

## Taking dyads seriously

Shahryar Minhas<sup>1\*</sup> , Cassy Dorff<sup>2</sup>, Max B. Gallop<sup>3</sup>, Margaret Foster<sup>4</sup>, Howard Liu<sup>5</sup>, Juan Tellez<sup>6\*\*</sup> and Michael D. Ward<sup>7†</sup>

<sup>1</sup>Department of Political Science, Michigan State University, East Lansing, MI, USA, <sup>2</sup>Department of Political Science, Vanderbilt University, Nashville, TN, USA, <sup>3</sup>Department of Government and Public Policy, University of Strathclyde, Glasgow, Scotland, UK, <sup>4</sup>Department of Political Science, University of North Carolina, Chapel Hill, NC, USA, <sup>5</sup>Department of Government, University of Essex, Colchester, England, UK, <sup>6</sup>Department of Political Science, University of South Carolina, Columbia, SC, USA and <sup>7</sup>Department of Political Science, Duke University, Durham, NC, USA

\*Corresponding author. Email: [minhash@msu.edu](mailto:minhash@msu.edu)

(Received 15 July 2019; revised 9 February 2021; accepted 18 February 2021; first published online 15 November 2021)

### Abstract

International relations scholarship concerns dyads, yet standard modeling approaches fail to adequately capture the data generating process behind dyadic events and processes. As a result, they suffer from biased coefficients and poorly calibrated standard errors. We show how a regression-based approach, the Additive and Multiplicative Effects (AME) model, can be used to account for the inherent dependencies in dyadic data and glean substantive insights in the interrelations between actors. First, we conduct a simulation to highlight how the model captures dependencies and show that accounting for these processes improves our ability to conduct inference on dyadic data. Second, we compare the AME model to approaches used in three prominent studies from recent international relations scholarship. For each study, we find that compared to AME, the modeling approach used performs notably worse at capturing the data generating process. Further, conventional methods misstate the effect of key variables and the uncertainty in these effects. Finally, AME outperforms standard approaches in terms of out-of-sample fit. In sum, our work shows the consequences of failing to take the dependencies inherent to dyadic data seriously. Most importantly, by better modeling the data generating process underlying political phenomena, the AME framework improves scholars' ability to conduct inferential analyses on dyadic data.

**Keywords:** Bayesian; latent variable models; multi-level and hierarchical models

### 1. Introduction

The aim of this study is to address how to estimate regression coefficients in a generalized linear model (GLM) context when there are network dependencies in dyadic data. Specifically, we discuss and evaluate how well the Additive and Multiplicative Effects (AME) model can be used to account for the interdependencies underlying the data generating process of dyadic structures (Hoff, 2005, 2008; Minhas *et al.*, 2019) in International Relations data. The AME works by including a set of parameters meant to capture network effects in the conditional mean equation of the GLM.

We focus on three types of network effects that can complicate dyadic analyses. First, dependencies may arise within a set of dyads if a particular actor is more likely to send or receive actions such as conflict.<sup>1</sup> Additionally, if the event of interest has a clear sender and receiver, we are likely to observe dependencies within a dyad; for example, if a rebel group initiates a conflict against a

\*\*Present address: Department of Political Science, University of California, Davis, USA

†Deceased

<sup>1</sup>In the case of undirected data where there is no clear sender or receiver, it is still essential to take into account the variance in how active actors are in the system.

government, the government will likely reciprocate that behavior. We capture these effects, often referred to as first- and second-order dependencies, respectively, within the additive effects portion of the model. Third-order dependencies capture relationships of transitivity, balance, and clusterability between different dyads. For example, we can only understand why Poland was involved in a dyadic conflict with Iraq in 2003 if we understand that the United States invaded Iraq in 2003 and that Poland often participates in US-led coalitions. The multiplicative effects capture these sorts of dependencies, specifically, those that result because the specified model has not accounted for a latent set of shared attributes that affect actors' probability of interacting with one another.

We begin with a discussion of these dependencies and an introduction to the AME model. Next, we conduct a simulation study to show how the AME approach can recover unbiased and well-calibrated regression coefficients in the presence of network-based dependence. Last, to highlight the utility of this approach, we apply the AME model to three recent studies in the international relations (IR) literature. Our comparison reveals that accounting for observational dependence leads to results that, at times, differ from those found in the original study as well as from the broader literature. Moreover, we demonstrate that the additional parameters included by AME in the conditional mean equation of a typical GLM can offer substantive insights that are often occluded by ignoring the interdependencies found in relational data. Finally, we show that for each replication our network-based approach provides substantially more accurate out-of-sample predictions than the models used in the original studies.

The AME approach advances statistical analysis of dyadic data by accounting for observational dependence while allowing scholars to test the substantive effect of variables of interest. Thus, the AME allows scholars to achieve a twofold goal: to continue to generate meaningful, substantive insights about political phenomena without abandoning a regression-based approach, while at the same time accounting for the data generating processes behind such events of interest. Perhaps most importantly, the AME approach concentrates on the relational aspect of international relations through a statistical framework that is familiar to most scholars.

## 2. Dependencies in dyadic data

When modeling dyadic data, scholars typically employ a GLM estimated via maximum-likelihood. This type of model is expressed via a stochastic and systematic component (Pawitan, 2013). The stochastic component reflects our 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, and we assume that each dyad in the sample is independently drawn from that particular distribution, given  $\theta_{ij}$ . The systematic component characterizes the model for the parameters of that distribution and describes how  $\theta_{ij}$  varies as a function of a set of nodal and dyadic covariates,  $\mathbf{X}_{ij}$ :  $\theta_{ij} = \beta^T \mathbf{X}_{ij}$ . A fundamental assumption we make when applying this modeling technique is that given  $\mathbf{X}_{ij}$  and the parameters of our 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, we construct the joint density function over all dyads, for example:

$$\begin{aligned} P(y_{ij}, y_{ik}, \dots | \theta_{ij}, \theta_{ik}, \dots) &= P(y_{ij} | \theta_{ij}) \times P(y_{ik} | \theta_{ik}) \times \dots \\ P(\mathbf{Y} | \boldsymbol{\theta}) &= \prod_{\alpha=1}^{n \times (n-1)} P(y_{\alpha} | \theta_{\alpha}) \end{aligned} \quad (1)$$

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

We can then estimate the parameters by maximizing the likelihood, or, more typically, the log-likelihood. However, the likelihood as defined above is only valid if we are able to make the assumption that, for example,  $y_{ij}$  is independent of  $y_{ji}$  and  $y_{ik}$  given the set of covariates

Sender Heterogeneity		Receiver Heterogeneity									
		<i>i</i>	<i>j</i>	<i>k</i>	<i>l</i>			<i>i</i>	<i>j</i>	<i>k</i>	<i>l</i>
<i>i</i>	NA	$y_{ij}$	$y_{ik}$	$y_{il}$		<i>i</i>	NA	$y_{ij}$	$y_{ik}$	$y_{il}$	
<i>j</i>	$y_{ji}$	NA	$y_{jk}$	$y_{jl}$		<i>j</i>	$y_{ji}$	NA	$y_{jk}$	$y_{jl}$	
<i>k</i>	$y_{ki}$	$y_{kj}$	NA	$y_{kl}$		<i>k</i>	$y_{ki}$	$y_{kj}$	NA	$y_{kl}$	
<i>l</i>	$y_{li}$	$y_{lj}$	$y_{lk}$	NA		<i>l</i>	$y_{li}$	$y_{lj}$	$y_{lk}$	NA	

Sender-Receiver Covariance		Reciprocity									
		<i>i</i>	<i>j</i>	<i>k</i>	<i>l</i>			<i>i</i>	<i>j</i>	<i>k</i>	<i>l</i>
<i>i</i>	NA	$y_{ij}$	$y_{ik}$	$y_{il}$		<i>i</i>	NA	$y_{ij}$	$y_{ik}$	$y_{il}$	
<i>j</i>	$y_{ji}$	NA	$y_{jk}$	$y_{jl}$		<i>j</i>	$y_{ji}$	NA	$y_{jk}$	$y_{jl}$	
<i>k</i>	$y_{ki}$	$y_{kj}$	NA	$y_{kl}$		<i>k</i>	$y_{ki}$	$y_{kj}$	NA	$y_{kl}$	
<i>l</i>	$y_{li}$	$y_{lj}$	$y_{lk}$	NA		<i>l</i>	$y_{li}$	$y_{lj}$	$y_{lk}$	NA	

Figure 1. Nodal and dyadic dependencies in relational data.

we specified, or the values of  $\theta_{ij}$ . Without the assumption of conditional observational independence the joint density function cannot be written in the way described above and a valid likelihood does not exist. 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. The importance of accounting for the underlying structure of our data has been a lesson well understood, at least when it comes to time-series cross-sectional data (TSCS) within political science (Beck and Katz, 1995; Beck *et al.*, 1998). As a result, it is now standard practice to take explicit steps to account for the complex data structures that emerge in TSCS applications and the unobserved heterogeneity that they cause.

