Modeling Asymmetric Relationships from Symmetric Networks

Arturas Rozenas

NYU

ar199@nyu.edu

Shahryar Minhas

MSU

minhassh@msu.edu

John Ahlquist

UCSD

jahlquist@ucsd.edu [*]

June 22, 2018

Many bilateral relationships requiring mutual agreement produce observable networks that are symmetric (undirected). However, the unobserved, asymmetric (directed) network is frequently the object of scientific interest. We propose a method that probabilistically reconstructs the latent, asymmetric network from the observed, symmetric graph in a regression-based framework. We apply this model to the bilateral investment treaty network. Our approach successfully recovers the true data generating process in simulation studies, extracts new, politically relevant information about the network structure inaccessible to alternative approaches, and has superior predictive performance.

1. Introduction

Social actors are often embedded in webs of relationships that profoundly shape political and economic outcomes (Franzese and Hays, 2008; Ward, Stovel and Sacks, 2011). One challenge in analyzing networks arises in situations where an analyst cannot fully observe the nature of relational ties. In many dyadic interactions — treaties, marriages — outsiders can observe ties only if both agents agree, that is, the payoff for forming a tie exceeds its cost for both members of a dyad (Jackson and Wolinsky, 1996). The observed network is therefore composed of symmetric ("undirected") ties even though the social process at work contains important relational asymmetries. The pursuit of a tie by one party may not be reciprocated to the same extent by another.

As an illustration, suppose A, B, and C are three warring factions deciding whether to sign bilateral peace agreements. We observe a network in which dyads (A, C) and (B, C) have signed agreements, but the dyad (A, B) continues fighting. This observed network of 'peaceful' ties could be generated from any of three unobserved sets of relations: it may be that B failed to reciprocate A's pursuit of peace, or vice versa, or neither A nor B pursued peace. To identify conditions that drive factions to sign peace agreements we must account for these unobserved asymmetries. Here, the observed symmetric graph is an incomplete representation of the underlying, asymmetric network, which is frequently the object of scientific interest. We refer to this situation as "partial observability."

We present the partial observability generalized bilinear mixed effects model (P-GBME) to address this challenge. The model is a synthesis of the generalized bilinear mixed effects (GBME) model (Hoff, 2005) and the bivariate or "partial observability" probit model (Poirier, 1980; Przeworski and Vreeland, 2002). The model can probabilistically reconstruct the directed network from which the observed, undirected graph emerged. The model enables the study of network ties in a regression framework by accounting for interdependencies as well as unobserved asymmetries in network relations. The stochastic actor-oriented model (SAOM) for networks (Snijders and Pickup, 2017) also allows for partial observability. However, SAOM was designed to assess how specific network features (e.g., k-star triangles) give rise to an observed network. The latent network approach is not used to study the role of specific network statistics. Rather, latent network models aim to account for broad patterns of network interdependence using a variance decomposition regression framework.[1]

We illustrate the P-GBME model by applying it to the bilateral investment treaties (BIT) network for each year in 1990-2012. The model substantially improves predictive accuracy relative to both conventional logit and standard GBME. As important, P-GBME extracts new information about the factors that drive treaty preferences, identifies important structural changes in the network, and highlights possible "hidden" agreements that are easily overlooked when latent network asymmetries are ignored.

2. The Model

Building on the random utility framework (McFadden, 1980), we model an actor $i = 1, \ldots, N$ as having net utility from forming a tie with another j: $z_{ij} = \mu_{ij} + \epsilon_{ij}$ with μ_{ij} representing the systematic component (that depends on observables) and ϵ_{ij} representing

---

[1]See Minhas, Hoff and Ward (2016) for detailed discussion. Other approaches based on generalized spatio-temporal dependence can also recover directed predicted probabilities Franzese, Hays and Kachi (2012).

the stochastic error. To account for interdependencies in actors' utilities from having ties, we use the "latent space" approach (Hoff, 2005) and model these utilities as follows:

$$
\begin{pmatrix} z_{ij} \\ z_{ji} \end{pmatrix} \sim \mathcal{N} \left(\begin{matrix} \mu_{ij} + a_i + b_j + \boldsymbol{u}_i' \boldsymbol{v}_j \\ \mu_{ji} + a_j + b_i + \boldsymbol{u}_j' \boldsymbol{v}_i \end{matrix} , \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right).
\tag{1}
$$

The correlation, ρ, captures the "reciprocity" between the utilities that actors derive from tie formation. Parameters a_i and b_j are sender- and receiver-specific random effects, respectively, and they capture second-order network dependencies. The vectors $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ represent the location of actor i in the latent space of 'senders' and 'receivers,' respectively. These random effects capture higher-order dependencies in network ties: i derives a large utility from forming a tie with j, if i's location in the latent space of 'senders' $\boldsymbol{u}_i$ is close to j's location in the latent space of 'receivers' $\boldsymbol{u}_j$ (so that the cross-product $\boldsymbol{u}_i' \boldsymbol{v}_j$ is large).[2] We express the systematic components of actors' utilities as linear functions of predictors:

$$
\mu_{ij} = \boldsymbol{\beta}^{(s)} \boldsymbol{x}_i^{(s)} + \boldsymbol{\beta}^{(r)} \boldsymbol{x}_j^{(r)} + \boldsymbol{\beta}^{(d)} \boldsymbol{x}_{ij}^{(d)},
\tag{2}
$$

$$
\mu_{ji} = \boldsymbol{\beta}^{(s)} \boldsymbol{x}_j^{(s)} + \boldsymbol{\beta}^{(r)} \boldsymbol{x}_i^{(r)} + \boldsymbol{\beta}^{(d)} \boldsymbol{x}_{ji}^{(d)}.
\tag{3}
$$

