Broader Research

Me

Predicting Intrastate Conflict: Evidence from Nigeria

Cassy Dorff (University of New Mexico),
Max Gallop (University of Strathclyde),
and Shahryar Minhas (Michigan State University)

October 7, 2017

Motivation

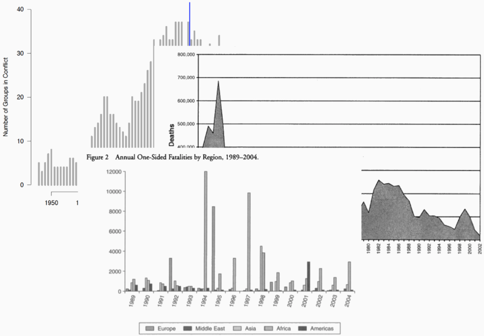
Predicting Intrastate Conflicts

Map of where conflicts have occurred?

Progress in the field

Extensive literature on the causes and prediction of intrastate conflict

Hegre et al. (2001)
Fearon & Laitin (2003)
Collier et al. (2004)
Salehyan (2008)
Cunningham (2013)
Sambanis & Shayo (2013)
Lacina (2014)
Prorok (2016)



Fearon & Laitin (2003) has been cited over 6,000 times!

Conflicts are complex

Roughly a **third** of all intrastate conflict between 1989 and 2003 have been fought with multiple warring parties (UCDP/PRIO 2007).

Conflicts involve multiple actors with changing relationships overtime

- Coordination (Bakke et al 2012; Findley & Rudloff 2012)
- Spoiler groups and veto-players (Cunningham, 2006)

Pairing Empirical Analysis to Theory

"Existence of multiple rebel groups means we can no longer understand civil wars with a sole focus on state attributes. In fact, the government's strategies leading to victory, defeat, or continuation of war can only be understood in relation to the rebel group/groups it is fighting."

Akcinaroglu (2012)

Our Project

Conflict processes are driven by the evolution of relationships overtime.

- We argue that intrastate conflicts are best understood as a single complex system composed of multiple actors in conflict with one another
- We conceptualize armed actors and battles as nodes and ties in a network
- We present a novel model that captures relationships endogenous to the conflict system, as well as covariate information at both the state and actor level.
- Our model out performs standard approaches and uncovers important substantive implications for the study of conflict processes

Networks & Conflict Processes

From dyads to networks

Dyadic data consists of a set of:

- nodes (e.g., rebel group actors)
- measurements specific to a pair of actors (e.g., the occurrence of a battle)

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

$\longrightarrow$

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

Dyadic data assumptions

GLM: $y_{ij} \sim \beta^T X_{ij} + e_{ij}$

Networks typically show evidence against independence of dyadic interactions

Not accounting for dependence can lead to:

- biased effects estimation
- uncalibrated confidence intervals
- poor predictive performance
- inaccurate description of network phenomena

We've been hearing this concern for decades now:

Thompson & Walker (1982)
Frank & Strauss (1986)
Kenny (1996)
Krackhardt (1998)
Beck et al. (1998)
Signorino (1999)
Li & Loken (2002)
Hoa & Ward (2004)
Snijders (2011)
Erikson et al. (2014)
Aronow et al. (2015)
Athey et al. (2016)

Network Effects

How does evolution in the structure of relationships influence conflict over time?

- 1st-order: Sender effects
- 2nd-order: Reciprocity
- 3rd-order: Homophily & Stochastic equivalence
- System level: Changing actor composition

What network phenomena? Sender heterogeneity

Values across a row, say $\{y_{ij}, y_{ik}, y_{il}\}$, may be more similar to each other than other values in the adjacency matrix because each of these values has a common sender i

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

What network phenomena? Receiver heterogeneity

Values across a column, say $\{y_{ji}, y_{ki}, y_{li}\}$, may be more similar to each other than other values in the adjacency matrix because each of these values has a common receiver i

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

What network phenomena? Sender-Receiver Covariance

Actors who are more likely to send ties in a network may also be more likely to receive them

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

What network phenomena? Reciprocity

Values of y_{ij} and y_{ji} may be statistically dependent

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

Social Relations Model (The "A" in AME)

