The Additive and Multiplicative Effects Network Model*

Seminar on Statistical Modeling of Social Networks

Daniel A. Seussler Becerra

June 2021

Abstract

This seminar paper introduces the Additive and Multiplicative Effects (AME) network model as summarized in P. D. Hoff (2021). First, a general introduction to network and dyadic data is given along with the statistical challenges of related data analysis. Then, the AME network model is introduced and motivated as a Bayesian approach to network analysis. Second, an application case to the interstate alliance network in 2000 is presented along with a comparison of restricted modeling approaches. The paper concludes with a discussion of the limitations of chosen approach and outlook. *Keywords: Dyadic Data, Bayesian Estimation, Latent Factors, International Relations.*

1 Introduction

Networks are ubiquitous in political science, economics, and sociology, but a variety of applications can also be found in natural sciences. In its most basic form, we consider a set of actors with relationships between each other. A relationship between two nodes is called a dyad[1] or a dyadic observation, hence the name dyadic data analysis. This requires specific statistical approaches, the center of research in statistical network analysis. Three aspects of networks warrant further research. First, clustering or community detection could be of interest in explorative data analysis. Second, network composition, uncovering determinants of relationship formation could yield information about the structure of interest. Third, the dynamic evolution of networks can be informative.

One popular approach is presented here, the Additive and Multiplicative Effects (AME) Network Model. This model has been introduced in P. D. Hoff (2021) and is based on previous work in P. D. Hoff (2007) and P. D. Hoff (2005). It lies in the intersection of mentioned research approaches; in that, it can describe patterns in networks while accounting for the effects of external variables and influences of network effects. To accomplish this, an extended error component is embedded in a generalized linear regression framework, designed to capture the additional statistical dependencies of observations inherent in network or dyadic data.

In panel data analysis, we are concerned that the repeated measurement of a certain unit could introduce bias into our estimates, which requires adjustments to the models employed. Similarly, the composition of network or dyadic data, specifically the possibility, that an actors' relationship with another actor, can depend on the latter one's relationships in the network. Therefore, we classify three types of effects, also called statistical dependencies, in networks, based on the number of nodes involved. Consider the relationships of actor i to actors j, k, and l, stated by the set $\{y_{ij}, y_{ik}, y_{il}\}$. Then this set could exhibit higher similarity *within* itself than with other dyadic observations, due to the common actor i. I. e., we could see heterogeneity in the propensity to send (or receive) network ties, depending on if the actor is the sender of the receiver of a given network tie. These are first-order network effects that are observed on a nodal or actor level. Involving two nodes, i.e., dyadic relationships, we could see patterns of reciprocity. This is characterized by the statistical dependency between observation y_{ij} and y_{ji}. Third, there is reason to assume dependence involving triadic relationships within a given network. These could be transitivity, balance, and clustering effects, and are best described by proverbs such as *the enemy of my enemy is my friend* and *my friends' friend is also my friend.*

The rest of this paper is structured in two sections. In the following theory section, I introduce common network statistics for the above-illustrated network effects and the AME network model as specified in P. D. Hoff (2021). Then, in an application section, I use data of the interstate alliance network in 2000 and apply the beforehand introduced AME network model. A comparison of

---

[1]Node and dyad are commonly named actor and tie in the Social Sciences, I use the names interchangeably. In Mathematical Graph Theory, we would refer to vertices V with edges E linking two vertices at a time. The set $G = (V, E)$ is called a graph G.

alternative restricted specifications illustrates the necessity of network approaches in the analysis of network data. The seminar paper concludes with a discussion on the limitations and further approaches.

2   Theory

Let $Y \in \mathbb{R}^{n \times n}$, for network actors labeled $1, \ldots, n$ where y_{ij} describes the directed relationship between actor i and j. The former is taken to be the sender, the latter the receiver of the network tie. Because we assume the actors to be acting (or embedded) within a network environment, there is *a priori* no reason to assume (conditional) independence of the observations y_{ij}.[2] For a given matrix Y we define the following network statistics. Note that, we rule out self-ties within a network, i.e., for any given matrix representing a network we assume the diagonal to be undefined.[A network matrix with this property is commonly referred to as a sociomatrix.] Undefined observations should therefore be dropped in the following calculations.

First-order effects characterize the actor specific heterogeneity in sending (or receiving) ties within a network. In a homogenous sender setting, we would expect the row means $\bar{Y}_{i \cdot}$ to be similar (or column means). Heterogeneity in sending or receiving ties is therefore summarised by the standard deviation of the respective row and column means:

$$sd.rowmean(Y) = \frac{1}{n} \sum_{i=1}^{n} \left(\bar{Y}_{i \cdot} - \overline{Y_{i \cdot}} \right)^2 \text{, and } sd.colmean(Y) = \frac{1}{n} \sum_{i=1}^{n} \left(\bar{Y}_{\cdot j} - \overline{Y_{\cdot j}} \right)^2 .$$

Second-order effects, i.e., reciprocity, describe the statistical dependency of the directed relationship between two actors in the network. For the relationship between actor i and j, the correlation with the relationship of actor j to i, if $i \neq j$, captures this effect. For a network it is given as

$$dyadic.dependency(Y) = corr \left(vec(Y), vec \left(Y^\top \right) \right) .$$

Here, $vec(A)$ is a $mn \times 1$ column vector, which is obtained by stacking the columns of $A \in \mathbb{R}^{n \times m}$ on top of another. Naturally, for an undirected network dyadic dependency is 1.

Third-order effects are patterns among three nodes and corresponding statistics shall capture triadic relationships as described above. Let $E = Y - \bar{Y}$. Then

$$cycle.dependency(Y) = \frac{1}{n(n-1)(n-2)} \sum_{i \neq j \neq k \neq i} E_{ij} E_{jk} E_{ki},$$

$$trans.dependency(Y) = \frac{1}{n(n-1)(n-2)} \sum_{i \neq j \neq k \neq i} E_{ij} E_{jk} E_{ik}.$$

---

[2]The observation y_{ij} refers to the ij-th observation in the matrix Y. Capital letters will be used to emphasise the matrix structure and boldface for vector variables.

Optionally the statistics can be standardized with $sd(vec(Y))$.[3] Naturally, for an undirected network cycle dependency equals transitive dependency. In the process of model selection, the network statistics above can be used to assess model fit. Based on draws from the posterior distribution (we later follow a Bayesian approach to model estimation), sociomatrices are simulated and respective goodness of fit statistics computed. This yields a distribution of network dependencies which is then compared to the observed statistic.

Lastly, actors i and j are *stochastically equivalent*, if both exhibit a similar probability to send (or receive) network ties with other groups of actors. This contrasts with *homophily* between actors, which describes a higher likelihood of a network tie based on similar characteristics of them.

The Additive and Multiplicative Effects network model is based on two types of random effects, additive and multiplicative, which are added to an error component in a generalized linear regression framework. The inclusion guarantees that the error structure considers the types of network dependencies inherent in dyadic observations as specified above. I introduce both separately and start with the additive component.

2.1 Additive Effects: The Social Relations Regression Model

To accommodate actor-specific heterogeneity beyond fixed covariates, sender and receiver additive random effects (a_i, b_j) are included in an extended error component. Let μ be a mean or a linear predictor $\boldsymbol{\beta}^\top \mathbf{x}_{ij}$.[3] This yields the following model, which is also referred to as the Social Relations Regression (SRRM) model:

$$y_{ij} = \mu + e_{ij}$$
$$e_{ij} = a_i + b_j + \epsilon_{ij}$$
$$\{(a_1, b_1), \dots, (a_n, b_n)\} \sim \mathcal{N}_2(\mathbf{0}, \Sigma_{ab}),$$
$$\{(\epsilon_{ij}, \epsilon_{ji}) : \ i \neq j\} \sim \mathcal{N}_2(\mathbf{0}, \Sigma_\epsilon),$$
$$\Sigma_{ab} = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix}, \quad \Sigma_\epsilon = \sigma_\epsilon^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

for $\{1 \leq i, j \leq n, i \neq j\}$. Parameter interpretation is straightforward. Nodal random effects a_i, b_j are independent normal gaussian variables with zero mean and covariance matrix Σ_{ab} for $i, j = 1, \dots, n$. The covariance matrix Σ_ϵ of the error term ϵ is composed of the errorvariance σ_ϵ^2 and a reciprocity parameter ρ. Moreover

$$Cov[Y_{ik}, Y_{il}] = \sigma_a^2,$$
$$Cov[Y_{ik}, Y_{lk}] = \sigma_b^2,$$
$$Cov[Y_{ik}, Y_{kl}] = \sigma_{ab},$$
$$Cov[Y_{ij}, Y_{ji}] = 2\sigma_{ab} + \rho\sigma^2,$$

---

[3]The p-dimensional covariate vector x_{ij} may include nodal or dyadic specific values.

are within-row variance, within-column variance, and row-column variance without and with reciprocity.