A researcher cannot directly observe net utility (the z's). We only observe agents' behaviors, in this case undirected bilateral ties. A directional tie $i \rightarrow j$ is formed if and only if i's net gain from doing so is strictly positive, $z_{ij} > 0$. Accordingly, the bilateral tie $i \leftrightarrow j$ is formed if and only if both actors derive a net positive payoff from having a tie so that $z_{ij} > 0$ and $z_{ji} > 0$. A researcher observes an undirected (bilateral) tie $y_{ij} = y_{ji} = \{0, 1\}$ arising from the following data generating process:

$$
y_{ij} = y_{ji} = \begin{cases} 1 & \text{if } z_{ij} > 0 \text{ and } z_{ji} > 0, \\ 0 & \text{else.} \end{cases}
\tag{4}
$$

Under a standard identifying restriction $\sigma^2 = 1$, the model is a partially observable probit regression (Poirier, 1980), augmented with random-effects to capture unobserved heterogeneity and inter-dyadic dependencies. Vectors $\boldsymbol{x}_i^{(s)}$ and $\boldsymbol{x}_i^{(r)}$ represent the sender-specific and receiver-specific covariates, respectively. A model for the directional link $i \rightarrow j$ (eq. 11), uses variables $\boldsymbol{x}_i$ as sender-specific predictors, but these same predictors become receiver-specific in the model for the directional link $j \rightarrow i$ (eq. 12). Vector $\boldsymbol{x}_{ij}^{(d)}$ contains dyad-specific variables. These dyad-specific variables might be symmetric, $x_{ij} = x_{ji}$ (e.g., distance between countries), or not (e.g., export-import).

A partially observed probit model requires at least one of the following identifying restrictions: (1) regression equations 11 and 12 must have the same parameters and/or (2) one equation contains a predictor not included in another equation (Poirier, 1980). If the dyadic predictors are asymmetric, $\boldsymbol{x}_{ij} \neq \boldsymbol{x}_{ji}$, then condition (2) is satisfied. Furthermore, regression equations 11 and 12 have the same parameters, and so condition (1) holds as well; thus, the above model is parametrically identified. However, we impose an additional restriction that $\rho = 0$. While this restriction is not required for parametric identification, Rajbhandari (2014) showed that finite sample estimates of ρ are sensitive to the starting

---

[2]As in previous literature, these random effects are modeled as $(a_i, b_i) \sim \mathcal{N}(0, \boldsymbol{\Sigma_{ab}})$, $\boldsymbol{u}_i \sim \mathcal{N}_K(\boldsymbol{0}, \sigma_u^2 \boldsymbol{I})$, and $\boldsymbol{v}_i \sim \mathcal{N}_K(\boldsymbol{0}, \sigma_v^2 \boldsymbol{I})$, where $\boldsymbol{\Sigma_{ab}}$, σ_u^2, and σ_v^2 are unknown parameters. The choice of K, dimensionality of the latent space, is discussed supplementary materials.

values and generally cannot be treated as reliable.[3]

We estimate the model in a Bayesian framework using Markov chain Monte Carlo. In the supplemental materials give a more detailed exposition of the model, prior assumptions, the sampling algorithm. We also provide results from a simulation study demonstrating that the model successfully recovers known parameter values.

3. Application: Bilateral Investment Treaties

We apply the P-GBME model to the network of bilateral investment treaties (BITs) from 1990 to 2012 using the standard United Nations BIT database. There is a vibrant debate on whether BITs boost FDI (Jandhyala, Henisz and Mansfield, 2011; Simmons, 2014; Minhas, 2016), but a proper resolution of this debate requires a convincing empirical model of treaty formation (Rosendorff and Shin, 2012). Partial observability is one key challenge in building such model: the observed network of signed bilateral treaties is symmetric, while the underlying preferences for these treaties are asymmetric.

We fit the P-GBME model separately for each year of data using a suite of covariates that closely follows the existing empirical literature (see supplementary materials). Our model improves on the previous literature by accounting for both network interdependence and partial observability.

3.1. Predictive Performance

In Table 1 we compare the in-sample and out-of-sample predictive performance of the P-GBME to that of pooled probit, which assumes dyadic independence and ignores partial observability, and GBME, which models dyadic interdependencies, but not partial observability. The predictive accuracy of the GBME model in this case is similar to the pooled probit. Adding the partial observability component to the GBME model, however, produces an additional substantial improvement in the predictive accuracy as shown by all metrics for the P-GBME model.

	In-sample		Out-of-sample	
	ROC	PR	ROC	PR
Pooled probit	0.75	0.48	0.73	0.44
GBME	0.76	0.47	0.77	0.48
P-GBME	0.90	0.71	0.91	0.78

Table 1: Predictive performance in BIT data: area under the receiver operating characteristic curve (ROC) and area under the precision recall curve (PR).

3.2. Regression Parameters

Existing models, including GBME, estimate a single coefficient for each predictor. This assumes away the possibility that the same factor differentially "affects" i's demand for a treaty with j and i's attractiveness to j. The P-GBME recovers directed sender- and receiver-effects for node-level covariates. For instance, our estimates suggest that,

---

[3]The estimation algorithm provided with this paper allows ρ to be estimated, but caution should be used when utilizing this option.

countries faster growth in GDP per capita were no more inclined to sign BITs with others (sender effect). But high-growth countries were more attractive BIT partners to others (receiver effect). Supplementary materials describes regression parameter estimates in detail.

3.3. The Structure of Latent Treaty Preferences

The P-GBME model allows us to extract "latent preferences" for treaty formation – the estimated probability that country i demands a treaty from j, and vice versa. Figure 1 displays the mean posterior predicted probabilities relevant to China in 1995 and 2010. The horizontal axis represents a country's attractiveness to China as a BIT partner and the vertical axis is China's attractiveness to that country. Countries above the diagonal line find China a more attractive BIT partner than China finds them, and vice versa. Color identifies observed BITs.