Additive effects portion of AME (Warner et al. 1979; Li & Loken 2002):

$$
\begin{aligned}
y_{ij} &= \mu + e_{ij} \\
e_{ij} &= a_i + b_j + \epsilon_{ij} \\
\{(a_1, b_1), \ldots, (a_n, b_n)\} &\sim N(0, \Sigma_{ab}) \\
\{(\epsilon_{ij}, \epsilon_{ji}) : \; i \neq j\} &\sim N(0, \Sigma_{\epsilon}), \text{ where}
\end{aligned}
$$

$$
\Sigma_{ab} = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix} \qquad \Sigma_{\epsilon} = \sigma_{\epsilon}^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}
$$

- μ baseline measure of network activity (for the purpose of regression we turn this into $\boldsymbol{\beta}^T \mathbf{X}_{ij,t}$)
- e_{ij} residual variation that we will use the SRM to decompose

Social Relations Model (The “A” in AME)

$$
\begin{aligned}
y_{ij} &= \mu + e_{ij} \\
e_{ij} &= a_i + b_j + \epsilon_{ij} \\
\{(a_1, b_1), \ldots, (a_n, b_n)\} &\sim N(0, \Sigma_{ab}) \\
\{(\epsilon_{ij}, \epsilon_{ji}) : \; i \neq j\} &\sim N(0, \Sigma_{\epsilon}), \text{ where}
\end{aligned}
$$

$$
\Sigma_{ab} = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix} \qquad \Sigma_{\epsilon} = \sigma_{\epsilon}^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}
$$

- row/sender effect (a_i) & column/receiver effect (b_j)
- Modeled jointly to account for correlation in how active an actor is in sending and receiving ties

Social Relations Model (The "A" in AME)

$$\begin{aligned} y_{ij} &= \mu + e_{ij} \\ e_{ij} &= a_i + b_j + \epsilon_{ij} \\ \{(a_1, b_1), \ldots, (a_n, b_n)\} &\sim N(0, \Sigma_{ab}) \\ \{(\epsilon_{ij}, \epsilon_{ji}) : \ i \neq j\} &\sim N(0, \Sigma_\epsilon), \text{ where} \end{aligned}$$

$$\Sigma_{ab} = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix} \qquad \Sigma_\epsilon = \sigma_\epsilon^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

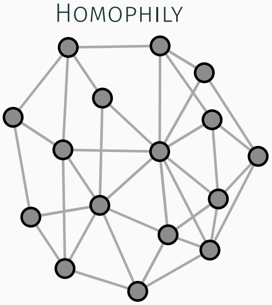
- σ_a^2 and σ_b^2 capture heterogeneity in the row and column means
- σ_{ab} describes the linear relationship between these two effects (i.e., whether actors who send [receive] a lot of ties also receive [send] a lot of ties)

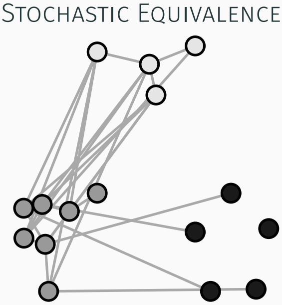
Social Relations Model (The "A" in AME)

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

- ϵ_{ij} captures the within dyad effect
- Second-order dependencies are described by σ_ϵ^2
- Reciprocity, aka within dyad correlation, represented by ρ

Third Order Dependencies





To account for these patterns we can build on the SRM framework and find an expression for γ:

$$y_{ij} \approx \beta^T X_{ij} + a_i + b_j + \gamma(u_i, v_j)$$

Latent Factor Model: The "M" in AME

Each node i has an unknown latent factor

$$\mathbf{u}_i, \mathbf{v}_j \in \mathbb{R}^k \quad i, j \in \{1, \ldots, n\}$$

The probability of a tie from i to j depends on their latent factors

$$\begin{aligned} \gamma(\mathbf{u}_i, \mathbf{v}_j) &= \mathbf{u}_i^T D \mathbf{v}_j \\ &= \sum_{k \in K} d_k u_{ik} v_{jk} \\ &D \text{ is a } K \times K \text{ diagonal matrix} \end{aligned}$$