2.2 Multiplicative Effects: The Latent Factor Model

The nodal random effects can capture first- and second-order network effects, but cannot account for triadic relationships within the data. This is illustrated in the application case below, but can also be seen as follows. For an estimated model the extended error component $\hat{e}_{ij} = \hat{a}_i + \hat{b}_j + \hat{\epsilon}_{ij}$ is 0 in expectation, and because it is approximate normal gaussian it's third moments equal 0. Therefore, an SRRM model will *not* capture third-order network dependencies by design. This issue is alleviated by the inclusion of latent factors as multiplicative effect for network nodes. The idea stems from the Latent Space Model, where network actors are associated in a lower-dimensional space through latent variables. In a Latent Space approach with dimension r, we define an additional component:

$$\alpha(\mathbf{u}_i, \mathbf{u}_j) = -|\mathbf{u}_i - \mathbf{u}_j|, \qquad \mathbf{u}_i \in \mathbb{R}^r, \qquad i = 1, \dots, n.$$

Because the L_1 norm obeys the triangle inequality, [. . .]

The Latent Space Approach comes with one caveat. If $\mathbf{u}_l, \mathbf{u}_k$ are close in the latent space, then $|\mathbf{u}_l - \mathbf{u}_k|$ is small. But then every other actor with latent variable $\mathbf{u}_j$ such that $|\mathbf{u}_l - \mathbf{u}_j| \approx |\mathbf{u}_k, \mathbf{u}_j|$ which implies all three nodes exhibit similar stochastic patterns. Thus, this approach cannot distinguish between homophily and stochastic equivalence.

The multiplicative component is therefore specified as a latent factor, first introduced in P. D. Hoff (2005). Let $r \in \mathbb{N}$ be the dimension of the latent factor, each for the sender and receiver, i.e., $\mathbf{u}_i, \mathbf{v}_j \in \mathbb{R}^r$ for $i, j = 1, \dots, n$. It can be interpreted as the unobserved characteristic of the node as sender or receiver. The term

$$\alpha(\mathbf{u}_i, \mathbf{v}_j) = \mathbf{u}_i^T \mathbf{v}_j$$

is the multiplicative component. In a symmetric sociomatrix, we introduce L as an $r \times r$ diagonal matrix, i.e.,

$$\alpha(\mathbf{u}_i, \mathbf{v}_j) = \mathbf{u}_i^T L \mathbf{v}_j$$

to properly generalize symmetric matrices. Higher values of α indicate higher likelihood of a tie and depend on the diagonal values of L aswell as u_{ik} and v_{kj} for $k \in \{1, \dots, r\}$. Determining the dimension of the latent space is not a trivial task and is to be considered alongside other modeling decisions. However, selecting lower rank dimensions facilitates visualization of the estimated effects, which can be informative by itself. To minimise overfitting and permit interpretation as random effects, let

$$\{(\mathbf{u}_1, \mathbf{v}_1), \dots, (\mathbf{u}_n, \mathbf{v}_n)\} \sim \mathcal{N}_{2r}(\mathbf{0}, \Psi).$$

This multiplicative effect is introduced in P. D. Hoff (2007) as the latent eigenmodel, which also proves the generalizability of this approach regarding the latent space component.

2.3 The AME Model: Additive plus Multiplicative Effects

Merging both types of random effects yields the Additive and Multiplicative network effects model, which can account for the network effects described above. The model is specified as

$$y_{ij} = \beta^T \mathbf{x}_{ij} + e_{ij}$$
$$e_{ij} = a_i + b_j + \alpha(\mathbf{u}_i, \mathbf{v}_j) + \epsilon_{ij}$$

for $\{1 \le i, j \le n, i \ne j\}$ and a_i, b_j, $\alpha(\mathbf{u}_i, \mathbf{v}_j)$ with the respective covariance structures as defined above. Using matrices, we can rewrite the model for an appropiate M as

$$Y = M(X, \boldsymbol{\beta}) + \mathbf{a}\mathbf{1}^\top + \mathbf{1}\mathbf{b}^\top + UV^\top$$

and

$$Y = M(X, \boldsymbol{\beta}) + \mathbf{a}\mathbf{1}^\top + \mathbf{1}\mathbf{a}^\top + ULU^\top$$

in the symmetric case. Again, we take the diagonal to be not defined in matrix notation. In fact, as the algorithm presented below does take this into account, it is viable to handle missing values off the diagonal. This is a welcome feature for observational studies, where data may be missing at random. Of course, until now the directed tie y_{ij} is taken to be a continuous outcome variable in $\mathbb{R}$. This is not necessary and the model can be extended to include other outcome variables as well, see the part on further extensions below. This does require the introduction of a link function, analogous to a generalized linear regression approach, as well as modifications to the estimation strategy. Generally, the AME network model is estimated using a Bayesian approach with Markov chain Monte Carlo (MCMC) simulation, which is described in the following part. I restrict this again to continuous outcome variables.

An empirical comparison to other popular approaches, such as the Exponential Random Graph Model, can be found in Minhas, Hoff, and Ward (2019).

2.4 Parameter Estimation

The MCMC estimation approach is achieved through the implementation of a Gibbs sampling algorithm, summarized for the AME *without* multiplicative effects in Algorithm 1. This is in turn extended to include the multiplicative effects in Algorithm 2. All details and derivations for the full conditional distributions necessary to implement this sampler are outlined in P. D. Hoff (2021). Here, I provide only the (semi-) conjugate prior- and full conditionals for each step. First, let

$$\beta \sim \mathcal{N}_p(\boldsymbol{\beta_0}, Q_0^{-1}),$$
$$\frac{1}{\sigma^2} \sim \text{Gamma}\left(\frac{\nu_0}{2}, \frac{\nu_0 \sigma^2}{2}\right),$$
$$\Sigma^{-1} \sim \text{Wishart}\left(\frac{\Sigma_0^{-1}}{\eta_0}, \eta_0\right)$$

be our prior distributions for appropiate $\beta_0, Q_0, \nu_0, \Sigma_0, \eta_0$. Note that we consider Σ_ϵ through σ^2 and ρ, as well as $\Sigma := \Sigma_{ab}$.

Algorithm 1: Gibbs Sampling for the SRRM, P. D. Hoff (2021).

Initalize unkown variables;

1. Simulate $\{\boldsymbol{\beta}, \mathbf{a}, \mathbf{b}\}$ given $Y, \Sigma, \sigma^2, \rho$;

2. Simulate σ^2 given $Y, \boldsymbol{\beta}, \mathbf{a}, \mathbf{b}, \rho$;

3. Simulate ρ given $Y, \boldsymbol{\beta}, \mathbf{a}, \mathbf{b}, \sigma^2$;

4. Simulate Σ given $\mathbf{a}, \mathbf{b}$;

5. Simulate missing values of Y given $\boldsymbol{\beta}, \mathbf{a}, \mathbf{b}, \rho, \sigma^2$ and observed values of Y.

To approximate a posterior distribution by the way of Gibbs Sampling, knowledge of every full conditional, i.e., the probability distribution of one parameter conditioned on *every* other parameter, has to be available. Thus, below I provide every full conditional required for this estimation strategy. Starting from *Step 2*, let

$$SS_1 = \sum_{i<j}(e_{ij}e_{ji})\begin{pmatrix}1 & \rho \\ \rho & 1\end{pmatrix}\begin{pmatrix}e_{ij} \\ e_{ji}\end{pmatrix}, \qquad SS_2 = \sum_{i=1}^{n}\frac{1}{1+\rho}e_{ii}^2,$$

and e_{ij} the ij-th value of $E = Y - (M(X, \boldsymbol{\beta}) + \mathbf{a}\mathbf{1}^\top + \mathbf{1}\mathbf{b}^\top)$. Then the full conditional for the error variance is:

$$\frac{1}{\sigma^2} \mid Y, \boldsymbol{\beta}, \mathbf{a}, \mathbf{b}, \rho \sim \text{Gamma}\left(\frac{1}{2}\left(\nu_0 + n^2\right), \frac{1}{2}\left(\nu_0\sigma_0^2 + SS_1 + SS_2\right)\right)$$