To uncover the underlying structure of relational data, it is helpful to restructure dyadic data in the form of a matrix—often referred to as an adjacency matrix—as shown in Figure 1. Rows designate the senders of an event and columns the receivers. The cross-sections in this matrix represent the actions that were sent by an actor in the row to those designated in the columns. Thus  $y_{ij}$  designates an action  $y$ , such as a conflictual event or trade flow, that is sent from actor  $i$  to actor  $j$ . In many applications, scholars are interested in studying undirected (i.e., symmetric) outcomes in which there is no clear sender or receiver, these types of outcomes still can, and we argue should, be studied using the type of framework we discuss below.

Using the structure of an adjacency matrix, Figure 1 visualizes the types of first- and second-order dependencies that can complicate the analysis of relational data in traditional

GLMs. The adjacency matrix on the top left highlights a particular row to illustrate that these values may be more similar to each other than other values because each has a common sender. Interactions involving a common sender also manifest heterogeneity in how active actors are across the network when compared to each other. In most relational datasets (e.g., trade flows, conflict, participation in international organizations, even networks derived from Twitter or Facebook), we often find that there are some actors that are much more active than others (Barabási and Réka, 1999). For example, in an analysis of international trade, certain countries (e.g., China) export much larger volumes than other countries for a variety of structural, contextual, and idiosyncratic reasons. Unless one is able to develop a model that can account for the variety of explanations that may play a role in determining why a particular actor is more active than others, parameter estimates from standard statistical models will be biased.<sup>2</sup>

For similar reasons one also needs to take into account the dependence between observations that share a common receiver. The bottom-left panel in Figure 1 illustrates that sender and receiver type dependencies can also blend together. Specifically, actors who are more likely to send ties in a network tend to also be more likely to receive them. As a result, the rows and columns in an adjacency matrix are often correlated. For example, consider trade flows both from and to many wealthy, developed countries. The bottom-right panel highlights a second-order dependence, specifically, reciprocity. This is a dependency occurring within dyads involving the same actors whereby values of  $y_{ij}$  and  $y_{ji}$  are correlated. The concept of reciprocity has deep roots in the study of relations between states (Richardson, 1960; Keohane, 1989).

For most relational data, however, dependencies do not simply manifest at the nodal or dyadic level. More often we find significant evidence of higher-order structures that result from dependencies between multiple groups of actors. These dependencies arise because there may be a set of latent attributes between actors that affect their probability of interacting with one another (Zinnes, 1967; Wasserman and Faust, 1994). In Figure 2, we provide a visualization of a simulated relational dataset wherein the nodes designate actors and edges between the nodes indicate that an interaction between the two took place. To highlight third-order dependence patterns, nodes with similar latent attributes are colored similarly.

The visualization illustrates that actors belonging to the same group have a higher likelihood of interacting with each other, whereas interactions across groups are rarer. A prominent example of a network with this type of structure is discussed by Adamic and Glance (2005), who visualize linkages between political blogs preceding the 2004 United States Presidential Election. Adamic and Glance (2005) find that the degree of interaction between right- and left-leaning blogs is minimal and that most blogs are linked to others that are politically similar. This showcases the types of higher-order dependencies that can emerge in relational data. First, the fact that interactions are determined by a shared attribute, in this case political ideology, is an example of *homophily*. Homophily explains the emergence of patterns such as transitivity (“a friend of a friend is a friend”) and balance (“an enemy of a friend is an enemy”), which also have a long history in international relations. The other major type of meso-scopic features that emerge in relational data is community structure (Mucha *et al.*, 2010), which is often formalized through the concept of stochastic equivalence (Anderson *et al.*, 1992). Stochastic equivalence refers to a type of pattern in which actors can be divided into groups such that members of the same group have similar patterns of relationships. In the example above, each of the left leaning blogs would be considered stochastically equivalent to one another because any given left-leaning blog is more likely to interact with a blog of a similar political position and less likely to interact with one of a divergent political position.

These types of patterns frequently emerge in IR contexts.<sup>3</sup> For example, a perennial finding in the interstate trade literature emphasizes the role that geography plays in determining trade flows.

<sup>2</sup>In an undirected setting instead of studying sender and receiver heterogeneity, we would just be concerned with actor heterogeneity in general.

<sup>3</sup>For example, see: Manger *et al.* (2012), Kinne (2013), and Chyzh (2016).

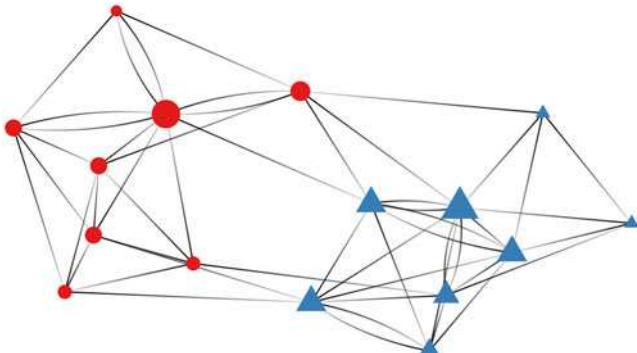


Figure 2. Higher-order dependence patterns in a network.

Geographic proximity in the network context is an example of homophily—a shared attribute between actors that corresponds to a greater likelihood of the event of interest taking place. Alternatively, in the interstate conflict literature, we may find that actors who are each a member of a particular (formal or informal) alliance are likely to act similarly in the conflict network. Specifically, they will tend to initiate conflictual events with actors that their fellow alliance members initiate conflict with, and they will be unlikely to initiate conflict with the members of their alliance—an example of stochastic equivalence. In both these examples, we are able to explicitly parameterize the attribute that might explain the emergence of higher order dependence patterns. While sometimes the conditions driving these patterns, such as geography, are easy to identify, at other times it can be difficult to describe exactly why higher order dependence patterns in networks may develop.

### 3. Additive and multiplicative effect models for networks

To account for the dependencies that are prevalent in dyadic data, we turn to the AME model. The AME approach can be used to conduct inference on cross-sectional and longitudinal networks with binary, ordinal, or continuous linkages. It is flexible and easy to use for analyzing the kind of relational data often found in the social sciences. It accounts for nodal and dyadic dependence patterns, as well as higher-order dependencies such as homophily and stochastic equivalence.<sup>4</sup> The AME model combines the social relations regression model (SRRM) to account for nodal and dyadic dependencies and the latent factor model (LFM) for third-order dependencies.<sup>5</sup> For details on the SRRM see Li and Loken (2002), Hoff (2005), and Dorff and Minhas (2017). The AME model is specified as follows:

$$\begin{aligned}
 y_{ij} &= f(\theta_{ij}), \quad \text{where} \\
 \theta_{ij} &= \beta_d^\top \mathbf{X}_{ij} + \beta_s^\top \mathbf{X}_i + \beta_r^\top \mathbf{X}_j \quad (\text{Exogenous parameters}) \\
 &\quad + a_i + b_j + \epsilon_{ij} \quad (\text{SRRM parameters}) \\
 &\quad + \mathbf{u}_i^\top \mathbf{D} \mathbf{v}_j \quad (\text{LFM parameters})
 \end{aligned} \tag{2}$$

where  $y_{ij,t}$  captures the interaction between actor  $i$  (the sender) and  $j$  (the receiver) at time  $t$ . We model a latent variable,  $\theta_{ij}$ , first using a set of exogenous dyadic ( $\beta_d^\top \mathbf{X}_{ij}$ ), sender ( $\beta_s^\top \mathbf{X}_i$ ), and receiver covariates ( $\beta_r^\top \mathbf{X}_j$ ).  $f$  is typically a mapping function, and can be one that applies to dichotomous, ordinal, or continuous distributions.

<sup>4</sup>Minhas *et al.* (2019) detail how this framework contrasts with alternative network-based approaches.

<sup>5</sup>An earlier version of the LFM used in AME is presented as the general bilinear mixed effects (GBME) model in Hoff (2005). The GBME model is more limited in the types of dependence patterns that it can capture due to the formulation of the matrix decomposition procedure.

Next, to account for the dependencies that emerge in dyadic data that may complicate inference on the parameter associated with exogenous covariates, we add parameters from the SRRM and LFM.  $a_i$  and  $b_j$  in Equation 2 represent sender and receiver random effects incorporated from the SRRM framework:

$$\begin{aligned} \{(a_1, b_1), \dots, (a_n, b_n)\} &\text{iid} \sim N(0, \Sigma_{ab}) \\ \{(\epsilon_{ij}, \epsilon_{ji}) : i \neq j\} &\text{iid} \sim N(0, \Sigma_{\epsilon}), \quad \text{where} \\ \Sigma_{ab} &= \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix} \quad \Sigma_{\epsilon} = \sigma_{\epsilon}^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \end{aligned} \quad (3)$$

The sender and receiver random effects are modeled jointly from a multivariate normal distribution to account for correlation in how active an actor is in sending and receiving ties. Heterogeneity in the sender and receiver effects is captured by  $\sigma_a^2$  and  $\sigma_b^2$ , respectively, and  $\sigma_{ab}$  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  $\sigma_{\epsilon}^2$  and a within dyad correlation, or reciprocity, parameter  $\rho$ .