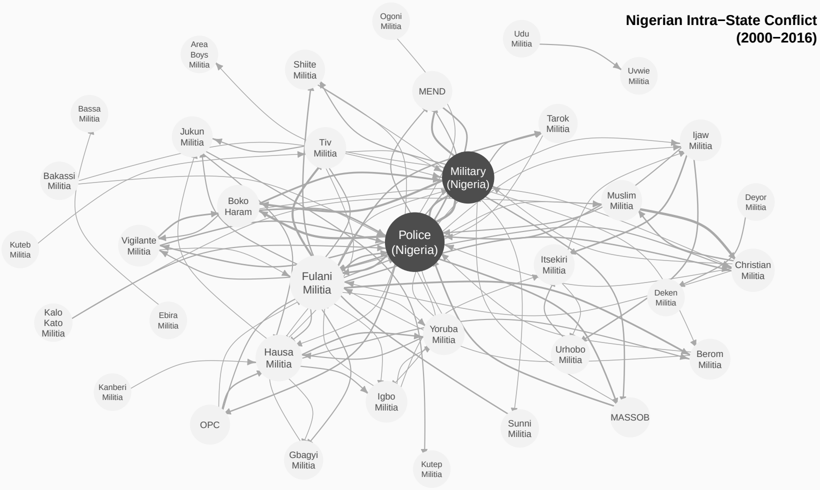
Accounts for both stochastic equivalence and homophily (Hoff 2008)

Additive and Multiplicative Effects (AME) Model

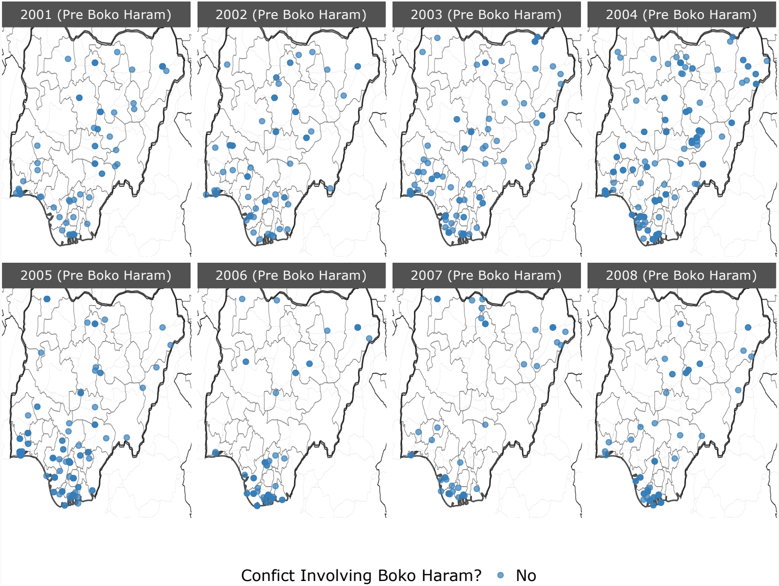
$$
\begin{aligned}
y_{ij,t} &= g(\theta_{ij,t}) \\
\theta_{ij,t} &= \boldsymbol{\beta}^T \mathbf{X}_{ij,t} + e_{ij,t} \\
e_{ij,t} &= a_i + b_j + \epsilon_{ij} + \alpha(\mathbf{u}_i, \mathbf{v}_j) \text{ , where} \\
\alpha(\mathbf{u}_i, \mathbf{v}_j) &= \mathbf{u}_i^T \mathbf{D} \mathbf{v}_j = \sum_{k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

(Hoff 2005; Hoff 2008; Hoff et al. 2013; Minhas et al. 2016)

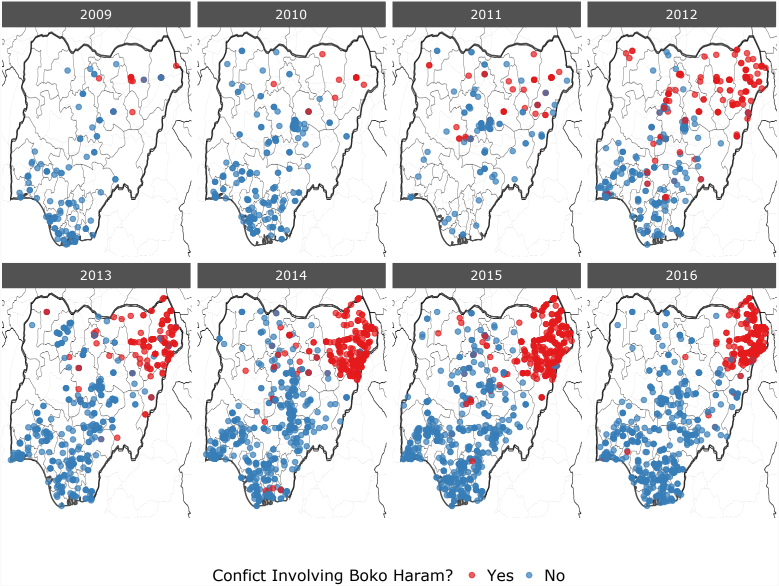
Application case: Nigerian intrastate conflict system