For *Step 3*, note that $\rho \in [-1, 1]$, thus an empirical CDF sampler on $[-1, 1]$ is sufficient to approximate the posterior for ρ. This circumvents the issue of specifying a prior distribution. The covariance matrix of the additive random effects in *Step 4* is simulated using

$$\Sigma^{-1} \mid [\mathbf{a}, \mathbf{b}] \sim \text{Wishart}\left((\eta_0\Sigma_0 + C^\top C)^{-1}, \eta_0 + n\right)$$

where $C = [\mathbf{a}, \mathbf{b}] \in \mathbb{R}^{n \times 2}$ is the columnwise concatenation of $\mathbf{a}, \mathbf{b}$. *Step 5* requires a definition by cases, as there can be three different constellations of missing data. First, values on the diagonal, a missing pair (y_{ij}, y_{ji}), and half a missing pair. Hence,

$$y_{ii} \mid \boldsymbol{\beta}, \mathbf{a}, \mathbf{b}, \rho, \sigma^2 \sim \mathcal{N}\left(m_{ij} + a_i + b_j, \ \sigma^2(1 + \rho)\right),$$

$$(y_{ij}\ y_{ji})^\top \mid \boldsymbol{\beta}, \mathbf{a}, \mathbf{b}, \rho, \sigma^2 \sim \mathcal{N}_2\left(\begin{pmatrix}m_{ij} + a_i + b_j \\ m_{ji} + a_j + b_i\end{pmatrix}, \ \sigma^2\begin{pmatrix}1 & \rho \\ \rho & 1\end{pmatrix}\right),$$

$$y_{ji} \mid y_{ij}, \boldsymbol{\beta}, \mathbf{a}, \mathbf{b}, \rho, \sigma^2 \sim \mathcal{N}\left(\rho(y_{ij} - m_{ij} - a_i - b_j) + m_{ji} + a_j + b_i, \ \sigma^2\left(1 - \rho^2\right)\right).$$

And lastly, the simulation of $\{\boldsymbol{\beta}, \mathbf{a}, \mathbf{b}\}$, *Step 1*. This step requires an additional transformation of

the data and is done for $\boldsymbol{\beta}$ and $\{\mathbf{a}, \mathbf{b}\}$ seperately. Consider the transformation of Y:

$$\tilde{c} = \frac{1}{2\sigma} \left((1+\rho)^{-1/2} + (1-\rho)^{-1/2} \right)$$

$$\tilde{d} = \frac{1}{2\sigma} \left((1+\rho)^{-1/2} - (1-\rho)^{-1/2} \right)$$

$$\tilde{Y} = \tilde{c}Y + \tilde{d}Y^\top$$

which yields the following model. It does does not exhibit dyadic correlation in the error term Z, which facilitates the derivation of the full conditionals. Noted in matrix type equations:

$$\tilde{Y} = M(\tilde{X}, \boldsymbol{\beta}) + \tilde{\mathbf{a}}\mathbf{1}^\top + \mathbf{1}\tilde{\mathbf{b}}^\top + Z$$

$$\tilde{x}_{ij} = \tilde{c}x_{ij} + \tilde{d}x_{ji}$$

$$Z \sim \mathcal{N}_{n \times n}(\mathbf{0}, I)$$

$$(\tilde{a}_1, \tilde{b}_1), \dots, (\tilde{a}_n, \tilde{b}_n) \sim \mathcal{N}_2(\mathbf{0}, \tilde{\Sigma})$$

$$\tilde{\Sigma} = \Sigma_e^{-1/2} \Sigma \Sigma_e^{-1/2}$$

It is then possible to simulate $\{\beta, \tilde{a}, \tilde{b}\}$ and obtain $(a_i, b_i)^\top = \Sigma_e^{1/2}(\tilde{a}_i, \tilde{b}_i)^\top$ by retransforming into the original space. Dropping the $\sim$ briefly to facilitate notation, let $\mathbf{y}$ be the vectorization of $Y = M(X, \boldsymbol{\beta})\mathbf{a}\mathbf{1}^\top + \mathbf{1}\mathbf{v}^\top + Z$, i.e.,

$$\mathbf{y} = \mathbf{m} + WC + \mathbf{z}$$

with $C = [\mathbf{a}, \mathbf{b}]$ as before. Define:

$$G = \left(\Sigma^{-1} + nI \right)^{-1}, \text{ and } H = \left(\Sigma^{-1} + n\mathbf{1}\mathbf{1}^\top \right)^{-1} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} G$$

The full conditional to simulate $\mathbf{a}, \mathbf{b}$ is then

$$[\mathbf{a}, \mathbf{b}] \mid Y, M, \Sigma \sim \mathcal{N}\left(\left(W^\top W + \Sigma^{-1} \otimes I \right)^{-1} W^\top (\mathbf{y} - \mathbf{m}), \; G \otimes I - H \otimes \mathbf{1}\mathbf{1}^\top \right).$$

Retransformation yiels $[\mathbf{a}, \mathbf{b}]$. To simulate $\boldsymbol{\beta}$ let

$$Q_1 = X^\top X$$

$$Q_2 = n^4 h \bar{x} \bar{x}^\top$$

$$Q_3 = -n^2 \left(g_{11} \bar{X}_r^\top \bar{X}_r + g_{12}(\bar{X}_r^\top \bar{X}_c + \bar{X}_c^\top \bar{X}_r) + g_{22} \bar{X}_c^\top \bar{X}_c \right)$$

$$\mathbf{l_1} = X^\top y$$

$$\mathbf{l_2} = n^4 h \bar{x} \bar{y}, \quad h = \mathbf{1}^\top H \mathbf{1}$$

$$\mathbf{l_3} = -n^2 \left(g_{11} \bar{X}_r^\top \bar{y}_r + g_{12}(\bar{X}_r^\top \bar{y}_c + \bar{X}_c^\top \bar{y}_r) + g_{22} \bar{X}_c^\top \bar{y}_c \right)$$

and $Q = Q_1 + Q_2 + Q_3$, $\mathbf{l} = \mathbf{l_1} + \mathbf{l_2} + \mathbf{l_3}$. Then the full conditional density is given as

$$\boldsymbol{\beta}|y, \Sigma \sim \mathcal{N}_p\left((Q_0 + Q)^{-1}(Q_0\beta_{\mathbf{0}} + \mathbf{l}), \ (Q_0 + Q)^{-1}\right).$$

Adding multiplicative effects to the SRRM model is accomplished as follows.

Algorithm 2: Gibbs Sampling for the AME, P. D. Hoff (2021).

Initalize unkown variables;

1. Update $\{\boldsymbol{\beta}, \mathbf{a}, \mathbf{b}, \sigma^2, \rho, \Sigma\}$ and the missing values of Y using Algorithm 1, but with Y replaced by Y - UV^T;

2. Simulate $\Psi^{-1} \sim$ Wishart $((\Psi_0\kappa_0 + [UV]^\top[UV])^{-1}, \kappa_0 + n)$, where $[UV]$ is the $n \times 2r$ matrix equal to the column-wise concatenation of U and V;

3. For each $k = 1, \ldots, r$, simulate the r-th columns of U and V from their full conditional distributions.

Only *Step 3* requires an additional full conditional, the derivation is illustrated in the case of u_1 and the rest is constructed analogous. Define the residual as

$$R = Y - \left(M(X, \boldsymbol{\beta}) + \sum_{k=2}^{r} u_k v_k^\top + \mathbf{a1}^\top + \mathbf{1b}^\top\right) = u_1 v_1^\top + E.$$

Again, to cirumvent issues with the dyadic correlation, R is decorrelated using the transformation introduced above.

$$\tilde{R} = \tilde{c}R + \tilde{d}R^\top = \tilde{c}u_1 v_1^\top + \tilde{d}v_1 u_1^\top + Z$$

$$\tilde{\mathbf{r}} = \left(\tilde{c}(v_1 \otimes I) + \tilde{d}I \otimes v_1)\right)u_1 + \mathbf{z}$$

The second equation is the vectorization of R. Using the conditional mean and variance of $\mathbf{u}_1 \ \mathbf{v}_1$, this yields the conditional distribution

$$u_1|\tilde{R}, v_1 \sim \mathcal{N}\left(\Sigma_{u|v}^{-1} + (\tilde{c}^2\tilde{d}^2)||v_1||^2 I + 2\tilde{c}\tilde{d}v_1 v_1^\top\right)^{-1},$$

$$(\Sigma_{u|v}^{-1} + (\tilde{c}^2\tilde{d}^2)||v_1||^2 I + 2\tilde{c}\tilde{d}v_1 v_1^\top)^{-1}(\Sigma_{u|v}^{-1}\mu_{u|v} + (\tilde{c}R + \tilde{d}R^\top)v_1).$$

Which concludes the section.

2.5 Extensions

Note that by now we have assumed a normal distribution of the outcome variable. However, the presented framework can be extended in several different ways. For one thing, additional types of outcome variables can be specified, such as binary, count, ordinal, and censored data. Ordinal data includes ranked categorical outcomes, for example, *reject, neutral, agree.* This requires changes to the sampling algorithms described above. For the specific changes to the algorithms and full conditionals, as required in the binary case for the following application case, see P. D. Hoff (2021). Furthermore, estimation of missing data points could be an objective. In the Bayesian estimation approach,

imputation of missing data is easily available through posterior simulation, here in Algorithm 1 Step 5. A third possibility is an extension to include longitudinal or replicated networks, where a (possibly changing) set of actors is observed at different time points. One approach to include replicated data is taking it into account as lagged variables. A different approach is presented in Kim et al. (2018), where the AME network model is embedded in a Gaussian process.

The estimation approach above as well as these extensions have been implemented in the R-package `amen` from P. Hoff, Fosdick, and Volfovsky (2020).