The plots reveal how China's position in the BIT network changed over time. In 1995, China was moving aggressively to demand BITs around the world, forming ties that the model views as relatively unlikely. By 2010, China had many more BITs in place and is more likely to be a treaty target by the remaining countries, an indication of China's expanded role in the global economy.
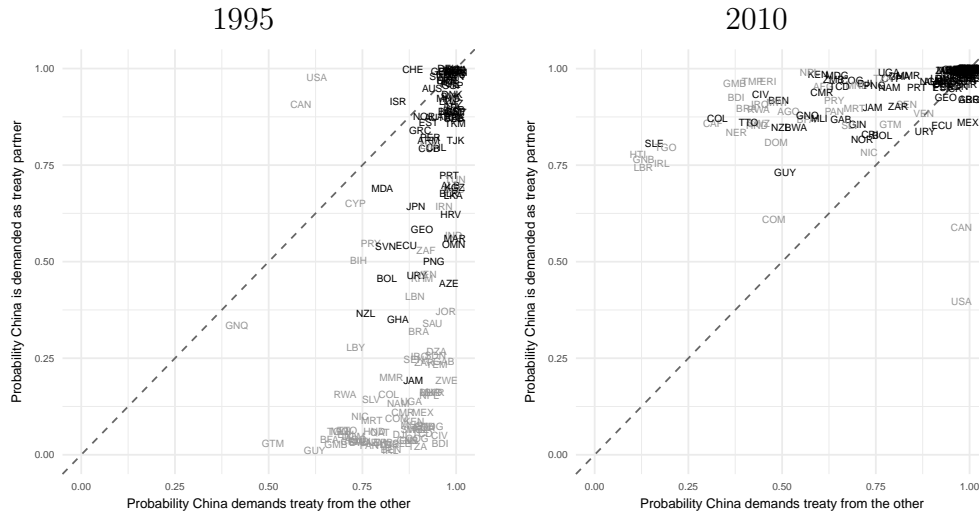

Figure 1: Each panels displays the posterior mean probability that China demands and will be demanded as a BIT partner. Country labels in black designate those that had formed a BIT with China by the specified year.

Figure 2 illustrates a different use of the P-GBME model. It shows the predicted probabilities that the USA is demanded (left) and demands (right) a BIT in 2010. Observe that the model predicts Peru, Mexico, and Chile—the top 3 non-BITs—demand a BIT with the USA at probabilities close to one, and are also likely BIT targets of the USA with probabilities exceeding 0.5. Closer inspection of UN treaty data reveals that, by 2010, all three had signed other agreements with the US that contain provisions functionally equivalent to BITs; these agreements do not appear in the BIT dataset commonly used in the literature.[4] The P-GBME model nevertheless highlights these "hidden" agreements as

---

[4]The treaties are 2006 Peru-USA Free Trade Agreement (FTA), NAFTA in 1992, and the 2003 Chile-USA FTA. These agreements included investment provisions that mirror the terms of a BIT almost exactly (see `http://investmentpolicyhub.unctad.org/Download/TreatyFile/5454`).

dyads likely to have a BIT. This suggests that researchers studying BITs need to carefully examine the dataset they employ and perhaps expand the set of treaties considered relevant.
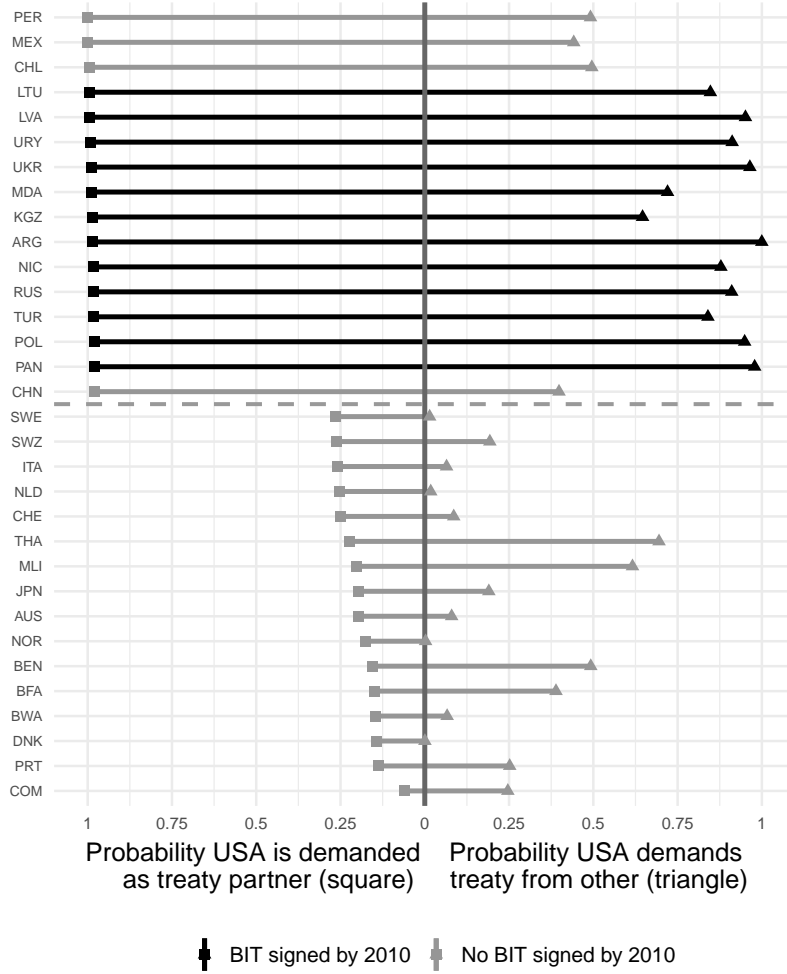

Figure 2: The posterior probability a country demands a BIT (left) and is demanded (right) with the USA in 2010 (the top and bottom decile of countries). P-GBME is good at separating likely from unlikely treaties as well as identifying "hidden agreements."

4. Conclusion

Partial observability occurs whenever the observed graph is undirected yet the underlying process implies directed relationships. We introduced a model that can reconstruct the latent directed network ties, and illustrated its advantages on an example of bilateral trade agreements. The future work in this area could focus on several extensions. First, we accounted for network dependencies using bilinear mixed effects framework (GBME), which could be generalized using the recently developed additive and multiplicative effects network model (Hoff, 2015; Minhas, Hoff and Ward, 2016). Second, estimating this type of network model in a fully dynamic setting (as opposed to slicing data by time, as we did here) remains a challenge, especially when the set of nodes changes in time.