The LFM contribution to the AME is in the multiplicative term:  $\mathbf{u}_i^\top \mathbf{D} \mathbf{v}_j = \sum_{k \in K} d_k u_{ik} v_{jk}$ .  $K$  denotes the dimensions of the latent space. The construction of the LFM here is actually quite similar to work on low rank approximations in computer science and has been applied to the development of recommender systems that companies like Amazon and Netflix use to model customer behavior (Resnick and Varian, 1997; Bennett and Lanning, 2007).<sup>6</sup> This model posits a latent vector of characteristics  $\mathbf{u}_i$  and  $\mathbf{v}_j$  for each sender  $i$  and receiver  $j$ . The similarity or dissimilarity of these vectors will then influence the likelihood of activity, and provides a representation of third-order interdependencies. The LFM parameters are estimated by a process similar to computing the singular value decomposition (SVD) of the observed network. When computing the SVD we factorize our observed network into the product of three matrices:  $\mathbf{U}$ ,  $\mathbf{D}$ , and  $\mathbf{V}$ . This provides us with a low-dimensional representation of our original network.<sup>7</sup> Values in  $\mathbf{U}$  provide a representation of how stochastically equivalent actors are as senders in a network or, for example, how similar actors are in terms of who they initiate conflict with.  $\hat{\mathbf{u}}_i \approx \hat{\mathbf{u}}_j$  would indicate that actors  $i$  and  $j$  initiate battles with similar third actors.  $\mathbf{V}$  provides a similar representation but from the perspective of how similar actors are as receivers. The values in  $\mathbf{D}$ , a diagonal matrix, represent the levels of homophily in the network.<sup>8</sup> Note that this model easily generalizes to the case, common in IR, where interactions are undirected (e.g., the presence of conflict or a bilateral investment treaty). In the case of the SRRM,  $\rho$  is constrained to be one and instead of separate sender and receiver random effects a single actor random effect is utilized. For the LFM, an eigen-decomposition scheme is used to capture higher-order dependence patterns. In the application section, we show the applicability of the AME approach to both directed and undirected dyadic data. Parameter estimation in the AME takes place within the context of a Gibbs sampler in which we iteratively sample from the posterior distribution of the full conditionals for each parameter.<sup>9</sup>

Non-iid observations in relational data result from the fact that there is a complex structure underlying the dyadic events or processes that we observe. Accounting for this structure is

<sup>6</sup>The LFM also shares similarities with work in the econometric literature on interactive fixed effects (Bai, 2009; Pang, 2014). In this stream of work, interactive fixed effects are used to deal with cross-sectional dependence in TSCS data, in such a way that a latent factor for time can be used to capture common shocks to actors and a latent factor on actors can capture varying responses to those shocks.

<sup>7</sup>The dimensions of  $\mathbf{U}$  and  $\mathbf{V}$  are  $n \times K$  and  $\mathbf{D}$  is a  $K \times K$  diagonal matrix.

<sup>8</sup>Unlike traditional SVD, in the LFM the singular values are not restricted to be positive, thus allowing us to account for both positive and negative homophily.

<sup>9</sup>Details on the sampling procedure and the full conditional distributions can be found in: Hoff (2005, 2008, 2021) and Minhas *et al.* (2019).

necessary if we are to adequately represent the data generating process. If one can specify each of the nodal, dyadic, and triadic attributes that influence interactions then the conditional independence assumption underlying standard approaches will be satisfied. However, it is rarely the case that this is possible even for TSCS data and thus modeling decisions must account for underlying structure. Failing to do so in either TSCS or dyadic data leads to a number of well-known challenges: (a) biased estimates of the effect of independent variables, (b) uncalibrated confidence intervals, and (c) poor predictive performance. Additionally, by ignoring these potential interdependencies, we often ignore substantively interesting features of the phenomena we investigate. The study of international relations is founded on the relations among actors. Why ignore the interdependencies that led to the study of IR in the first place?

#### 4. Simulation study

We explore the utility of AME as an inferential tool for dyadic analysis via a simulation exercise.<sup>10</sup> Most scholars working with dyadic data are primarily concerned with understanding the effect of a particular independent variable on a dyadic-dependent variable. The goal of our simulation is to assess how well AME can provide unbiased and well-calibrated estimates of regression coefficients in the presence of unobserved dependencies, specifically, homophily. As discussed in the previous section, homophily is the idea that actors are more likely to have a tie if they have similar values on a particular variable, and in networks the presence of homophily can lead to third-order dependencies such as transitivity. Homophily can be operationalized by creating a dyadic covariate via the multiplication of a nodal covariate with its transpose. For example, if the nodal covariate was a binary indicator for democracy, multiplying it by its transpose would give us a dyadic covariate that represents whether any dyad is jointly democratic or not.<sup>11</sup>

Assume that the true data-generating process for a binary variable,  $Y$ , is given by:

$$\begin{aligned} Z_{ij} &= \mu + \beta X_{ij} + \gamma W_{ij} + \epsilon_{ij}, \quad \epsilon \sim \text{normal}(0, 1) \\ Y_{ij} &= I(Z_{ij} > 0) \end{aligned} \tag{4}$$

$X_{ij} = x \times x^T$ , where  $x$  is a nodal covariate that is drawn from a standard normal distribution. Similarly,  $W_{ij} = w \times w^T$ , where  $w$  is also a nodal covariate that is independently drawn from a standard normal distribution. We generate our binary dependent variable,  $Y$ , within a probit framework with  $Z$  serving as the latent variable.  $X$  and  $W$  are both dyadic covariates that are a part of the data-generating process for  $Y$ , but  $W$  is not observed. We compare inference for  $\mu$  and  $\beta$ —the latter parameter would be of primary concern for applied scholars—using three models:

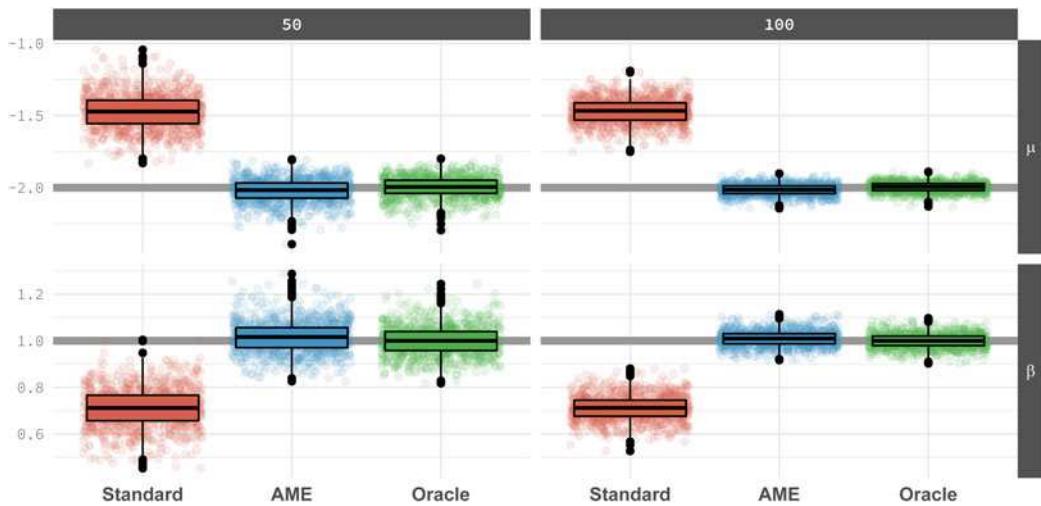
- the “standard” international relations approach estimated through a generalized linear model;<sup>12</sup>
- the AME approach outlined in the previous section with a unidimensional latent factor space ( $K = 1$ );<sup>13</sup>
- and an “oracle” regression model that assumes we have measured all sources of dependencies and thus includes both  $x_{ij}$  and  $w_{ij}$ .

<sup>10</sup>Alternative network-based approaches for dyadic data are exponential random graph models and the related stochastic actor-oriented model. While both these models have led to numerous contributions to a variety of literatures, the applicability of these approaches may be limited to certain types of networks and individual-level characteristics. Specifically, Block *et al.* (2017) note that these types of models may not be appropriate in situations where network and behavioral data depend on unobserved latent variables, which is explicitly the focus of our analysis here.

<sup>11</sup>This process of operationalizing homophily is equivalent to the “nodematch” function in the “ergm” and “latentnet” packages. There are many other options of operationalizing homophily including calculating the difference in scores that a pair of actors may have on a particular nodal variable.

<sup>12</sup>Specifically, here we are just regressing  $Y$  on  $X$  assuming independent errors.