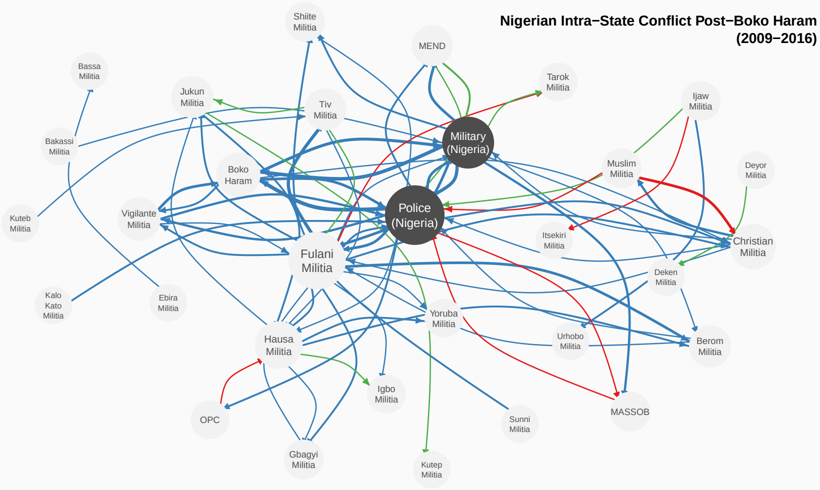
Spatial Distribution of Conflict Pre Boko Haram



Spatial Distribution of Conflict Post Boko Haram



Boko Haram’s Entrance in Network



Data

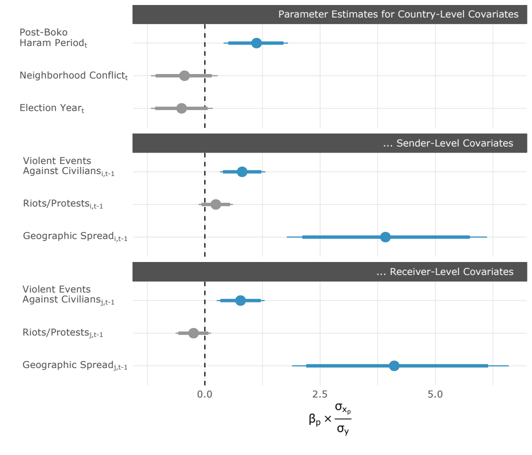
Armed Conflict Location and Event Data Project (ACLED) developed by Raleigh et al. (2010)

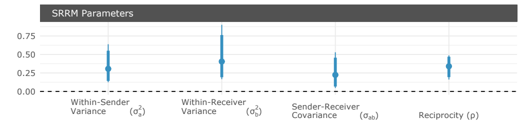
- ACLED records armed conflict and protest events in over 60 developing countries
- We use ACLED *battles* data for Nigeria to generate a measure of conflict where:
    - $y_{ij,t} = 1$ indicates that a conflict occurred when actor i attacked actor j at time t
    - $y_{ij,t} = 0$ if no conflict occurred
- We focus only on modeling the interactions between armed groups that are engaged in battles for at least 5 years during the 2000-2016 period, which results in a total of 37 armed groups

Covariates

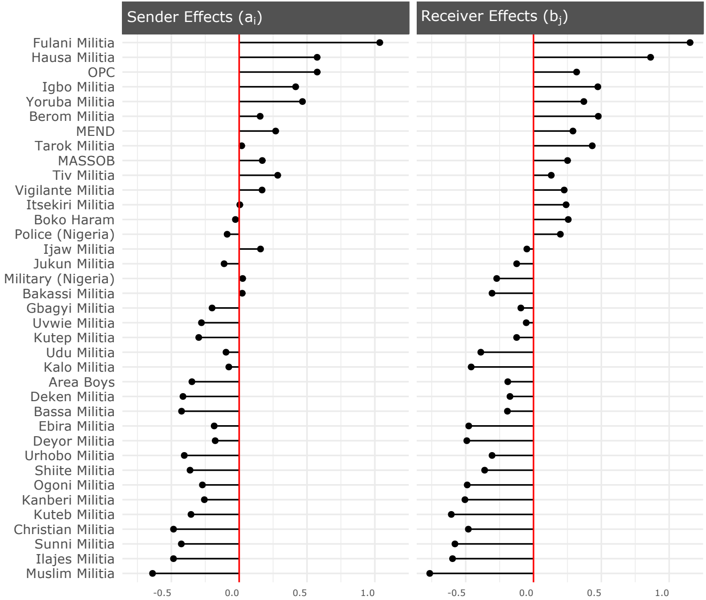
- Country-Level covariates:
    - Post Boko-Haram
    - Neighborhood conflict
    - Election year
- Sender and Receiver-Level Covariates:
    - Violence against civilians
    - Riots/Protests directed against actor
    - Geographic spread