3   Application Case: Interstate Defence Alliances

To illustrate the purpose of statistical network approaches we turn our attention to the field of international relations. More specifically, I apply the beforehand introduced AME network model to the interstate defense alliances network in the year 2000. Alliance formation in an interstate framework provides an interesting application case, both because of the availability of data and the assumed extent of network effects inherent in the strategic decisions of state actors. Neglecting the latter effects by presuming conditional independence of observations, as in the classical regression framework, could seriously bias parameter estimates (Signorino (1999)).

The data is an excerpt from the data used in S. J. Cranmer, Desmarais, and Menninga (2012) and S. Cranmer, Desmarais, and Kirkland (2012) and contains the yearly interstate defense network as a time series from 1981 to 2000.[4] It furthermore includes information on military capabilities, political regime types, geographic borders, and interstate conflict. From here on out, I refer to this data set as the alliances data set.

The case study is structured as follows. First, restricted to a cross-sectional study of the year 2000, I examine the alliances' structure with respect to the network statistics and preliminary models. Second, to motivate the statistical network approach, I estimate four different specifications and compare these for their goodness of fit. I conclude with a discussion of the challenges and improvements of the AME network model. The computational analysis was implemented with the R-Software (R Core Team (2021)) and the AMEN R-Package (P. Hoff, Fosdick, and Volfovsky (2020)).

3.1   Network Statistics

The alliances data set includes 158 countries[5] and a total of 767 observed interstate alliances. In the corresponding sociomatrix, a value of 1 indicates the presence of an alliance between two countries, 0 its absence. Note that we have a symmetric matrix, as an alliance between two states is taken to be not directed. 55 countries have no alliance recorded in the data set, the United States and

---

[4]The data set `alliances` can be conveniently accessed as network data in the R-Package `xergm.common`.

[5]This data set covers the period of 1981 - 2000, it includes all countries which were selected at any time of given series. I exclude the Yemen Arab Republic, Yemen People's Republic, German Democratic Republic, German Federal Republic, Yugoslavia, and Czechoslovakia as former countries, which yields a total of 158 countries.
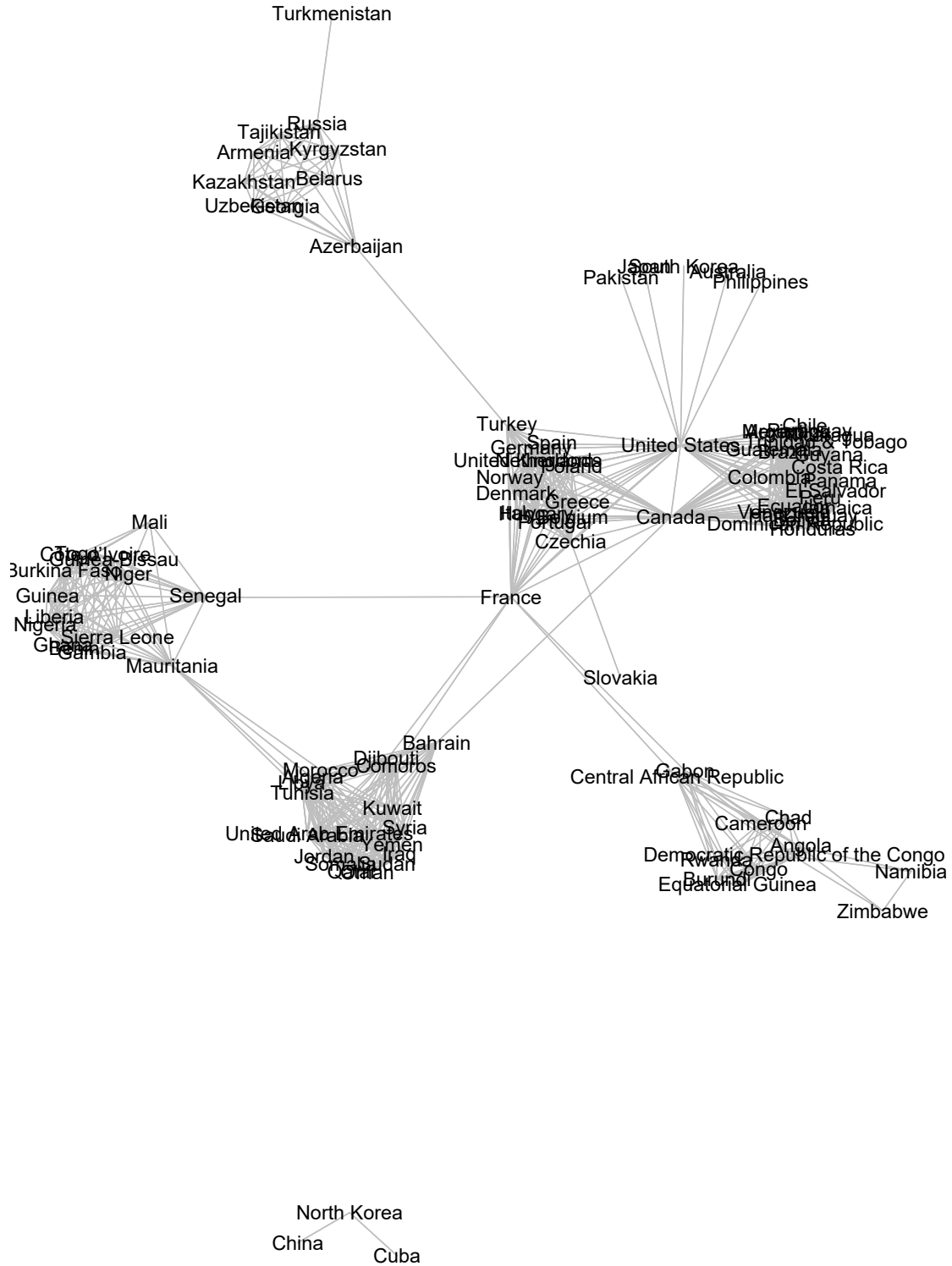
Figure 1: The interstate defense alliance network in the year 2000. Countries with no interstate alliance recorded are ommitted from the plot.

Canada declare the most, with 43 and 39 alliances respectively. The median number of alliances is 9 and the network density is 0.061.

A visualization of the interstate alliance network is provided in Figure 1. Strong clustering on based geographic proximity is immediately apparent. Located in the upper right of the figure we find the Middle East, below France in the center are the European states. On the left, we find the former USSR countries, in the lower part Latin America and the Caribbean states. Perhaps more surprisingly there are two distinct clusters of African states.

Observed statistics of network dependence will be used later on to assess model fit and are computed as described above. First-order statistics, the standard deviation of the row- and column means is 0.0596. As our sociomatrix is symmetric, second-order dependency or reciprocity is 1. Third-order statistics are 0.3881, both for the cycle and transitive dependency, because of symmetry.

3.2 Model Comparison: Interstate Defence Alliances

I follow the specification of Kim et al. (2018) and Warren (2010) to define nodal and dyadic covariates which are both sensitive to the approach and relevant to the theoretic discussion of interstate alliance networks.[6]

1. Nodal Covariates. As covariate of interest-specific to each actor in the network, I include Gross Domestic Product (GDP) in per capita terms and on a log scale for each country using data from The World Bank (2021b). Missing data points are imputed on an individual basis.[7]

2. Dyadic Covariates. As dyadic effects I include covariates on geographic factors, political and military similarity, past conflicts, economic dependence, and cultural similarity. The covariates are defined as follows:

1. $GeoDistance_{ij}$ is the logarithmic geographic distance between two countries measured by the distance of the respective capitals. If the two countries share a border, the log distance is set 0.
2. $CulturalSim_{ij}$ is a dichotomous variable which takes the value 1, if the most spoken language of two countries is the same, 0 else.[8]
3. $EconomicDep_{ij}$ is a measure for the economic dependence of two states, calculated by the share of Imports and Exports to the respective GDP:

$$EconomicDep_{ij} = min\left(\frac{Trade_{ij}}{GDP_i}, \frac{Trade_{ij}}{GDP_j}\right) \cdot 100\%$$

---

[6]Kim et al. (2018) present a dynamic extension of the AME network model. In Warren (2010) the author proceeds with a *stochastic actor-oriented* approach for the analysis of longitudinal network data. The covariates of interest reflect the research field, but are not directly comparable.

[7]Adding the *Composite Index of National Capability* (CINC) as a measure for military capability, appended in the alliances data set, did not improve model fit during preliminary modeling. It is therefore omitted from the analysis.

[8]I use data from the replication files of Warren (2010) and choose the year 1985, since the year 2000 did not seem to be accurate. Henderson (1997) is provided as the original data source.

To construct this index, I take data from the Correlates of War Project on trade flows (Barbieri and Keshk (2016)) and The World Bank (2021a). Missing data points for GDP are imputed on individual basis, negative values for trade are set 0.

4. $SharedAllies_{ij}$ is the number of shared allies between country i and j.

5. $ConflictInd_{ij}$ is an indicator variable that takes the value 1 if a militarized interstate dispute was recorded in the preceding 10 years between country i and j, 0 else.