<sup>13</sup>Results with higher values of  $K$  are similar.



**Figure 3.** Regression parameter estimates for the standard, AME, and oracle models from 1000 simulations. Summary statistics are presented through a traditional box plot, and the estimates from each simulation are visualized as well as points.

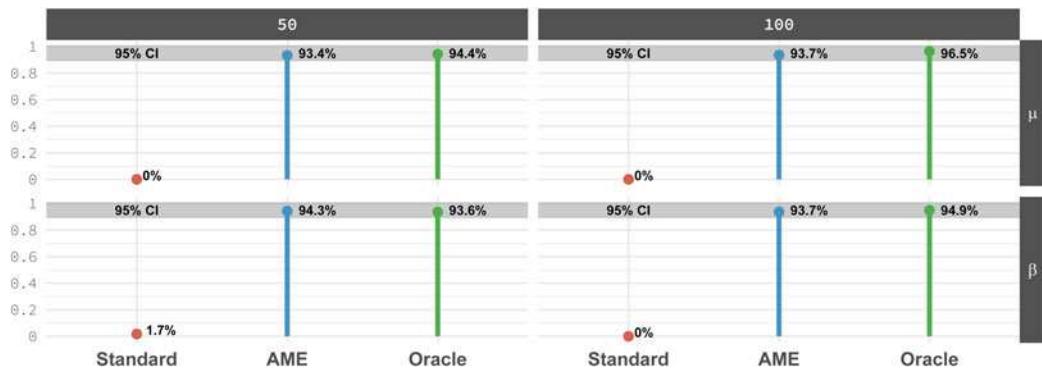
The first model corresponds to the “standard” approach in which little is explicitly done to account for dependencies in dyadic data. In the second model, we use the AME framework described in the previous section. For both the first and second models, we are simply estimating a linear model of  $X$  on  $Y$ , and assessing the extent to which inferences on the regression parameters are complicated by the presence of unobserved dependencies,  $W$ . In the last model, we provide an illustration of the ideal case in which we have observed and measured  $W$  and include it in our specification for  $Y$ . The oracle case provides an important benchmark for the standard and AME approaches.

For the simulation, we set the true value of  $\mu$  (the intercept term) to  $-2$  and  $\beta$  (the effect of  $X$  on  $Y$ ) to  $1$ .<sup>14</sup> We conduct two sets of simulations, one in which the number of actors in the network is set to 50 and the other at 100. In total, we run 1000 simulations where we begin by simulating  $Y$  from the specification given in Equation 4 and then for each simulated  $Y$  we estimate a standard, AME, and oracle model.

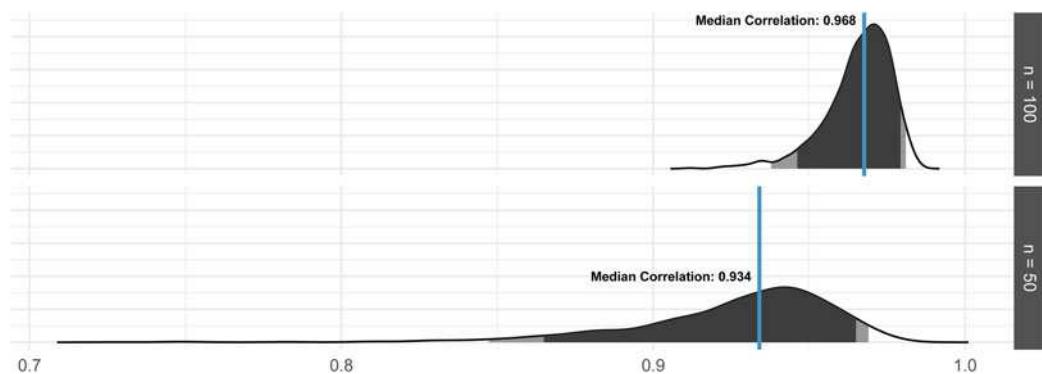
We first compare the performance of the models in terms of how well they estimate the true values of  $\mu$  and  $\beta$  in Figure 3. The panels on the left show the results for when the number of actors is set to 50 and on the right for 100 actors. The top pair of panels represents the estimates for  $\mu$  while the bottom pair do the same for  $\beta$ . In each case, we find that the estimates for  $\mu$  and  $\beta$  produced by the standard approach are notably off from their true values. On the other hand, the AME model performs just as well as the oracle at estimating the true parameter values.

Next, we estimate the 95 percent credible interval for the three models in each of the simulations and estimate the proportion of times that the true value fell within those intervals. The results are summarized in Figure 4, and again we see that the AME approach performs as well as the oracle, while the standard approach performs poorly by comparison. The implication of the results presented in Figures 3 and 4 is that standard approaches can often fail at estimating parameter values and conducting inferential tasks in the presence of unobserved dependencies. The AME approach by comparison can be used as a tool for scholars working with dyadic data to still estimate the true effects of their main variables of interest, while accounting for dependencies that do often emerge in dyadic data.

<sup>14</sup>The value of  $\gamma$  is also set to 1, which corresponds to an example where the  $W$  character is associated with homophily.



**Figure 4.** Proportion of times the true value fell within the estimated 95 percent confidence interval for the standard, AME, and oracle models from 1000 simulations.



**Figure 5.** Distribution of correlation between missing variable and multiplicative random effect in AME across the 1000 simulations. Vertical line through the distribution represents the median value across the simulations.

Moreover, the AME approach allows scholars to better understand what parameters their model may be missing. In the case of the simulation here,  $W$  is set as an unobserved dyadic covariate that has a homophilous effect on  $Y$ . The effect of  $W$  is homophilous within this framework because it is a dyadic attribute involving  $i$  and  $j$  that positively affects the degree to which actors interact with one another, i.e.,  $y_{ij}$ . This type of unobserved dependency will be captured through the multiplicative effects portion of the model,  $U^\top DV$ . To estimate how well the model performs, we recover the multiplicative effects term for each simulation and calculate the correlation between it and the unobserved dependency,  $W$ .<sup>15</sup> We visualize the distribution of the correlations from each of the 1000 simulations in Figure 5 for the case where the number of actors is set to 100 (top pair of panels) and 50. Additionally, we calculate the median across the correlations and display the result using a vertical line. For both  $n = 50$  and  $n = 100$ , we find that the multiplicative effects perform very well in capturing the unobserved dependency, which indicates that the AME does not simply capture noise but also works as a tool to estimate unobserved structure.

The simulation shows that beyond obtaining less biased and better-calibrated parameter estimates, a key benefit of the AME framework is to directly estimate unobserved dependencies through the random effects structure of the model. Scholars can use this framework in an iterative

<sup>15</sup> Specifically, since both the multiplicative effects term and  $W$  are continuous dyadic variables, we calculate the Pearson correlation coefficient.

**Table 1.** Descriptive information about the replicated studies

	Model	Date range	N. actors	Dyads type	Clustering $\sigma_{\beta}$
Reiter and Stam (2003)	Logit	1945–1995	193	Directed	Robust
Weeks (2012)	Logit	1946–1999	197	Directed	Robust
Gibler (2017)	Logit	1816–2008	193	Undirected	Robust

fashion: beginning with an estimated model, they can then empirically study the structure of the random effects to assess whether there are unobserved covariates that they want to include in their model. Importantly, this simulation underscores how a careful consideration of a systems' interconnectedness, both through theoretical approaches and empirical models, can result in more precise estimates of direct and indirect effects across the system.

## 5. Applications with AME

We apply AME to three recent IR studies: Reiter and Stam (2003), Weeks (2012), and Gibler (2017). Each of these studies use relational data of state interactions and propose both dyadic, monadic, and structural explanations for the behavior of actors in the system. We demonstrate the capabilities of AME with reference to existing studies in order to highlight several features of our network-based approach. First, we show that simply by using the AME framework scholars can better model the data generating process behind events of interest. Second, the results of AME estimation are interpretable alongside results using standard approaches. Third, through using this approach, we can also quantify the degree to which first-, second-, and third-order dependencies are present.

We obtain the data for these studies from their replication archives.<sup>16</sup> The chosen studies have each gained over 100 or more citations and were recently published.<sup>17</sup> Each of these articles are well-known and posit a hypothesis in which interdependencies are consequential. Reflecting the dominant approach in the literature, the studies test their hypothesis by employing some form of a GLM. Table 1 provides descriptive information for the studies that we replicate.

For each of the studies listed above, we replicate the authors' key model using their original estimation procedure, a GLM.<sup>18</sup> Next, utilizing the same covariate specifications, we estimate the models with the AME framework.<sup>19</sup> For Reiter and Stam (2003) and Weeks (2012), we use a version of AME that accounts for the directed nature of the data and for Gibler (2017) we used an undirected version. In the directed formulation, separate random effects are used for senders and receivers in both the additive and multiplicative portions of the model.

