctly predict non-events. In the case of our application dataset, the vast majority of dyads, 80%, do not have a network linkage, which points to the relevance of assessing performance using the PR curves as we do in the right-most plot of Figure 4. We can see that the relative-ordering of the models remains similar but the differences in how well they perform become much more stark. Here we find that the AME approach performs notably better in actually predicting network linkages than each of the alternatives. Area under the curve (AUC) statistics are provided in Figure 4 and these also highlight AME's superior out-of-sample performance.

2.3. Capturing Network Attributes

For network data it is also important to assess whether a model adequately captures the network parameters of the dependent variable (Hunter et al., 2008). To do this one can compare the observed network with a set of networks simulated from the estimated models. We restrict our focus to the three approaches–LSM, ERGM, and AME–that explicitly seek to model network interdependencies. We simulate 1,000 networks from the three models and compare how well they align with the observed network in terms of four network statistics: (1) the empirical standard deviation of the row means (i.e., heterogeneity of nodes in terms of the ties they send); (2) the empirical standard deviation of the column means (i.e., heterogeneity of nodes in terms of the ties they receive); (3) the empirical within-dyad correlation (i.e., measure of reciprocity in the network); and (4) a normalized measure of triadic dependence. A comparison of the LSM, ERGM, and AME models among these four statistics is shown in Figure 5.[20]

[20]Scholars in the networks field usually test for more specific dependencies in order to ascertain whether a particular endogenous covariate needs to be added or modified. We perform this same performance exericise on a more specific set of statistics and visualize the results in Figure **??** of the Appendix. There we also find that the AME does as well as ERGM, and that the LSM model lags behind.

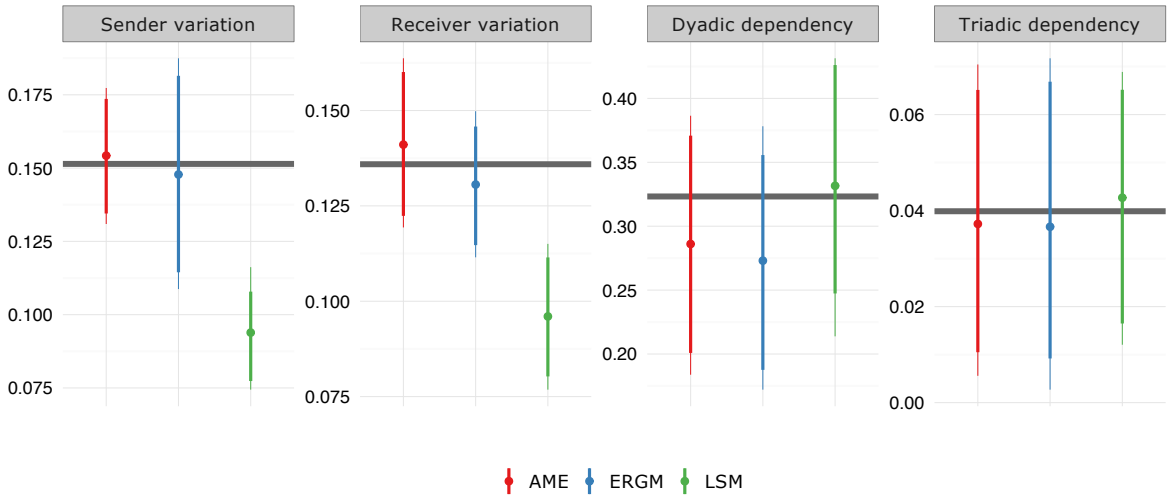

Figure 5: Network goodness of fit summary using **amen**.

Here it becomes quickly apparent that the LSM model fails to capture how active and popular actors are in the Swiss climate change mitigation network.[21] The AME and ERGM specifications again both tend to do equally well.[22] If when running this diagnostic, we found that the AME model did not adequately represent the observed network this would indicate that we might want to increase K to better account for network interdependencies. No changes to the model specification as described by the exogenous covariates a researcher has chosen would be necessary. If the ERGM results did not align with the diagnostic presented in Figure 5, then this would indicate that an incorrect set of endogenous dependencies have been specified. Failing to identify (or find) the right specification will leave the researcher with the problems we introduced earlier.

3. Conclusion

The AME approach to estimation and inference in network data provides a number of benefits over alternative approaches. Specifically, it provides a modeling framework for dyadic data that is based on familiar statistical tools such as linear regression, GLM, random effects, and factor models. We have an understanding of how each of these tools work, they are numerically more stable than ERGM approaches, and more general than alternative latent variable models such as the latent distance or class frameworks. Further the estimation procedure utilized in AME avoids confounding the effects of

[21] Interestingly, even after incorporating random sender and receiver effects into the LSM framework this problem is not completely resolved, see Figure **??** in the Appendix for details.