6. $PoliticalSim_{ij}$ is constructed using data from the Polity IV Index:

$$PoliticalSim_{ij} = |POLITY_i - POLITY_j|$$

The Polity IV index classifies countries on a 21 point scale from -10 to 10, where -10 to -6 corresponds to autocracies, -5 to 5 to anocracies, and 6 to 10 to democracies. Note that a high value in PoliticalSim corresponds to a high *dissimilarity.*

7. $CapabilityRat_{ij}$ The capability ratio is defined as the log of the relative *Composite Index of National Capability* of both countries, with the stronger state (indicated by the subscript s) as the numerator:

$$CapabilityRat_{ij} = \log\left(\frac{CINC_s}{CINC_w}\right).$$

At the provided precision, three countries are indistinguishable from 0. I impute these data points with 0.5 times of the second to lowest value.

Note that each dyadic fixed effect is symmetric. To validate the network approach taken in this application case, I estimate four models with different restrictions. One with additive and multiplicative effects (AME), one with only additive effects (AE), one with only multiplicative effects (ME), and one without any types of random effects (NO). The latter one yields a classical (bayesian) probit regression, which does not take into account any type of network effect and assumes conditional indepedence of observations. This serves as a benchmark for our validation. Based on preliminary modeling, I choose rank two for the multiplicative effects.[9] The results presented in this section are based on 100,000 iterations and a burn-in of 10,000. Output density is set at 100. Long sampling and stark thinning are due to symmetric matrices of the multiplicative part, which result in considerably slow mixing of the MCMC chains (see P. D. Hoff 2015a, 44) and high autocorrelation of the estimates.

As elaborated in the theory section, model assessment is performed using the goodness of fit statistics of posterior predictive draws. Figure 2 illustrates this comparison for our models. The red line indicates the observed goodness of fit statistic for the sociomatrix. On the left, we have a comparison of first-order network statistics, on the right, third-order. As our matrix is symmetric, second-order dependence, i.e. reciprocity, is constant, obsolete and thus omitted.

---

[9]A two-dimensional latent space facilitates the visualization in the analysis. Moreover, choosing a higher dimension of R = 5 for the latent space did not improve model fit considerably (not shown here, see code appendix for further comparisons).

Table 1: Model Comparison

	AME Network Model (R=2)			Probit Regression		
	pmean	psd	p - value	pmean	psd	p - value
Intercept	-33.186	6.369	0	-0.199	0.144	0.166
GDP (log p.c.).node	1.232	0.303	0	0.044	0.008	0
GeoDistance.dyad	-0.511	0.11	0	-0.191	0.008	0
CulturalSim.dyad	0.484	0.898	0.59	0.69	0.082	0
EconomicDep.dyad	0.454	0.236	0.055	-0.024	0.031	0.437
SharedAllies.dyad	0.261	0.067	0	0.105	0.037	0.005
ConflictInd.dyad	0.201	0.443	0.65	-0.05	0.05	0.32
PoliticalSim.dyad	-0.283	0.059	0	-0.083	0.005	0
CapabilityRat.dyad	0.003	0.143	0.981	-0.074	0.014	0
Variance Parameters						
va	26.493	9.52				
ve	1	0				

[1] Effective Sample Size of the AME Network Model (R=2): 14, 46, 41, 124, 52, 5, 10, 30, 74.

[2] Effective Sample Size of the Probit Regression: 1000, 1000, 1095, 1000, 1000, 7, 624, 1000, 929.

The first row compares the ordinary (bayesian) probit regression (NO) with the additive effects (AE) model. Predictably, the former (pink) does barely represent the network dependencies present in the data. Adding additive effects does improve this representation (blue). An overestimated first-order effect does probably account for the high third-order network effect. Still, both models are not able to capture the extent of statistical dependency in the network. In the second row, the additive effects (AE) model is compared to the multiplicative effects model (ME, R = 2, green), which does improve model fit considerably. First-order effects are captured very well, representation of third-order effects has improved. Comparing multiplicative effects with the full AME Model (grey, third row), both R = 2, show the best fit to the goodness of fit statistics.

Table 1 and 2 provide parameter estimates for the four specifications discussed above. Estimates are provided as posterior means, along with the posterior standard deviation and a p - value.[10] Different specifications lead to vastly different estimates, not only in magnitude but also in changes of the sign.

We restrict our attention to the AME network model with rank two as, based on the network dependencies in figure 2, it exhibits the best fit in the above comparison. In the AME framework, we can interpret the fixed effects similar to the generalized linear regression framework, with the addition that we condition the effect estimate on the additive and multiplicative random effects. All estimates are directionally consistent with our expectations. A low intercept reflects the low

---

[10]The p - value is calculated under the assumption of an approximately normal posterior distribution, i.e., $z = \frac{pmean}{psd} \sim \mathcal{N}(0,1)$. Yet, the p - value should be interpreted with caution, as the density plots do not always suggest approximate normality.
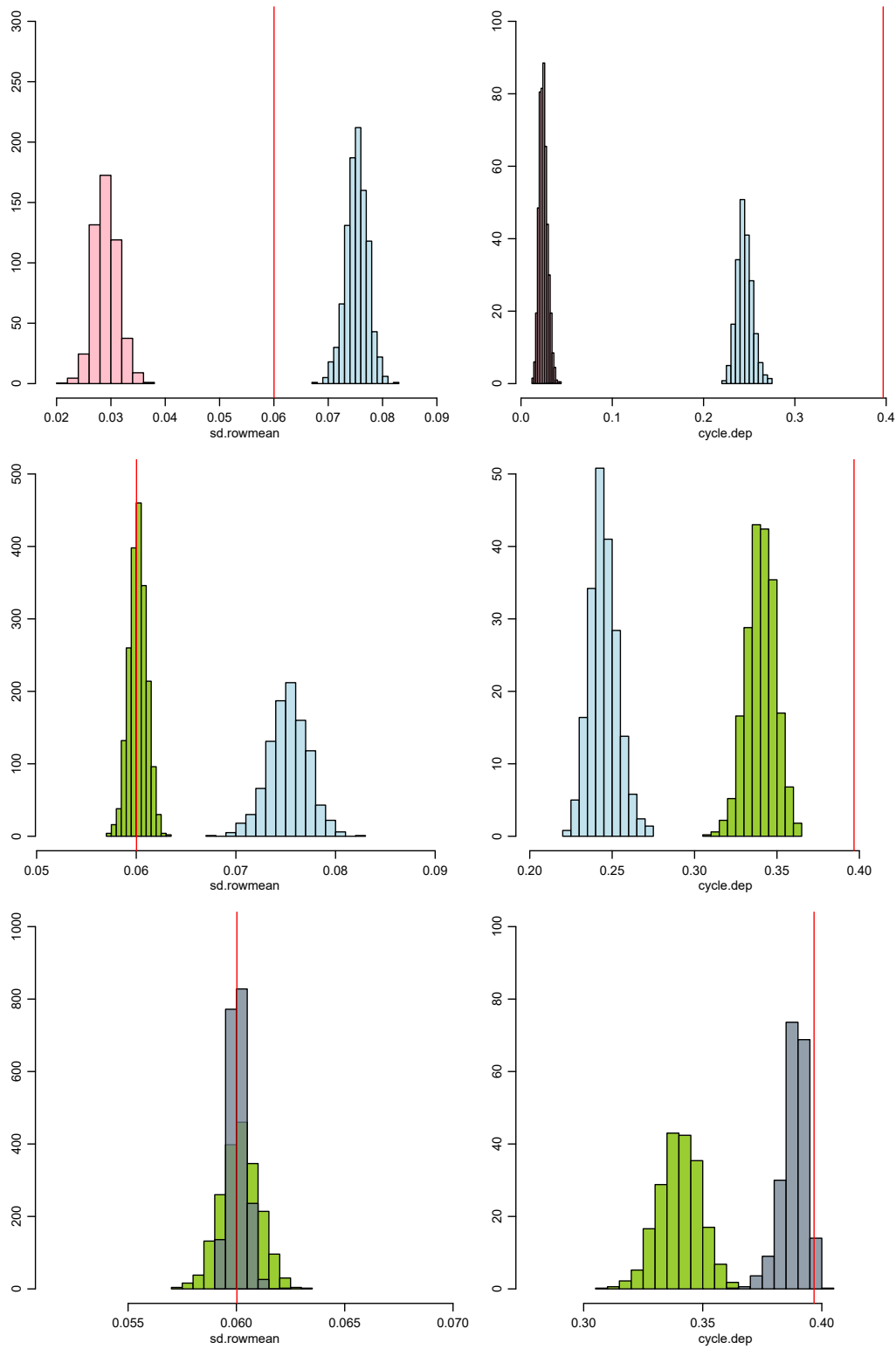
Figure 2: Pairwise comparison (row-wise) of posterior predictive distributions of the goodness of fit statistics for the ordinary regression (pink), only additive effects (blue), only multiplicative effects (green), and both additive and multiplicative effects (grey).