A key claim that we have made is that by accounting for dependencies inherent to relational data, we can better capture the data generating process behind events of interest. To assess whether or not the AME approach successfully does this, we turn to an out-of-sample cross-validation strategy. An out-of-sample approach is essential since relying on in-sample procedures would enable models with more parameters, such as AME, to simply overfit the data. The cross-validation procedure is executed as follows. For each study, we randomly divide the data into  $k = 30$  sets, letting  $s_{ij,t}$  be the set to which pair  $ij, t$  is assigned. Then for each  $s \in \{1, \dots, k\}$ , we estimate model parameters with  $\{y_{ij,t} : s_{ij,t} \neq s\}$ , the data not in set  $s$ , and predict  $\{\hat{y}_{ij,t} : s_{ij,t} = s\}$  from these estimated parameters.

<sup>16</sup>Without exception, the data were easy to retrieve thanks to the authors' transparency and an increasing norm in the social sciences of open data sharing.

<sup>17</sup>We selected our cases based on a few criteria. We focus on studies that are explicitly about International Relations, were published after the year 2000, and were published in a top ranking general political science outlet (for consistency in editorial standards and reviews, we focus on one journal *American Political Science Review*). We hope that these three criteria ensure that our paper is readable and interpretable to an applied audience.

<sup>18</sup>Replicating the key models from each study was straightforward because of the authors' assembled replication scripts.

<sup>19</sup>In estimating AME, we show results when setting  $K = 2$ . Results with alternative values of  $K$  are similar. We recommend that authors try out a range of  $K$ s.

The result of this procedure is a set of sociomatrices  $\hat{Y}$ , in which each entry  $\hat{y}_{ij,t}$  is a predicted value obtained from a subset of the data that does not include  $y_{ij,t}$ . Next we conduct a series of analyses to discern whether or not the AME model provides any benefit for each study. These analyses are summarized in [Figure 6](#). The left-most plot in each of the panels evaluates performance using Receiver Operating Characteristic (ROC) curves. Models that have a better fit according to this metric will follow the upper-left corner border of the ROC space. In addition to ROC curves, we also use separation plots (Greenhill *et al.*, 2011). Separation plots visualize each of the observations in the dataset according to their predicted probability. In this graph, the shaded panels correspond to the occurrence of an event. Darker panels are events, lighter panels are non-events. If a model performs well, then the events that actually occur would stack to the far right of the graph, where the predicted probabilities generated by the model are highest. The right-most plot in each of the panels evaluates performance using precision-recall (PR) curves. 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 our applications here, as they each are trying to model conflict within dyads, which is an infrequent occurrence.

For each of the replications, we find that the AME approach substantially outperforms the original models in terms of out-of-sample predictive performance. This indicates that switching to the AME framework—even when using the same covariate specification as the original studies—enables scholars to better represent the data generating process of their events of interest. The fact that this analysis is done in an out-of-sample context ensures that the AME framework is not simply overfitting with more parameters. Instead, it suggests that the AME are capturing underlying structure previously missed by the exogenous covariates in the models.

Ignoring this underlying structure has consequences for inferential analysis.<sup>20</sup> The fact that there is such a divergence in the performance of AME versus the original estimation procedures highlights that there are unobserved sources of bias in each of these studies. We hone in on the main finding of each study to draw into focus the potential consequences for ignoring these sources of bias and the inferential benefits of the AME estimation procedure. In [Table 2](#), we present the overall results; the term *Unconfirmed* indicates that the statistical significance of the crucial finding in the original study is not found to hold in the AME estimation.<sup>21</sup>

An important takeaway here is that many scholars are forced to make knowledge claims based on the statistical significance of a small set of covariates or the differences between these covariates. These differences may change dramatically once interdependencies are taken into account. This outcome follows from AME's ability to better account for the dependencies discussed in the previous section, whereas GLM approaches explicitly assume observational independence conditional on the specified covariates. As this is a widely-known limitation of GLM approaches, scholars often attempt to account for clustering of observations by including additional variables and adjusting the standard errors of the resulting estimates. At best, this method introduces noise and imprecision into results and at worst can produce misleading outcomes.

Next we discuss each of the replications in more detail and highlight the substantive insights drawn from the AME framework.

### 5.1. Reiter and Stam (2003)

Reiter and Stam (2003) examine the relationship between democracy, dictatorship, and the initiation of militarized disputes. Their work contests prior scholarship claiming that interstate dyads containing democracies and personalist dictatorships were particularly prone to conflict

<sup>20</sup>As shown in the simulation section, inference in a dyadic setting can become complicated when there are unobserved dependencies.

<sup>21</sup>Full tabular results for each of the original and re-estimated models are presented in the online Supplementary Appendix.

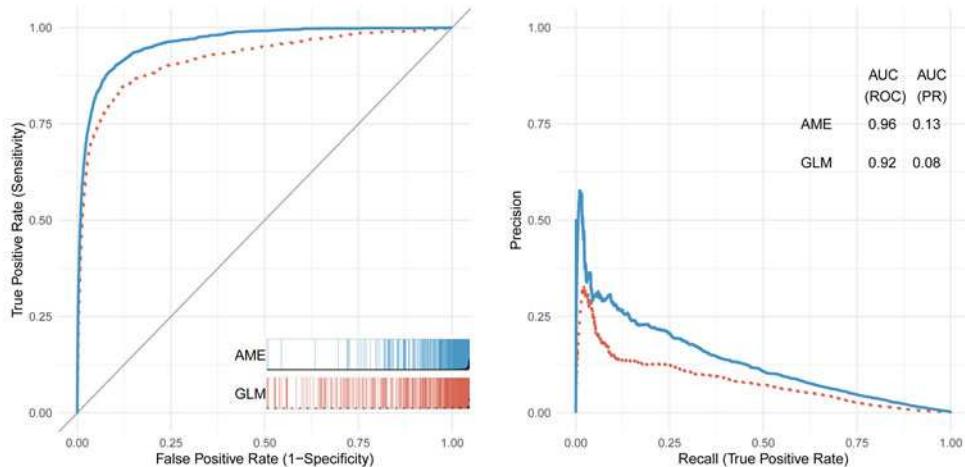
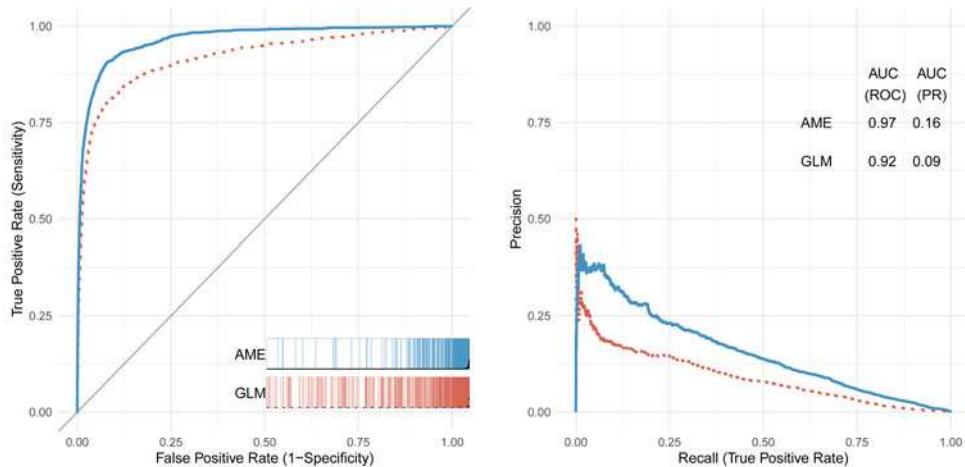
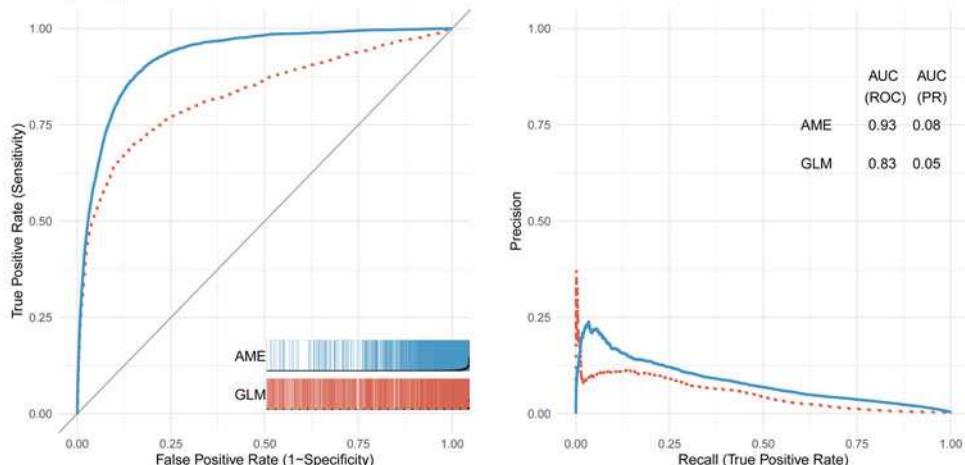
**Reiter & Stam (2003)****Weeks (2012)****Gibler (2017)**

Figure 6. Assessments of out-of-sample predictive performance using ROC curves, PR curves, and separation plots.