[22] Not surprisingly, if we increase K in the AME approach we are able to better account for triadic dependencies, see Figure **??** in the Appendix for details.

nodal and dyadic covariates with actor positions in the latent space as the latent distance variable does. For researchers in international relations and more broadly across political science this is of primary interest, as many studies that employ relational data still have conceptualizations that are monadic or dyadic in nature. Additionally, through the application dataset utilized herein we show that the AME approach outperforms both ERGM and latent distance models in out-of-sample prediction, and also is better able to capture network dependencies than the latent distance model.

More broadly, relational data structures are composed of actors that are part of a system. It is unlikely that this system can be viewed simply as a collection of isolated actors or pairs of actors. The assumption that dependencies between observations occur can at the very least be examined. Failure to take into account interdependencies leads to biased parameter estimates and poor fitting models. By using standard diagnostics such as shown in Figures **??** and 5, one can easily assess whether an assumption of independence is reasonable. We stress this point because a common misunderstanding that seems to have emerged within the political science literature relying on dyadic data is that a network based approach is only necessary if one has theoretical explanations that extend beyond the dyadic. This is not at all the case and findings that continue to employ a dyadic design may misrepresent the effects of the very variables that they are interested in. The AME approach that we have detailed here provides a statistically familiar way for scholars to account for unobserved network structures in relational data. Additionally, through this approach we can visualize these dependencies in order to better learn about the network patterns that remain in the event of interest after having accounted for observed covariates.

When compared to other network based approaches such as ERGM, AME is easier to specify and utilize. It is also more straightforward to interpret since it does not require interpretation of unusual features such as *three-stars* which fall outside of the normal language for discussing social science. Further, the **amen** package facilitates the modeling of longitudinal network data. In sum, excuses for continuing to treat relational data as conditionally independent are no longer valid.

Aronow, P. M., Samii, C., Assenova, V. A., Autumn 2015. Cluster–robust variance estimation for dyadic data. Political Analysis 23 (4), 564–577.

Barabási, A.-L., Réka, A., October 15 1999. Emergence of scalin in random networks. Science 286, 509–510.

Beck, N., Katz, J. N., Tucker, R., 1998. Taking time seriously: Time-series-cross-section analysis with a binary dependent variable. American Journal of Political Science 42 (2), 1260–1288.

Breunig, C., Cao, X., Luedtke, A., October 2012. Global migration and political regime type: A democratic disadvantage. British Journal of Political Science 42 (4), 825–854.

Cao, X., 2012. Global networks and domestic policy convergence: A network explanation of policy changes. World Politics 64 (03), 375–425.

Chatterjee, S., Diaconis, P., 2013. Estimating and understanding exponential random graph models. The Annals of Statistics 41 (5), 2428–2461.

Cranmer, S. J., Leifeld, P., McClurg, S. D., Rolfe, M., tba 2016. Navigating the range of statistical tools for inferential network analysis. American Journal of Political Science In Press (tba), tba.

Davis, J., Goadrich, M., 2006. The relationship between precision-recall and roc curves. In: Proceedings of the 23rd international conference on Machine learning. ACM, pp. 233–240.

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

Diehl, P. F., Wright, T. M., 2016 in press. A conditional defense of the dyadic approach. International Studies Quarterly.

Erikson, R. S., Pinto, P. M., Rader, K. T., 2014. Dyadic analysis in international relations: A cautionary tale. Political Analysis 22 (4), 457–463.

Goodreau, S. M., Handcock, M. S., Hunter, David R .and Butts, C. T., Morris, M., 2008. A statnet tutorial. Journal of Statistical Software 24 (9), 1.

Greenhill, B. D., 2015. Transmitting rights; International Organizations and the Diffusion of Human RIghts Practices. Oxford University Press, Oxford, UK.

Hammersley, J. M., Clifford, P., 1971. Markov fields on finite graphs and lattices.

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., Morris, M., 2008. `statnet`: Software tools for the representation, visualization, analysis and simulation of network data. Journal of Statistical Software 24 (1), 1548.

Hoff, P., 2009. Multiplicative latent factor models for description and prediction of social networks. Computational and Mathematical Organization Theory 15 (4), 261–272.