Table 2: Model Comparison

	Additive Effects			Multiplicative Effects (R=2)		
	pmean	psd	p - value	pmean	psd	p - value
Intercept	-0.331	1.738	0.849	-7.136	1.683	0
GDP (log p.c.).node	0.014	0.115	0.905	0.319	0.096	0.001
GeoDistance.dyad	-0.386	0.019	0	-0.396	0.048	0
CulturalSim.dyad	0.524	0.154	0.001	0.239	0.351	0.497
EconomicDep.dyad	0.38	0.103	0	0.009	0.066	0.893
SharedAllies.dyad	1.177	0.106	0	0.05	0.027	0.062
ConflictInd.dyad	-0.263	0.167	0.116	-0.194	0.276	0.482
PoliticalSim.dyad	-0.141	0.009	0	-0.197	0.025	0
CapabilityRat.dyad	-0.162	0.029	0	0.001	0.073	0.985
Variance Parameters						
va	4.76	1.051				
ve	1	0				

[1] Effective Sample Size of the Additive Effects Model: 755, 672, 133, 1000, 41, 2, 53, 123, 264.

[2] Effective Sample Size of the Multiplicative Effects Model: 51, 52, 25, 270, 745, 19, 26, 75, 109.

density (0.061) of the present interstate alliance network. A higher GDP (log p. c.) is associated *ceteris paribus* and conditional on the random effects with a higher propensity of interstate alliance formation. Geographic distance has a negative effect. Cultural similarity as reflected by joint language, a higher economic dependence, number of shared allies, and previous conflict are all associated with the presence of alliances. In political similarity, a higher dissimilarity is associated with a negative effect on alliance formation, whereas differences in military capability show no effect.

In the AME network model framework, statistical network dependencies will be encoded in the additive and multiplicative, i.e., the random-effects, component of the model. We expect the multiplicative component to capture third-order network effects and cluster types of states which share similar characteristics with respect to their tie formation in the network to other groups, while additive effects should capture the actor-specific propensity of tie formation. Figure 3 visualizes the latent space obtained with the AME (R=2) model. The multiplicative effect (matrix notation) in a symmetric model is $ULU^\top$, where $U \in \mathbb{R}^{158 \times 2}$ and $L \in \mathbb{R}^{2 \times 2}$. Both U and L are provided as posterior means in the package output, and to visualize the latent factors the decomposition of this matrix is computed: $\mathbf{u}_i \sqrt{L}$ for $i = 1, \ldots, 158$.

Apart from the scatter in the center of the plot, the multiplicative component captures some interesting geographical factors, which were already apparent in Figure 1, and are not captured within the $GeoDistance_{ij}$ fixed effect. This includes the Americas on the left and at least two distinct clusters of African states. Russia and some of the former USSR states are clustered in the lower left, the right side has the Middle East and North African countries.

Additive effects are plotted in Figure 4. Among the five countries with the highest posterior mean additive effects are Canada, the United States, and France. Switzerland, Sweden, and Austria list among the five countries with the lowest values. This is as expected, as additive effects are added to capture first-order effects, i.e., the propensity to form a network tie.
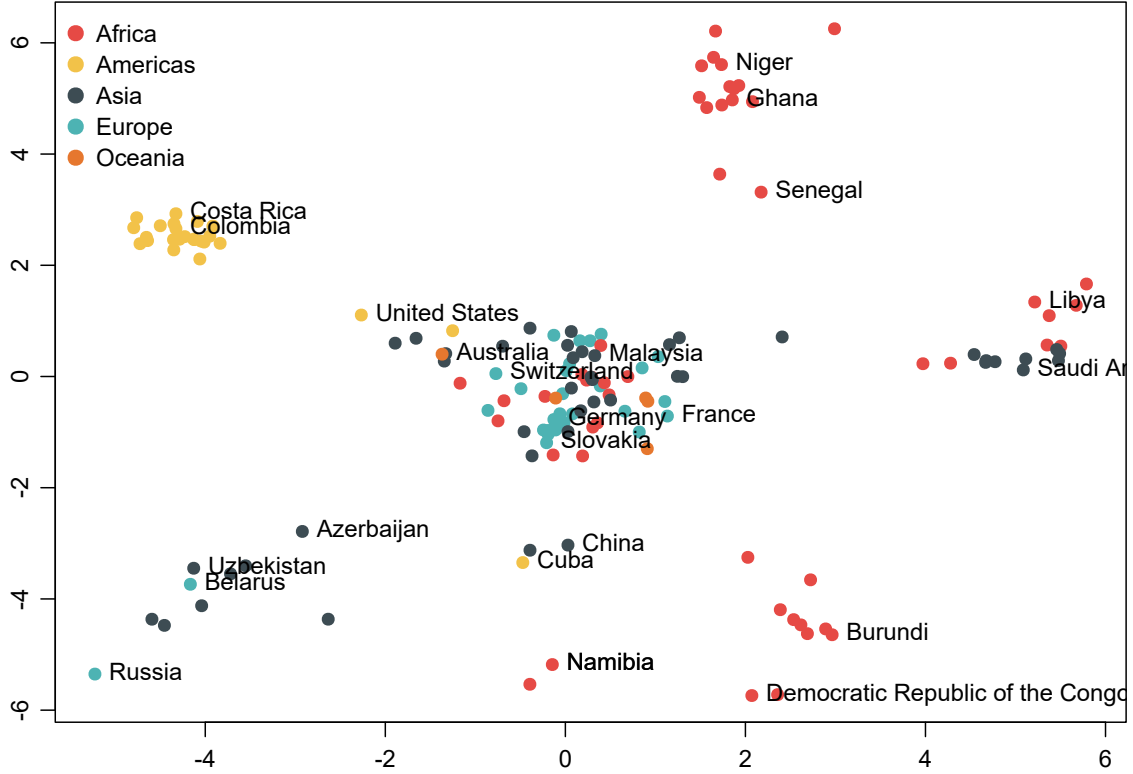

Figure 3: Visualization of the latent space in the AME network model with rank 2. Values are posterior mean estimates for the multiplicative random effects. Proximity in the latent space indicates similarity with respect to tie formation.

Figure 5 shows the trace plots for the nine fixed effects and the variance of the additive effect in this estimation. Both multimodality and persistent trending are apparent, even *after* strong thinning of 1 to 100. Specifically for the variance of the additive effect and the intercept, which suggests issues with the mixing of the MCMC chains. This is also indicated by the low effective sample size (Table 1) for the specifications including multiplicative effects.

To adapt the model, I estimated the following changed specifications of the AME rank 2 model. First, I dropped the Intercept, which yielded worse effective sample sizes. Predictably, the variance of the additive effect increased, along with a change of sign in the fixed effects estimates of *GDP (log p.c.), ConflictInd*, and *CapabilityRat*. Overall it did not seem to improve the estimation.[11]

Next, I dropped the two covariates with the lowest effective sample size, *SharedAllies* and *ConflictInd*.

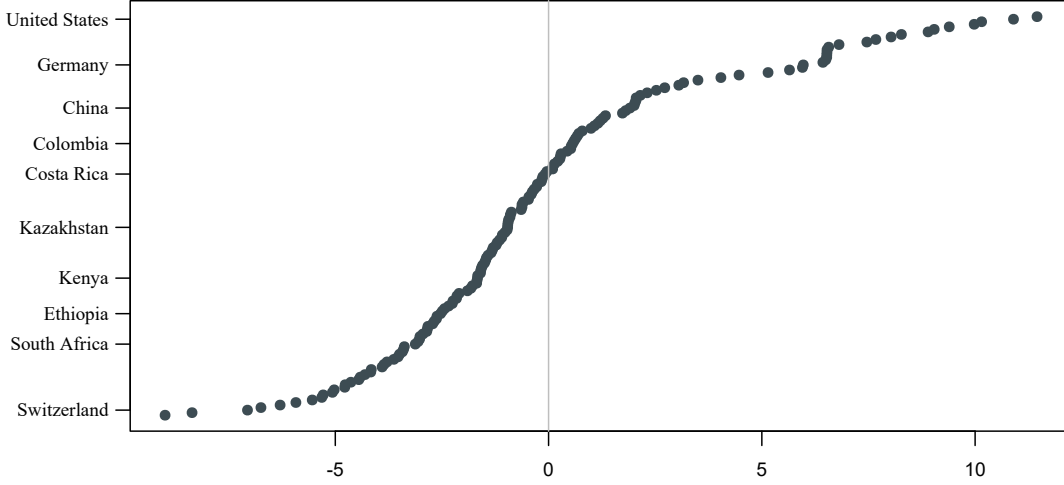[11]This and the following model can be found in the online code appendix and are not shown here.

Figure 4: Additive Random Effects in the AME network model with rank 2. Values are posterior mean estimates for the additive random effects. AE effects capture first-order network effects, i.e., actor-specific propensity to form a network tie.