**Table 2.** Here we provide a brief summary of the key variable in each of the replications and a note about whether or not the highlighted finding remains when using our network-based approach

Study	Central finding	Confirmed after accounting for dependencies?
Reiter and Stam (2003) Weeks (2012)	Personalist regimes attack democracies, not vice versa Bosses, juntas, and strongmen are more aggressive, machines are not	Confirmed Unconfirmed
Gibler (2017)	Entry to the international system, not power parity, determines conflict	Partially confirmed

because of aggression on the part of the democratic state (Peceny *et al.*, 2002). Using a directed dyadic dataset of almost a million observations, they find evidence against this hypothesis: dictators are in fact more likely to challenge democracies, but not the other way around. In addition, military regimes and single-party regimes are more prone to initiate disputes with democracies, but the opposite is not true.<sup>22</sup> As is prevalent in this literature, Reiter and Stam employ a logistic regression that includes an indicator of the time since the last dispute as well as three cubic splines. Based on their statistical analysis, they conclude that institutional constraints affect the propensity of democratic and non-democratic leaders to engage in military conflict.

The key variables in the original model measure whether or not the sender in the directed dyad is a personalist regime and the target a democratic regime (“Pers/Democ directed dyad”) or whether the opposite is true (“Democ/Pers directed dyad”). The authors find that coefficient of the Pers/Democ directed dyad indicator is positive, while the Democ/Pers directed dyad is too imprecisely measured to indicate a direction. In our re-estimation using the AME framework, we confirm these results, indicating that dictators are likely to initiate or engage in conflict with democratic regimes but not vice versa.

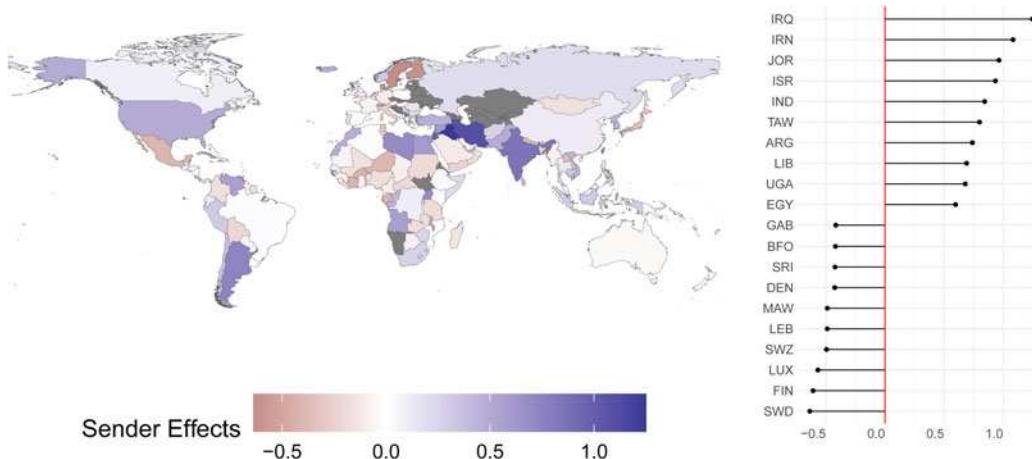
Even though we are able to confirm the original results, employing the AME model offers clear benefits in this case. As already shown in Figure 6, our approach performs notably better in reflecting the data generating process. The reason for this is that there is still underlying structure within this conflict system that the Reiter and Stam model does not fully capture. To highlight this we visualize the estimated sender random effects ( $a_i$ ) from the SRRM portion of the AME framework in Figure 7.

The visualization of the sender random effects highlights that the behavior of countries does not fully accord with predictions from the covariates specified by Reiter and Stam. Specifically, we can see that countries such as Iraq, Israel, and Iran are more likely to be involved in initiating or continuing conflict with other countries than the model would predict. Further, other countries such as Sweden, Finland, and Swaziland are less likely to engage in conflict than the exogenous covariates in the model would suggest. In this case, the finding that countries in the middle east experience more conflict with other countries might lead one to more carefully examine the effects of geography on conflict initiation or to account for Colgan (2010)’s theory that revolutionary petrostates are more aggressive. None of these findings change the key conclusion from Reiter and Stam’s work, but by using the AME framework we are able to better understand the limitations of their model.

### 5.2. Weeks (2012)

Weeks examines the influence of domestic institutions on the initiation of military conflicts by autocratic leaders. She argues that in some circumstances autocrats are held accountable for their foreign policy decisions, and that this is dependent on the audiences of autocrats. When

<sup>22</sup>Independent variables focus on various encodings of regime types, contiguity, alliance, and capability measures. For a full tabular display of the results see online Supplementary Appendix.



**Figure 7.** Estimates of sender random effects ( $a_i$ ) from AME for the Reiter and Stam (2003) model. Positive values indicate that the particular country is more likely to be involved in conflict than predicted by the covariates in the model. Negative values indicate that the country is less likely to be involved in conflict. The left panel shows  $a_i$  estimates for all countries and the right panel highlights the top ten countries in terms of positive values of  $a_i$ .

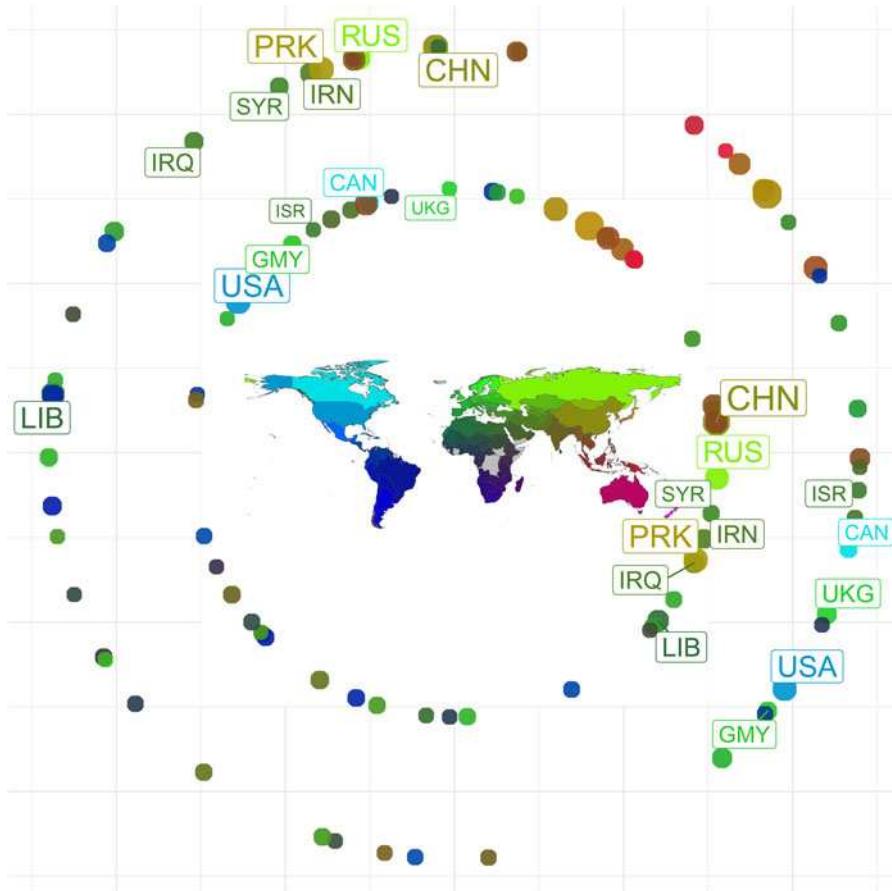
the autocratic regime is nonmilitary, domestic audiences do not favor military actions, but in military autocracies this is not the case. Further, she argues that in personalistic regimes without a military or civilian domestic audience, the leaders are more likely to employ military force in their foreign policy. To study this question, Weeks employs a directed dyad design of conflict similar to that used by Reiter and Stam.

The major innovation in her study resides in the nuanced way she conceptualizes and codes regimes into four types: (a) Machine, (b) Junta, (c) Boss, and (d) Strongmen.<sup>23</sup> She uses a logistic regression, following Beck *et al.* (1998) and includes splines to capture temporal covariation in the dependent variable along with dyad clustered standard errors. The key findings of her work are that (a) juntas, boss, and strongmen regimes are more likely to initiate conflict than machine-type regimes and (b) machine-type regimes are no more belligerent than democracies.<sup>24</sup> In the empirical analysis, Weeks finds that machines are less prone to initiate conflict than the reference category, whereas juntas, bosses, and strongmen are more conflict-prone. When analyzing the results using AME, however, we find that the parameters on each of her autocratic regime type variables are too imprecisely measured to draw any inference about their putative causal effects. Consequentially, none of the findings from her original analysis are confirmed once known dependencies among the data are taken into account via AME.