Supplemental Materials

A. Details on the P-GBME

As detailed in the paper, the partial observability generalized bilinear mixed effects (P-GBME) framework treats the observed symmetric outcome, $y_{ij} = y_{ji}$, as resulting from a joint decision taken by a pair of actors. We formalize the joint decision making process using a bivariate probit model with a standard normal link function:

$$y_{ij} = y_{ji} = \begin{cases} 1 & \text{if } z_{ij} > 0 \text{ and } z_{ji} > 0, \\ 0 & \text{else,} \end{cases} \tag{5}$$

$$\begin{pmatrix} z_{ij} \\ z_{ji} \end{pmatrix} \sim \mathcal{N}\left(\begin{matrix} \mu_{ij} + a_i + b_j + \boldsymbol{u}_i'\boldsymbol{v}_j \\ \mu_{ji} + a_j + b_i + \boldsymbol{u}_j'\boldsymbol{v}_i \end{matrix}, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right), \tag{6}$$

$$(a_i, b_i)' \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\Sigma}_{ab}\right), \tag{7}$$

$$\boldsymbol{u}_i \sim \mathcal{N}_K(\boldsymbol{0}, \sigma_u^2 \boldsymbol{I}), \tag{8}$$

$$\boldsymbol{v}_i \sim \mathcal{N}_K(\boldsymbol{0}, \sigma_v^2 \boldsymbol{I}). \tag{9}$$

a_i and b_j represent sender and receiver random effects that account for first order dependence patterns that often arise in relational data, while $\boldsymbol{u}_i'\boldsymbol{v}_j$ captures the likelihood of a pair of actors interacting with one another based on third order dependence patterns such as transitivity, balance, and clustering. For identification purposes, we fix $\sigma^2 = 1$ and $\rho = 0$. The former is a standard restriction in probit frameworks with a binary outcome. We undertake the latter restriction because Rajbhandari (2014) shows that in this framework it is difficult to recover reliable estimates for ρ as the parameter is highly sensitive to the initial value.

The sender and receiver random effects (a_i and b_j) are drawn from a multivariate normal distribution centered at zero with a covariance matrix, Σ_{ab}, parameterized as follows:

$$\Sigma_{ab} = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix} \tag{10}$$

The nodal effects are modeled in this way to account for the fact that in many relational datasets we often find that actors who send a lot of ties are also more likely to receive a lot of ties. Heterogeneity in the the sender and receiver effects is captured by σ_a^2 and σ_b^2, respectively, and σ_{ab} describes the covariance between these two effects.

μ represents the systematic component of actors' utilities and is expressed as a linear function of sender $^{(s)}$, receiver $^{(r)}$, and dyadic $^{(d)}$ covariates:

$$\mu_{ij} = \boldsymbol{\beta}^{(s)}\boldsymbol{x}_i^{(s)} + \boldsymbol{\beta}^{(r)}\boldsymbol{x}_j^{(r)} + \boldsymbol{\beta}^{(d)}\boldsymbol{x}_{ij}^{(d)}, \tag{11}$$

$$\mu_{ji} = \boldsymbol{\beta}^{(s)}\boldsymbol{x}_j^{(s)} + \boldsymbol{\beta}^{(r)}\boldsymbol{x}_i^{(r)} + \boldsymbol{\beta}^{(d)}\boldsymbol{x}_{ji}^{(d)}. \tag{12}$$

This formulation allows us to incorporate exogenous actor and dyad level characteristics into how actors make decisions within the partial probit framework. Following Hoff (2005), to enable a more efficient estimation, we reparameterize the model to implement

hierarchical centering of the random effects (Gelfand, Sahu and Carlin, 1995):

$$z_{i,j} \approx \mathcal{N}(\boldsymbol{\beta}^{(d)}\boldsymbol{x}_{ij}^{(d)} + s_i + r_j + \boldsymbol{u}_i'\boldsymbol{v}_j), \tag{13}$$

$$s_i = \boldsymbol{\beta}^{(s)}\boldsymbol{x}_i^{(s)} + a_i, \tag{14}$$

$$r_j = \boldsymbol{\beta}^{(s)}\boldsymbol{x}_j^{(r)} + b_j. \tag{15}$$

A.1. Parameters and Priors

To estimate the parameters discussed in the previous section, we utilize conjugate priors and a Monte Carlo Markov Chain (MCMC) algorithm. Prior distributions for the parameters are specified as follows:[5]

- $\boldsymbol{\beta}^{(s)}$, $\boldsymbol{\beta}^{(r)}$, and $\boldsymbol{\beta}^{(d)}$ are each drawn from multivariate normals with mean zero and a covariance matrix in which the covariances are set to zero and variances to 10

- $\Sigma_{a,b} \sim$ inverse Wishart$(I_{2\times2}, 4)$

- σ_u^2, and σ_v^2 are each drawn from an i.i.d. inverse gamma(1,1).

Starting values for each of the parameters are determined using maximum likelihood estimation.

A.2. The MCMC algorithm

To estimate this model a Gibbs sampler is used. This sampler follows the procedure laid out in Hoff (2005, 2009) with the exception of the first step in which we extend the GBME by accounting for the possibility that seemingly symmetric events are the result of a joint decision between a pair of actors. This first step involves sampling from a truncated normal distribution, we show the full conditional distribution below.