Hoff, P., Fosdick, B., Volfovsky, A., He, Y., 2015. amen: Additive and Multiplicative Effects Models for Networks and Relational Data. R package version 1.1.
URL `https://CRAN.R-project.org/package=amen`

Hoff, P. D., 2005. Bilinear mixed-effects models for dyadic data. Journal of the American Statistical Association 100 (4690), 286–295.

Hoff, P. D., 2008. Modeling homophily and stochastic equivalence in symmetric relational data. In: Platt, J. C., Koller, D., Singer, Y., Roweis, S. T. (Eds.), Advances in Neural Information Processing Systems 20. Processing Systems 21. MIT Press, Cambridge, MA, USA, pp. 657–664.

Hoff, P. D., 2015. Dyadic data analysis with `amen`. arxiv arxiv:1506.08237, 1–48.

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.

Hoff, P. D., Ward, M. D., 2004. Modeling Dependencies in International Relations Networks. Political Analysis 12 (2), 160–175.

Hunter, D., Handcock, M., Butts, C., Goodreau, S. M., Morris, M., 2008. `ergm`: A package to fit, simulate and diagnose exponential-family models for networks. Journal of Statistical Software 24 (3), 1–29.

Hunter, D., Krivitsky, P., Schweinberger, M., 2012. Computational statistical methods for social network models. Journal of Computational and Graphical Statistics 21 (4), 856–882.

Ingold, K., 2008. Les mécanismes de décision: Le cas de la politique climatique Suisse. Politikanalysen. Rüegger Verlag, Zürich.

Ingold, K., Fischer, M., 2014. Drivers of collaboration to mitigate climate change: An illustration of swiss climate policy over 15 years. Global Environmental Change 24, 88–98.

Kinne, B. J., 2013. Network dynamics and the evolution of international cooperation. American Political Science Review 107 (04), 766–785.

Kirkland, J. H., 2012. Multimember districts' effect on collaboration between US state legislators. Legislative Studies Quarterly 37 (3), 329–353.

Krivitsky, P. N., Handcock, M. S., 2015. latentnet: Latent Position and Cluster Models for Statistical Networks. The Statnet Project (http://www.statnet.org), r package version 2.7.1.
URL http://CRAN.R-project.org/package=latentnet

Lai, D., 1995. A structural approach to alignment: A case study of the china- soviet-u.s. strategic triangle, 1971-1988. International Interactions 20 (4), 349–374.

Li, H., Loken, E., 2002. A unified theory of statistical analysis and inference for variance component models for dyadic data. Statistica Sinica 12 (2), 519–535.

Manger, M. S., Pickup, M. A., Snijders, T. A., 2012. A hierarchy of preferences: A longitudinal network analysis approach to PTA formation. Journal of Conflict Resolution 56 (5), 852–877.

Metternich, N., Minhas, S., Ward, M., 2015. Firewall? or wall on fire? A unified framework of conflict contagion and the role of ethnic exclusion. Journal of Conflict Resolution Online version (doi:10.1177/0022002715603452).

Richardson, L. F., 1960. Arms and Insecurity. Quadrangle.

Salter-Townshend, M., White, A., Gollini, I., Murphy, T. B., 2012. Review of statistical network analysis: Models, algorithms, and software. Statistical Analysis and Data Mining 5 (4), 243–264.

Schweinberger, M., 2011. Instability, sensitivity, and degeneracy of discrete exponential families. Journal of the American Statistical Association 106 (496), 1361–1370.

Signorino, C., June 1999. Strategic interaction and the statistical analysis of international conflict. American Political Science Review 92 (2), 279–298.

Snijders, T. A., 2011. Statistical models for social networks. Annual Review of Sociology 37, 131–53.

Snijders, T. A., Pattison, P. E., Robins, G. L., Handcock, M. S., 2006. New specifications for exponential random graph models. Sociological Methodology 36 (1), 99–153.

Strauss, D., Ikeda, M., 1990. Pseudolikelihood estimation for social networks. Journal of the American Statistical Association 85, 204–212.

Van Duijn, M. A., Gile, K. J., Handcock, M. S., 2009. A framework for the comparison of maximum pseudo-likelihood and maximum likelihood estimation of exponential family random graph models. Social Networks 31 (1), 52–62.

Ward, M. D., Siverson, R. M., Cao, X., 2007. Disputes, democracies, and dependencies: A reexamination of the kantian peace. American Journal of Political Science 51 (3), 583–601.

Warner, R., Kenny, D., Stoto, M., 1979. A new round robin analysis of variance for social interaction data. Journal of Personality and Social Psychology 37, 1742–1757.

Wong, G. Y., 1982. Round robin analysis of variance via maximum likelihood. Journal of the American Statistical Association 77 (380), 714–724.