Except for the intercept, which increased by a third, the estimated variance of the additive effect and the fixed effect estimates remained virtually the same. The issues with the trace plots and the low effective sample size remained.

Then, I removed all non-significant fixed covariates, leaving only the *Intercept, GeoDistance, EconomicDep, SharedAllies*, and *PoliticalSim*. Again, these changes did not seem to substantially improve the model, the results are shown in Table 3. This posits the general question of the selection of fixed covariates in the model. Estimating an AME rank 2 network model without *any* fixed effects did *not* depress substantially the representation of network dependencies in the estimated model. In practice, this makes covariate selection troublesome.

Lastly, longer burn-in and sampling durations could potentially improve the estimates. However, drastically raising these numbers would push computational feasibility. To verify robustness with respect to the length of the MCMC estimation, I raise the number of iterations to 400,000 after a burn-in period of 100,000.[12] Output density is kept the same. Posterior estimates for this model are in Table 3. The intercept estimate is lower, and the estimate for *ConflictInd* is now negative, although still non-significant. The other fixed-effects estimates are qualitative as before. Moreover, the effective sample size has *not* improved. In the random effects, we see drastically increased variance of the additive effect and the corresponding standard error. This is also apparent upon further inspection of the trace plots in Figure 6. Figure 7 shows a comparison of the posterior goodness of fit statistics for both AME rank 2 models, with the only difference in the length of the

---

[12]The models of the first comparison took about two hours to estimate. Quadrupling the number of iterations in the last model required about ten hours to compute.
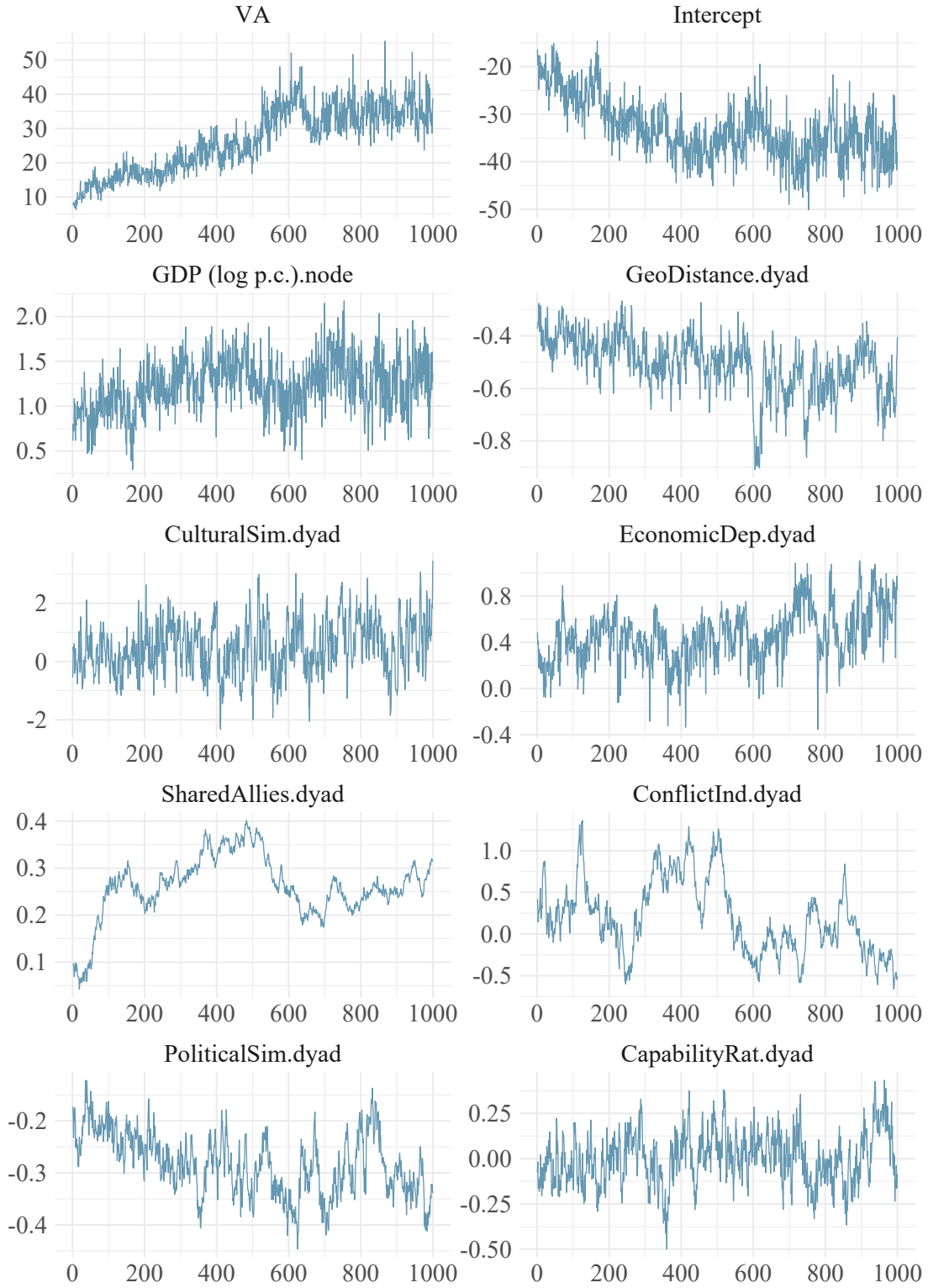
Figure 5: Trace plots for the AME rank 2 model. Trending and multimodality is visible for several parameter trace plots of the MCMC estimation.

Table 3: Model Comparison

	AME (R=2, selected)			AME (R=2, 4x)		
	pmean	psd	p - value	pmean	psd	p - value
Intercept	-18.382	10.263	0.073	-37.519	6.168	0
GDP (log p.c.).node	0.478	0.482	0.321	1.196	0.345	0.001
GeoDistance.dyad	-0.752	0.115	0	-0.681	0.129	0
CulturalSim.dyad				0.439	1.205	0.715
EconomicDep.dyad	0.416	0.23	0.07	0.714	0.366	0.051
SharedAllies.dyad	0.252	0.056	0	0.294	0.128	0.022
ConflictInd.dyad				-0.383	0.585	0.513
PoliticalSim.dyad	-0.3	0.07	0	-0.375	0.071	0
CapabilityRat.dyad				0.025	0.211	0.907
AME Components						
va	25.65	11.908		54.731	18.034	
ve	1	0		1	0	

[1] Effective Sample Size: 2, 3, 39, 41, 6, 16.

[2] Effective Sample Size: 45, 73, 103, 193, 75, 2, 11, 56, 122.

MCMC estimation. The increase yields narrower distributions for the goodness of fit statistics but does not improve on model fit.

3.3 Limitations and Further Approaches

Before entering the limitations and the discussion on the AME network model, two possible approaches to improve the present issues follow:

1. Changes in model specifications above did only affect dyadic covariates. Given the persistent trend in the variance of the additive effect and the intercept, one could hypothesize that the actor-specific propensity of tie formation is not sufficiently considered within the models' fixed effects. This could be alleviated by adding more nodal covariates. Unfortunately, there was no evident political or economic factor that could have been included. One option could be to construct an index of vote patterns in the United Nations, along the lines of the outcome variable in Kim et al. (2018). This is left for further exploration.

2. Wide posterior intervals indicate a wide distribution about the parameter estimates of the intercept and the variance of the additive effects. This could present problems with the posterior estimation.[13] A natural step would be to specify different starting values as well as stronger priors, specifically for the intercept. The network density[14] ranges from 0.046 to 0.057 in the time-frame 1981 - 2000. Assuming this low density is a reliable characteristic of the interstate alliance network,

---

[13]See for example Gelman (2021).