1. Modeling partially observable outcome. Conditional on there being an observed link between i and j, and conditional on other parameters, we draw the latent variables z_{ij} and z_{ji} from the bivariate normal distribution such that both latent variables are positive:

$$\begin{pmatrix} z_{ij} \\ z_{ji} \end{pmatrix} \Bigg| y_{ij} = 1 \sim \mathcal{N}\left(\begin{matrix} \mu_{ij} + a_i + b_j + \boldsymbol{u}_i'\boldsymbol{v}_j \\ \mu_{ji} + a_j + b_i + \boldsymbol{u}_j'\boldsymbol{v}_i \end{matrix}, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right) \mathbb{1}\{z_{ij} > 0 \cap z_{ji} > 0\}.$$

Conditional $y_{ij} = 0$ (there is no observed link between i and j), we sample the latent variables from the bivariate normal distribution where at least one of the latent variables, z_{ij} or z_{ji}, is constrained to be negative:

$$\begin{pmatrix} z_{ij} \\ z_{ji} \end{pmatrix} \Bigg| y_{ij} = 0 \sim \mathcal{N}\left(\begin{matrix} \mu_{ij} + a_i + b_j + \boldsymbol{u}_i'\boldsymbol{v}_j \\ \mu_{ji} + a_j + b_i + \boldsymbol{u}_j'\boldsymbol{v}_i \end{matrix}, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right) \mathbb{1}\{z_{ij} < 0 \cup z_{ji} < 0\}.$$

2. Additive effects

   - Sample $\boldsymbol{\beta}^{(d)}, s, r \mid \boldsymbol{\beta}^{(s)}, \boldsymbol{\beta}^{(r)}, \Sigma_{a,b}, Z, \boldsymbol{U}, \boldsymbol{V}$ (linear regression)

---

[5]For details on the full conditional distributions of each of the parameters see Hoff (2005).

- Sample $\boldsymbol{\beta}^{(s)}, \boldsymbol{\beta}^{(r)} \mid s, r, \Sigma_{a,b}$ (linear regression)
- Sample $\Sigma_{a,b}$ from full conditional distribution

3. Multiplicative effects[6]

- For $i = 1, \dots, n$:
    - Sample $u_i \mid \{u_j, j \neq i\}, Z, \boldsymbol{\beta}^{(d)}, s, r, \sigma_u^2, \sigma_v^2, \boldsymbol{V}$ (linear regression)
    - Sample $v_i \mid \{v_j, j \neq i\}, Z, \boldsymbol{\beta}^{(d)}, s, r, \sigma_u^2, \sigma_v^2, \boldsymbol{U}$ (linear regression)

A.3.   Simulation Exercise

To test the capabilities of the P-GBME framework in representing the data generating process for a partially observable outcome we conduct a simulation exercise. In each simulation, we randomly construct a directed network from a pair of dyadic covariates, nodal covariates, and the random effects structure detailed in the previous section. The regression parameters for the dyadic covariates are set at 1 and -1/2, and the parameters for the nodal covariates are set at 0 and 1/2. At this stage, the network simulated from this data generating process is directed. We modify the simulated network so that a link between a dyad only appears in the network if both the i, j dyad and the j, i dyad both have a link in the simulated network, thus making the network appear undirected.

Next, we examine whether the P-GBME model can recover the data generating process underlying the partially observed simulated network. We run the P-GBME model in every simulation for 20,000 iterations with a 10,000 burn-in period. We repeat this simulation process 100 times.

With the simulation results our first step is to examine whether the P-GBME accurately recovers the regression parameter estimates for the dyadic and nodal covariates. To test whether this is the case we calculate the mean regression parameter estimate from the MCMC results for each simulation, and we summarize these results in Figure 1. For each parameter we indicate its true value by a colored horizontal line and summarize the distribution of the mean regression values estimated from the P-GBME using a boxplot. Given that for each of the parameters the true value almost exactly crosses the median value indicated in the box plot, this simulation shows that the P-GBME is quite effective in estimating the true parameter values underlying a partially observed outcome.

---

[6]See Hoff (2009) for further details on how multiplicative effects are estimated in a directed context within the GBME framework.
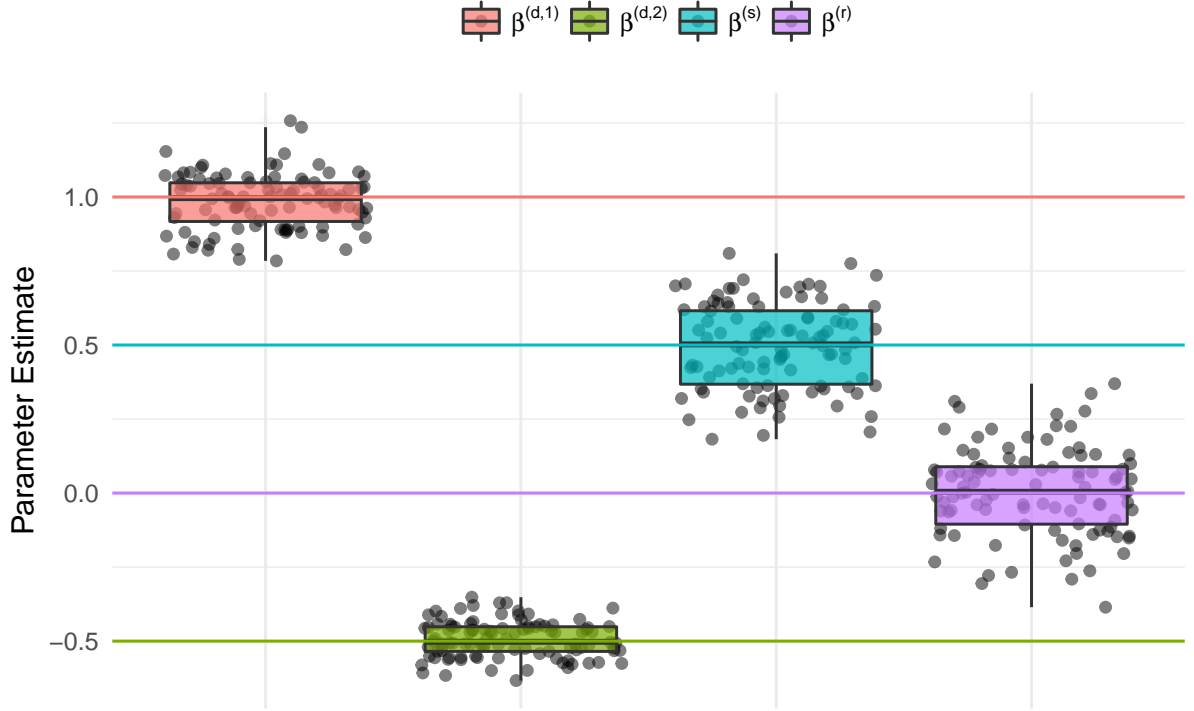
Figure 1: Boxplot of mean value of regression parameters estimated using the P-GBME across 100 simulations. Horizontal colored lines indicate the true parameter values.