There is also a striking difference between Weeks' original model and our estimation using AME in terms of capturing the data generating process. As with Reiter and Stam, the divergence is a result of the GLM framework's inability to account for underlying structure generating the event of interest. To uncover this structure, we illustrate another beneficial aspect of the AME framework, which is the multiplicative random effects estimated through the LFM portion of the model:  $\mathbf{u}_i^\top \mathbf{D} \mathbf{v}_j$ . These random effects account for higher order dependencies that manifest as a result of homophily and stochastic equivalence in dyadic data.

<sup>23</sup>Weeks also includes a variety of control variables focusing on capabilities for both sides of the dyad, alliances, geography, trade dependence, regime instability, and the regime type of "side B." For a full tabular display of the results see online Supplementary Appendix.

<sup>24</sup>These insights are mainly emphasized in the paper by the parameter estimates depicted in Tables 1 and 2 (pages 339–340) from the paper.



**Figure 8.** Visualization of multiplicative effects for Weeks (2012). Each circle designates a country and the color corresponds to the legend at the center of the visualization. Countries that cluster together in the outer ring are those that were found by the model to have similar sending patterns, meaning that they tend to send conflict to similar sets of countries. The inner ring clusters countries by the similarity of who they receive conflict from.

To visualize the multiplicative effects, we display the circular diagram shown in Figure 8. The nodes throughout the diagram represent countries and are colored according to their geographic position—a legend is provided in the center. The outer ring visualizes higher order dependence patterns through countries’ sender relationships—country positions here are estimated in the  $u_i$  random effects described in Equation 2. Countries that are more proximate to each other in this outer ring are more likely to send or initiate conflicts with similar targets. The inner ring, based on the estimates of  $v_i$ , is constructed such that countries closer together are more likely to receive conflict from the same sender countries. Last, the distance between countries in the inner and outer rings proportionately reflects who a country is more likely to be in conflict with.

Figure 8 reveals a number of notable clusters. For example, in the bottom right corner we see the US, UK, Germany, Canada, and Israel. These states cluster together in the outer ring of this visualization because they tend to send conflicts to similar targets. Conversely, in the top left of the outer ring, we see a cluster of authoritarian countries: Iraq, Russia, Syria, North Korea, and China. We observe similar clusters in the inner ring. Specifically, we again see the US, UK, Germany, Canada, and Israel clustering together indicating that they are more likely to receive conflict from the same countries. The cluster of democratic and authoritarian countries facing

each other in the inner and outer rings indicate that they are more likely to engage in conflict with one another.

Perhaps most critically, an evaluation of this visualization highlights Weeks' original expectation of how states behave in the conflict system. Specifically, Iraq, Syria, Libya, and North Korea all fell under Weeks' "boss" category, and each of these states tends to cluster together in the inner and outer rings. This indicates that even though we do not find support for Weeks' assertion that certain authoritarian regime types are more likely to initiate conflict, we do find that these regimes are more likely to behave similarly in terms of who they target and receive conflict from.

### 5.3. *Gibler (2017)*

The last replication we conduct with the AME model considers a study by Gibler (2017). Gibler argues that the long-standing relationship between the relative parity of capabilities and initiation of militarized interstate disputes (MIDs) is almost completely mediated by the initial conditions for the members of the dyad when they joined the international system as sovereign members. In most specifications, after taking into account the initial conditions for dyadic entry, the statistical significance of power parity vanishes. This finding calls into question many IR theories about the role of balance in generating international conflict (Organski, 1958).

To test this hypothesis, Gibler employs an undirected dyadic design and estimates his model using a GLM with dyad-clustered standard errors.<sup>25</sup> With this design, Gibler finds support for both the insignificance of contemporary power parity, and the effect of the initial conditions for entry on driving conflict behavior. When we re-estimate using AME power parity still has a small, imprecisely measured effect, supporting Gibler's argument that contemporary power parity is an artifact of other aspects of state relations, and does not drive contemporary conflict behavior. At the same time, the AME approach finds that the effect of power parity when a country enters into the international system is too imprecisely measured to draw any inferences about its effect on the initiation of MIDs. Not only is our estimate of the effect of this variable small, but it has a very large relative standard error—over a magnitude larger than the parameter itself. Thus, while we can confirm the argument that the effect of power parity vanishes when accounting for both initial conditions and network dependencies, we find less support for the argument that initial conditions drive the initiation of MIDs.

In the two previous replications, we showed how to parse apart underlying structure using the SRM and LFM portions of the AME framework. Here we turn our focus to parameter interpretation in a substantive context. Specifically, even when the GLM and AME frameworks produce results that may seem to be in accordance with one another, the substantive interpretations of the effects of covariates can differ notably since the AME model uses a set of random effects to account for unobserved factors. To explore this, we focus on the effect of rivalry on MID initiation. Both the GLM and AME estimations find that rivalry has a positive effect on MID initiation, but the expected effects between the two models differ greatly.

To clarify the difference, we turn to a simulation-based approach. We employ mean or modal values for all independent variables, except we change the rivalry variable to indicate that there was a rivalry when the actual data suggest there is none. This provides us with two scenarios, one in which rivalry is set to one and the other zero, while in both scenarios all other parameters are set to their measure of central tendency. The expected values of this scenario are essentially a first difference plot comparing results with the model when estimated in two different ways: the GLM estimation and our AME approach.

As Figure 9 illustrates, the substantive results differ notably. The expected value of the dependent variable—the probability of the onset of a MID, is considerably lower once interdependencies are taken into account with the AME model. These are rare events, so the probabilities are low, but the difference in predictions is notable.

<sup>25</sup>Full tabular results are shown in the online Supplementary Appendix, we focus on model 6 from Table 6 (2017, 34).

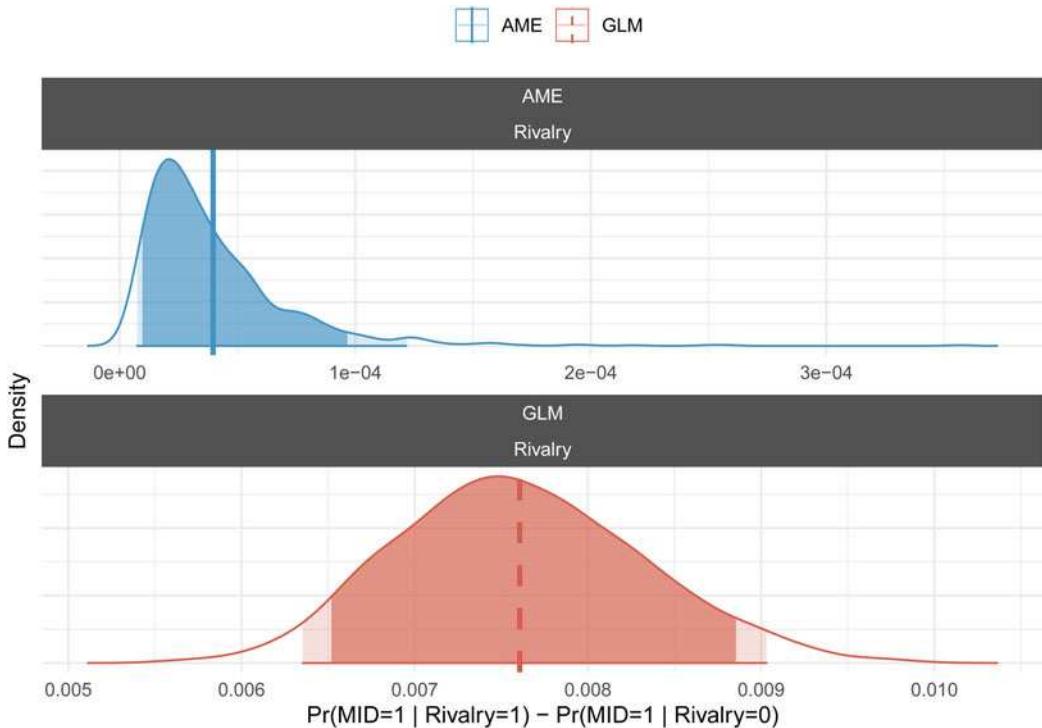


Figure 9. Marginal effects of a change in the Rivalry variable for both the AME and the GLM estimation.

#### 5.4. Lessons learned

First, utilizing the AME framework enables scholars to better model the underlying structure inherent to dyadic data. In each of the models, the AME substantially outperforms the original model out of sample. Not only does AME perform better at correctly identifying cases in which the dependent variable takes a value of 0 (via the ROC curves), but it also dominates at correctly identifying occurrences of the dependent variable in the data (seen via the PR curves). This may be because by ignoring dependencies, the original models are misspecifying the DGP, and the AME better accounts for it.

Second, in both the simulation and replication sections, we have shown that results can change notably when interdependencies are not taken into account. Not only are coefficients biased in the GLM approaches, but they are often imprecisely measured with poorly calibrated standard errors. This means that significance testing (for better or worse) is compromised when network effects are ignored.