[14]This calculation takes into account every state in the data, it therefore slightly understimates the density provided before.
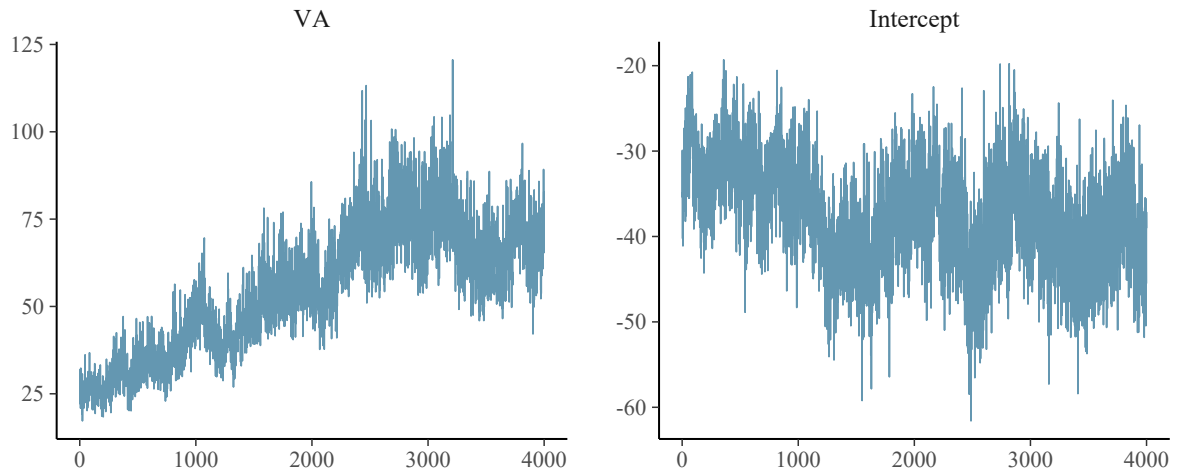
Figure 6: Trace plots for estimated posterior variance of the additive effect and the intercept. Results based on 400,000 iterations and a burn-in period of 100,000. Chains were thinned by keeping every 100th sample.
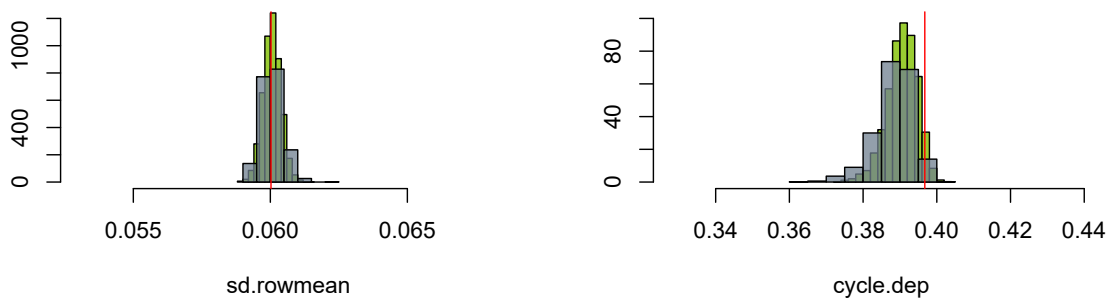

Figure 7: Posterior Goodness of Fit statistics for model estimates with 100,000 iterations (grey) and 400,000 iterations (green). A narrower distribution can be identified for the green model.

a narrower prior distribution, could be selected for the intercept.

The low effective sample size and the troublesome trace plots of the AME network model leave doubts about the validity of the statistical inference in this application case. Further analysis to improve the model fitting procedure would need to tackle the MCMC estimation procedure with more elaborated tools on the analysis of different chains as the visual approach provided in this seminar paper. The present issues could have their roots in different aspects of the chosen data set. The absence of strong predictors for the actor-specific propensity of tie formation, for example. A very low network density, along with large heterogeneity in the network actors too; as well as the symmetric, binary outcome.

4   Discussion

The Additive and Multiplicative Effects network model provides an extension to the already familiar generalized linear model framework to network data. By adding a node-specific additive random effect and a latent space as a multiplicative effect, an AME model can capture a range of statistical network effects that are inherent in dyadic data and ordinary regression frameworks that fail to account for (Figure 2). The deliberate omission of these effects can lead to significantly biased estimates of effect sizes. In this regard, the AME Model presented in P. D. Hoff (2021) provides an easy-to-use and interpretable model for network data.

However, the Bayesian estimation procedure makes it difficult to pin down the issues in the estimation, for example in this paper, and thus confidently establish valid statistical inference. It requires furthermore a different approach to model assessment and goodness of fit, specifically with covariate selection in mind. With the added benefit of the multiplicative effect capturing omitted variables, the omission of potentially relevant predictors is probably less distorting than in a classical regression approach. The individual quality of results although may vary and strongly depend on network size, density, and available predictors, in combination with the type of outcome variable and available data. Yet, as a versatile approach in statistical modeling of networks, it can uncover interesting features about the data at hand, for example in the preceding application. In conjunction with other modeling approaches, it thus can provide an important starting point to the analysis of network data.

The AME framework has found it's way into the applied network analysis literature. Dorff, Gallop, and Minhas (2020) apply this model to study civil conflict in Nigeria and the implications of the emergence of Boko Haram on the conflict structure. In Gade, Hafez, and Gabbay (2019) it is taken to study the Syrian civil conflict. The AME model framework has also been expanded to include longitudinal networks or replicated data. Minhas, Hoff, and Ward (2016) use a time series approach of the AME network model presented in P. D. Hoff (2015b) to study international relations, in this case, the type of interaction chosen by each state. The time dependency is modeled using a vector autoregression (VAR) approach. Kim et al. (2018) model UN Voting Behaviour in a similar fashion

with a longitudinal extension of the AME framework, relying on Gaussian processes.

References

Barbieri, Katherine, and Omar M. G. Omar Keshk. 2016. "Correlates of War Project Trade Data Set Codebook, Version 4.0." 2016. http://correlatesofwar.org.

Cranmer, Skyler J., Bruce A. Desmarais, and Elizabeth J. Menninga. 2012. "Complex Dependencies in the Alliance Network." *Conflict Management and Peace Science* 29 (3): 279–313. https://doi.org/10.1177/0738894212443446.

Cranmer, Skyler, Bruce Desmarais, and Justin Kirkland. 2012. "Toward a Network Theory of Alliance Formation." *International Interactions - INT INTERACT* 38 (July): 295–324. https://doi.org/10.1080/03050629.2012.677741.

Dorff, Cassy, Max Gallop, and Shahryar Minhas. 2020. "Networks of Violence: Predicting Conflict in Nigeria." *The Journal of Politics* 82 (2): 476–93. https://doi.org/10.1086/706459.

Gade, Emily Kalah, Mohammed M Hafez, and Michael Gabbay. 2019. "Fratricide in Rebel Movements: A Network Analysis of Syrian Militant Infighting." *Journal of Peace Research* 56 (3): 321–35. https://doi.org/10.1177/0022343318806940.

Gelman, Andrew. 2021. "When MCMC Fails: The Advice We're Giving Is Wrong. Here's What We You Should Be Doing Instead. (Hint: It's All about the Folk Theorem.)." https://statmodeling.stat.columbia.edu/2021/06/10/when-mcmc-fails-the-advice-were-giving-is-wrong-heres-what-we-you-should-be-doing-instead-hint-its-all-about-the-folk-theorem/.

Henderson, Errol A. 1997. "Culture or Contiguity: Ethnic Conflict, the Similarity of States, and the Onset of War, 1820-1989." *Journal of Conflict Resolution* 41 (5): 649–68. https://doi.org/10.1177/0022002797041005003.

Hoff, Peter D. 2005. "Bilinear Mixed-Effects Models for Dyadic Data." *Journal of the American Statistical Association* 100 (469): 286–95. https://doi.org/10.1198/016214504000001015.

———. 2007. "Modeling Homophily and Stochastic Equivalence in Symmetric Relational Data." In *Proceedings of the 20th International Conference on Neural Information Processing Systems*, 657–64. NIPS'07. Red Hook, NY, USA: Curran Associates Inc.

———. 2015a. "Dyadic Data Analysis with Amen." https://arxiv.org/abs/1506.08237.

———. 2015b. "Multilinear tensor regression for longitudinal relational data." *The Annals of Applied Statistics* 9 (3): 1169–93. https://doi.org/10.1214/15-AOAS839.

———. 2021. "Additive and Multiplicative Effects Network Models." *Statistical Science* 36 (1): 34–50. https://doi.org/10.1214/19-STS757.

Hoff, Peter, Bailey Fosdick, and Alex Volfovsky. 2020. *Amen: Additive and Multiplicative Effects Models for Networks and Relational Data.* https://CRAN.R-project.org/package=amen.

Kim, Bomin, X. Niu, D. Hunter, and X. Cao. 2018. "A Dynamic Additive and Multiplicative Effects Model with Application to the United Nations Voting Behaviors." *arXiv: Applications.*

Minhas, Shahryar, Peter D. Hoff, and Michael D. Ward. 2016. "A New Approach to Analyzing Coevolving Longitudinal Networks in International Relations." *Journal of Peace Research* 53 (3): 491–505. https://doi.org/10.1177/0022343316630783.

———. 2019. "Inferential Approaches for Network Analysis: AMEN for Latent Factor Models." *Political Analysis* 27 (2): 208–22. https://doi.org/10.1017/pan.2018.50.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *The American Political Science Review* 93 (2): 279–97. http://www.jstor.org/stable/2585396.

The World Bank. 2021a. "GDP (Current USD)." data retrieved from World Development Indicators, http://data.worldbank.org/indicator/NY.GDP.MKTP.CD, (Accessed on 12/06/2021).

———. 2021b. "GDP Per Capita (Current USD)." data retrieved from World Development Indicators, http://data.worldbank.org/indicator/NY.GDP.PCAP.CD, (Accessed on 12/06/2021).

Warren, T. Camber. 2010. "The Geometry of Security: Modeling Interstate Alliances as Evolving Networks." *Journal of Peace Research* 47 (6): 697–709. https://doi.org/10.1177/002234331038 6270.