We also examine the proportion of times that the true value falls within the 95% credible interval of the estimated regression parameter. In over ninety percent of the simulations, the true value falls within the 95% credible interval of each of the estimated regression parameters from the P-GBME. Specifically, for $\beta^{(d,1)}$ the coverage rate is 0.85, for $\beta^{(d,2)}$ 0.93, for $\beta^{(s)}$ 0.93, and for $\beta^{(r)}$ 0.97.

B. Estimation and Application

B.1. Data

Table B1 provides a description for each of the variables used in the analysis.

Table B1: Variables in the analysis

Variable	Level	Definition	Source
BIT	dyadic	1 if i & j Signed a BIT by year t	UNCTAD[7]
UDS (median)	dyadic	$|\text{UDS}_{it} - \text{UDS}_{jt}|$	Pemstein et al. (2010)
Law & Order	dyadic	$|\text{LO}_{it} - \text{LO}_{jt}|$	ICRG[8]
Log(GDP per capita)	dyadic	$|\text{Log(GDPcap)}_{it} - \text{Log(GDPcap)}_{jt}|$	WDI
OECD	dyadic	1 if i & j Both OECD members by year t	OECD
Distance	dyadic	Minimum distance between i & j	Gleditsch & Ward (2001) [9]
FDI/GDP	node	Net FDI inflow as % GDP in year t	WDI
ICSID Disputes	node	Cumulative number of disputes by year t	ICSID[10]
GDP per capita growth	node	Level of GDP per capita growth by year t	WDI
PTAs	node	Cumulative number of PTAs signed by year t	DESTA[11]

A shortcoming of the existing GBME framework is its inability to account for applications where there is missingness in the set of exogenous covariates used in the model. For our application, a number of the nodal covariates had varying levels of missingness. Additionally, most of the dyadic covariates that we construct from nodal variables, such as the unified democracy scores, also have varying levels of missingness. The table below shows how much missingness we had for the variables included in our analysis:

Table B2: Missingness among variables used in the analysis

Variable	Proportion of Cases Missing
Law & Order	16.6%
FDI/GDP	2.8%
GDP per capita	1.4%
GDP per capita growth	1.4%
Unified Democracy Scores (UDS)	0.7%
ICSID Disputes	0%
PTAs	0%
OECD	0%

In general, the level of missingness is not high. The only exception here is with the Law & Order variable from the ICRG dataset, for this variable we had approximately 17% of country-year observations missing from 1990 to 2012. The only true dyadic variable we include in our analysis is a calculation of the minimum distance between countries, and this variable has no missingness. Additionally, our dependent variable measuring whether or not two countries had signed a BIT by year t also has no missingness.

A number of works have noted the issues that can arise when simply using list-wise deletion,[12] thus before running the P-GBME sampler we impute missingness among the covariates used in our model with a Bayesian, semi-parametric copula imputation scheme.[13] We generate 1,000 imputed datasets from this imputation scheme and save 10 for use in the P-GBME MCMC sampler.

To account for missingness within the P-GBME, at the beginning of every iteration of

---

[12]See, for example, King et al. (2001).

[13]See Hoff (2007); Hollenbach et al. (2016) for details on this imputation scheme and how it differs from other approaches frequently utilized in political science.

the MCMC for model, we draw a randomly sampled imputed dataset from the posterior of the Copula, calculate the parameters associated with the P-GBME using the imputed dataset, and repeat this process for every iteration of the sampler for the model. This approach directly incorporates imputation uncertainty into our posterior distributions of the P-GBME parameters without having to run and combine separate models.

B.2. Estimation details

In our application we estimate the P-GBME separately for each year from 1990 to 2012 using the prior distributions and MCMC algorithm described above. For each year, we ran the P-GBME MCMC sampler for 20,000 iterations, discarding the first 10,000 iterations as burn-in. We thinned the chain by saving only every 10th value.

The following trace plot describes MCMC convergence for all parameters in the 2012 P-GBME model.
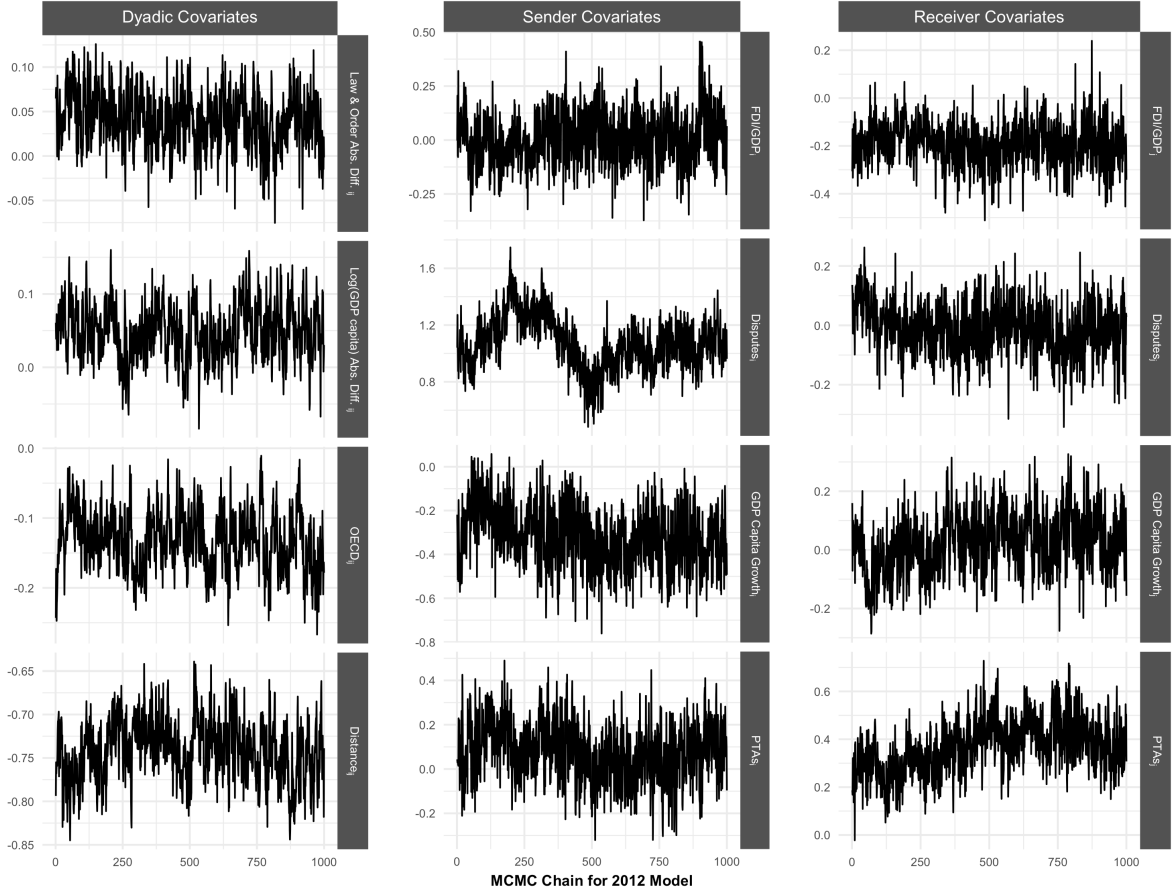