Third, even when the results from the AME estimation conform with those found in an OLS or logistic regression, new insights can emerge from the structure uncovered by the random effects of the AME framework. In particular, the AME provides information about the dependencies so that clusters can be identified. This information can then be used to generate new hypotheses.

Fourth, it is evident that the actual results—not the estimated coefficients and their covariances—generated by the models differ greatly in expectations. This implies that policy experimentations with the models, as well as scenario-based simulations and forecasting of GLM models are likely to give misleading results compared to the AME approach.

## 6. Conclusion

International relations is about the interactions, relationships, and dependencies among countries or other important international actors. This is particularly true of scholarship in the tradition of

the Correlates of War Project, but it is by no means limited to it.<sup>26</sup> Many scholars have debated the use and abuse of dyadic data.<sup>27</sup> A broad survey of the IR literature makes it clear that scholars find dyadic data to be an essential touchstone in the study of international relations (Erikson *et al.*, 2014; Aronow *et al.*, 2015). Our findings bolster a growing recognition in the field of International Relations that interdependence influences not only statistical estimations, but how scholars theorize about internationally relevant politics. Scholars have demonstrated the theoretical importance of interdependence through research on intrastate conflict (Dorff *et al.*, 2020), interstate bargaining (Gallop, 2017), economic interdependence (Maoz, 2009), and international treaties (Kinne, 2013) among other topics.

At the same time, we know that research designs focusing on the statistical analysis of dyadic data quickly go astray if the dyadic data are assumed to be iid. Virtually all of the standard statistical models—ordinary least squares and logistic regressions, to name a few—fail if the data are not conditionally independent. This fact has been accepted as it relates to temporal dependencies, but adoption of methods to account for network dependencies have seen less progress. By definition dyadic data are not iid and thus the standard approaches cannot be used cavalierly to analyze these data. Signorino (1999) showed why this is true of models of strategic interaction, but it is more broadly true of models that employ dyadic data. We show that the AME framework can be employed to account for the statistical issues that arise when studying dyadic data.

To explore this approach in the context of international relations we have presented two analyses. The first is a simulation where the characteristics of the network are known. This shows that when there are unobserved dependencies, the AME approach is less biased in terms of parameter estimation compared to the standard approach employed in international relations to study dyadic data (i.e., GLM models). The second analysis is a replication of recent studies that use a broad range of dyadic data to draw inferences about international relations. These studies have been replicated with the original research designs, each of which used a statistical method that assumes the dyadic data are all independent from one another. We then re-analyzed each study using the AME model. In every case, we found that the AME approach provided (a) increased precision of estimation, (b) better out-of-sample fit, and (c) evidence of 1st-, 2nd-, and 3rd-order dependencies that were overlooked in the original studies.<sup>28</sup>

**Supplementary material.** Online appendix material for this article can be found at <https://doi.org/10.1017/psrm.2021.56>.

**Funding.** Cassy Dorff acknowledges support from National Science Foundation (NSF) Award 2017162 and Shahryar Minhas acknowledges support from NSF Award 2017180.

## References

Adamic LA and Glance N (2005) The political blogosphere and the 2004 US election: divided they blog. *Proceedings of the 3rd International Workshop on Link Discovery*. New York, NY: ACM, for ACM, pp. 36–43.

Anderson CJ, Wasserman S and Faust K (1992) Building stochastic blockmodels. *Social Networks* **14**, 137–161.

Aronow PM, Samii C and Assenova VA (2015) Cluster-robust variance estimation for dyadic data. *Political Analysis* **23**, 564–577.

Bai J (2009) Panel data models with interactive fixed effects. *Econometrica* **77**, 1229–1279.

Barabási A-L and Réka A (1999) Emergence of scaling in random networks. *Science* **286**, 509–510.

Beck N and Katz JN (1995) What to do (and not to do) with pooled time-series cross-section data. *American Political Science Review* **89**, 634–647.

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

Bennett J and Lanning S (2007) The Netflix prize. *Proceedings of KDD Cup and Workshop*.

<sup>26</sup>See Singer (1972) for an early description of the project and also see the project's website for an history and more recent efforts <http://www.correlatesofwar.org/>.

<sup>27</sup>One recent on-line symposium can be found at <http://bit.ly/2wB2hab>.

<sup>28</sup>The online Supplementary Appendix contains performance data on all of these replications, as well as sample code illustrating how to undertake AME analysis using amen.

**Block P, Stadtfeld C and Snijders TAB** (2017) Forms of dependence: comparing SAOMs and ERGMs from basic principles. *Sociological Methods & Research* **48**(1), 202–239.

**Chyžh O** (2016) Dangerous liaisons: an endogenous model of international trade and human rights. *Journal of Peace Research* **53**, 409–423.

**Colgan JD** (2010) Oil and revolutionary governments: fuel for international conflict. *International Organization* **64**, 661–694.

**Davis J and Goadrich M** (2006) The relationship between precision-recall and ROC curves. *Proceedings of the 23rd International Conference on Machine Learning, ICML 2006*. New York, NY, USA: ACM, pp. 233–240.

**Dorff C and Minhas S** (2017) When do states say uncle? network dependence and sanction compliance. *International Interactions* **43**, 563–588.

**Dorff C, Gallop M and Minhas S** (2020) Networks of violence: predicting conflict in Nigeria. *The Journal of Politics* **82**, 476–493.

**Erikson RS, Pinto PM and Rader KT** (2014) Dyadic analysis in international relations: a cautionary tale. *Political Analysis* **22**, 457–463.

**Gallop M** (2017) More dangerous than dyads: how a third party enables rationalist explanations for war. *Journal of Theoretical Politics* **29**, 353–381.

**Gibler DM** (2017) State development, parity, and international conflict. *American Political Science Review* **111**, 21–38.

**Greenhill B, Ward MD and Sacks A** (2011) The separation plot: a new visual method for evaluating the fit of binary data. *American Journal of Political Science* **55**, 991–1002.

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

**Hoff PD** (2008) Modeling homophily and stochastic equivalence in symmetric relational data. In Platt JC, Koller D, Singer Y and Roweis ST (eds), *Advances in Neural Information Processing Systems 20, Processing Systems 21*. Cambridge, MA, USA: MIT Press, pp. 657–664.

**Hoff PD** (2021) Additive and multiplicative effects network models. *Statistical Science* **36**, 34–50.

**Keohane RO** (1989) Reciprocity in international relations. *International Organization* **40**, 1–27.

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

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

**Manger MS, Pickup MA and Snijders TAB** (2012) A hierarchy of preferences: a longitudinal network analysis approach to PTA formation. *Journal of Conflict Resolution* **56**, 852–877.

**Maoz Z** (2009) The effects of strategic and economic interdependence on international conflict across levels of analysis. *American Journal of Political Science* **53**, 223–240.

**Minhas S, Hoff PD and Ward MD** (2017) Influence networks in international relations. Working Paper.

**Minhas S, Hoff PD and Ward MD** (2019) Inferential approaches for network analysis: AMEN for latent factor models. *Political Analysis* **27**, 208–222.

**Mucha PJ, Richardson T, Macon K, Porter MA and Onnela JP** (2010) Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876ff.

**Organski AFK** (1958) *World Politics: The Stages of Political Development*. New York: Alfred A. Knopf.

**Pang X** (2014) Varying responses to common shocks and complex cross-sectional dependence: dynamic multilevel modeling with multifactor error structures for time-series cross-sectional data. *Political Analysis* **22**, 464–496.

**Pawitan Y** (2013) *In All Likelihood: Statistical Modeling and Inference Using Likelihood*, 1st Edn. Oxford, England: Oxford University Press.

**Peceny M, Beer CC and Sanchez-Terry S** (2002) Identifying the culprit: democracy, dictatorship, and dispute initiation. *American Political Science Review* **96**, 15–26.

**Reiter D and Stam AC** (2003) Identifying the culprit: democracy, dictatorship, and dispute initiation. *American Political Science Review* **97**, 333–337.

**Resnick P and Varian HR** (1997) Recommender systems. *Communications of the ACM* **40**, 56–58.

**Richardson LF** (1960) *Arms and Insecurity*. Chicago and Pittsburgh, PA: Quadrangle/Boxwood.

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

**Singer JD** (1972) The correlates of war project: continuity, diversity, and convergence. In Singer JD (ed), *Quantitative International Politics: An Appraisal. Special Studies in International Politics and Government*, Vol. VI. Praeger.

**Wasserman S and Faust K** (1994) *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

**Weeks JL** (2012) Strongmen and straw men: authoritarian regimes and the initiation of international conflict. *American Political Science Review* **106**, 326–347.

**Zinnes DA** (1967) An analytical study of the balance of power theories. *Journal of Peace Research* **3**, 270–288.

Copyright of Political Science Research & Methods is the property of Cambridge University Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.