Figure 2: Traceplot for 2012 P-GBME results.

B.3. Regression Parameters

Figure 3 displays the posterior mean and the 90% (thicker) and 95% (thinner) credible intervals (CIs). The first column contains the dyadic covariates; the second, sender-level covariates; and, the third, receiver covariates. In each panel, we show the parameter estimates for that variable from 1990 to 2012. The dotted horizontal line is 0 and the thicker grey line is the posterior mean, pooling the posterior draws across all years.

12

The P-GBME recovers directed sender- and receiver-effects for node-level covariates from an observed undirected network, something that other approaches, including the GBME, are unable to do. Our estimation in this application indicates substantial instability in these estimates over time, both relative to a baseline of 0 and relative to the pooled posterior mean. This instability is consistent with substantive arguments that the incentives to sign BITs have changed over time (Jandhyala, Henisz and Mansfield, 2011).
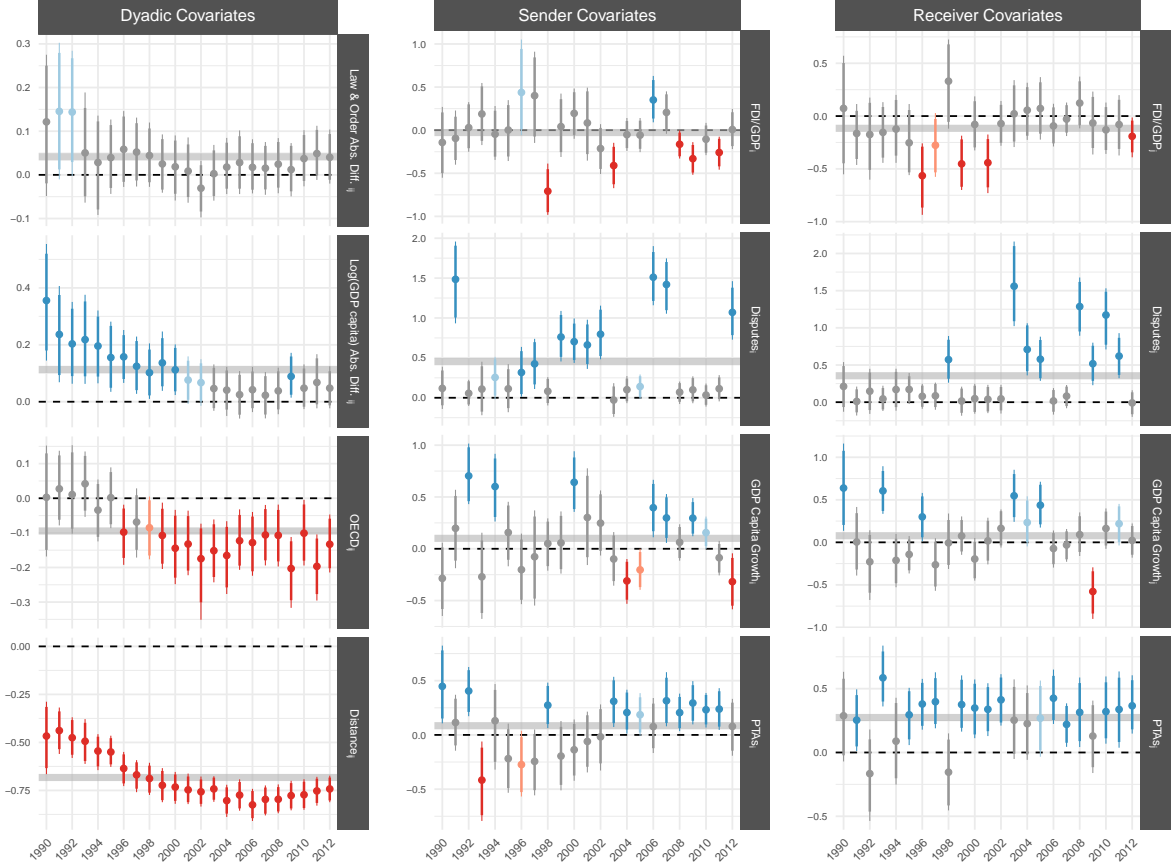

Figure 3: Left-most plot shows results by year for the dyadic parameters, next shows parameter results for sender covariates, and right-most plot show results for receiver covariates. Points in each of the plots represents the average effect for the parameter and the width the 90 and 95% credible intervals. The grey bar in the panels represents the average effect of the parameter across all years. Dark shades of blue and red indicate that the 95% CI does not contain 0 and lighter shades implies that the 0 is not in the 90% credible interval. Parameters in grey are are ones where 0 is inside the 90% credible interval.

Dyadic covariates tend to be more stable across years in this application. But they, too, show that the BIT formation process has changed over time. Economic and political "distance", for example, has become less important as the network evolved and more lower-income countries have signed BITs with each other. Geographic distance, on the other hand, continues to be strongly related to the formation of BITs.

P-GBME covariate estimates shed light on the evolving processes producing the observed BIT network in the 1990-2012 period. The changing values of covariate parameters over time also indicates that common practice of pooling dyads and assuming the existence of temporally stable parameters may be dangerous.

B.4. Choosing dimension of the multiplicative effects, K

One of the parameters that users are able to set within the P-GBME to account for third order dependence patterns is K – see the MCMC algorithm section above for more details on this parameter and its relation to the model. In the results reported in the paper, we set $K = 2$. To understand whether or not a higher value of K is necessary users of this approach can compare the in-sample fit of the model with varying values for K. In our application exercise, we varied K from 1 to 3 to settle on an appropriate value of K that can represent the data generating process of the network. Results are shown in Table B3 below.

Table B3: In-sample performance results from running the P-GBME on data from 2012 with varying values of K.

	AUC (ROC)	AUC (PR)
$K=1$	0.87	0.63
$K=2$	0.90	0.71
$K=3$	0.92	0.71

As you can see after $K=2$, the subsequent in-sample performance improvement notably declines. There is a slight increase in performance from $K=2$ to $K=3$, however, every time one increases K we are also adding $2*n$ more parameters to the P-GBME model. Adding this many more parameters can easily lead one to overfit the data in an out-of-sample context.

References

Franzese, Robert J. and Jude C. Hays. 2008. "Interdependence in Comparative Politics: Substance, Theory, Empirics, Substance." Comparative Political Studies 41(4/5):742–780.

Franzese, Robert J., Jude C. Hays and Aya Kachi. 2012. "Modeling History Dependence in Network-Behavior Coevolution." Political Analysis 20(2):175–190.

Gelfand, Alan E., Sujit K. Sahu and Bradley P. Carlin. 1995. "Efficient parametrisations for normal linear mixed models." Biometrika 82(3):479–488.

Gleditsch, Kristian S. and Michael D. Ward. 2001. "Measuring Space: A Minimum Distance Database and Applications to International Studies." Journal of Peace Research 38(6):749–768.

Hoff, Peter. 2009. "Multiplicative Latent Factor Models for Description and Prediction of Social Networks." Computational and Mathematical Organization Theory 15(4):261–272.

Hoff, Peter D. 2005. "Bilinear Mixed-Effects Models for Dyadic Data." Journal of the American Statistical Association 100(4690):286–295.

Hoff, Peter D. 2007. "sbgcop: Semiparametric Bayesian Gaussian copula estimation and imputation." R Package.

Hoff, Peter D. 2015. "Dyadic Data Analysis with `amen`." arxiv arxiv:1506.08237:1–48.

Hollenbach, Florian M., Iavor Bojinov, Shahryar Minhas, Nils W. Metternich, Michael D. Ward and Alexander Volfovsky. 2016. "Principled Imputation Made Simple: Multiple Imputation Using Gaussian Copulas." Working Paper .

Jackson, Matthew O. and Asher Wolinsky. 1996. "A strategic model of social and economic networks." Journal of Economic Theory 71(1):44–74.

Jandhyala, Srividya, Witold J. Henisz and Edward D. Mansfield. 2011. "Three waves of BITs: The global diffusion of foreign investment policy." Journal of Conflict Resolution 55(6):1047–1073.

King, Gary, James Honaker, Anne Joseph and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." American Political Science Review 95(1):49–69.

McFadden, Daniel. 1980. "Econometric models for probabilistic choice among products." Journal of Business pp. S13–S29.

Minhas, Shahryar. 2016. Learning from Incredible Commitments: Evolution and Impact of Bilateral Investment Treaties Ph. D. Thesis Duke University Durham, North Carolina: .

Minhas, Shahryar, Peter D. Hoff and Michael D. Ward. 2016. "Let's Say Amen for Latent Factor Models." Working paper.
URL: https://arxiv.org/abs/1611.00460

Pemstein, Daniel, Steven A. Meserve and James Melton. 2010. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." Political Analysis 18(4):426–449.

Poirier, Dale J. 1980. "Partial Observability in Bivariate Probit Models." Journal of Econometrics 12(2):209–0217.

Przeworski, Adam and James Raymond Vreeland. 2002. "A Statistical Model of Bilateral Cooperation." Political Analysis 10(2):101–112.

Rajbhandari, Ashish. 2014. Identification and MCMC Estimation of Bivariate Probit Models with Partial Observability. In Bayesian Inference in the Social Sciences, ed. Ivan Jeliazkov and Xin-She Yang. Wiley.

Rosendorff, B. Peter and Kongjoo Shin. 2012. "Importing transparency: The political economy of BITs and FDI Flows." Manuscript, New York University Political Science Department. New York, NY: NYU. .

Simmons, Beth A. 2014. "Bargaining over BITs, Arbitrating Awards: The Regime for Protection and Promotion of International Investment." World Politics 66(1):12–46.

Snijders, Tom A. B. and Mark Pickup. 2017. Stochastic Actor Oriented Models for Network Dynamics. In Oxford Handbook of Politicial Networks, ed. Jennifer Nicoll Victor, Alexander H. Montgomery and Mark Lubell. Oxford University Press.

Ward, Michael D, Katherine Stovel and Audrey Sacks. 2011. "Network analysis and political science." Annual Review of Political Science 14:245–264.

Weidmann, Nils B. and Kristian Skrede Gleditsch. 2015. *cshapes*: CShapes Dataset and Utilities. $\mathcal{R}$ package version 0.5-1.
URL: https://CRAN.R-project.org/package=cshapes