Model Results

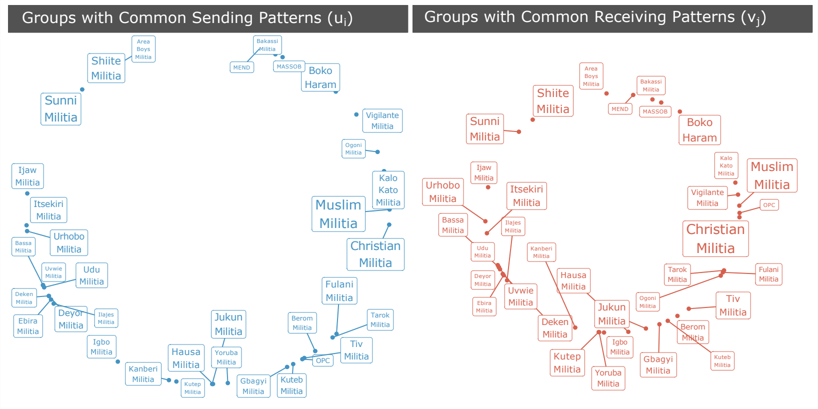




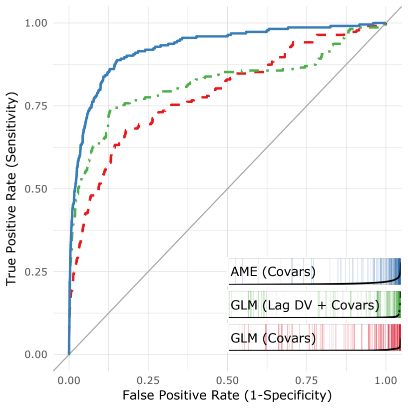
Additive Sender/Receiver Random Effects

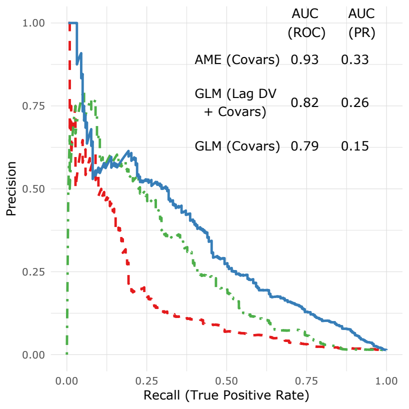


Multiplicative Effects

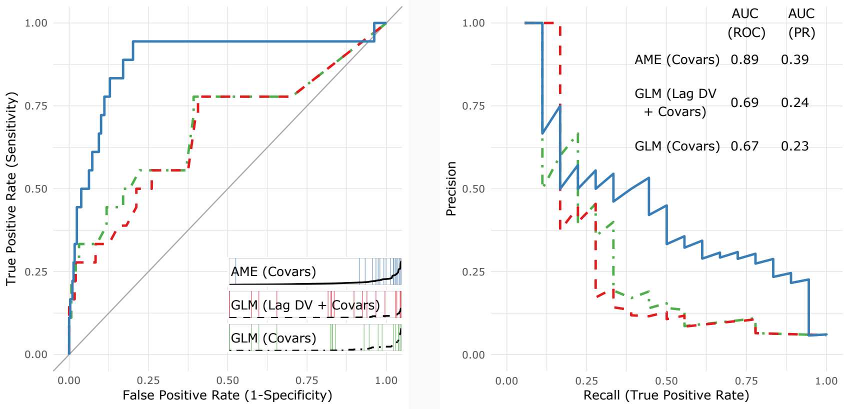


Out of Sample Cross-Validation



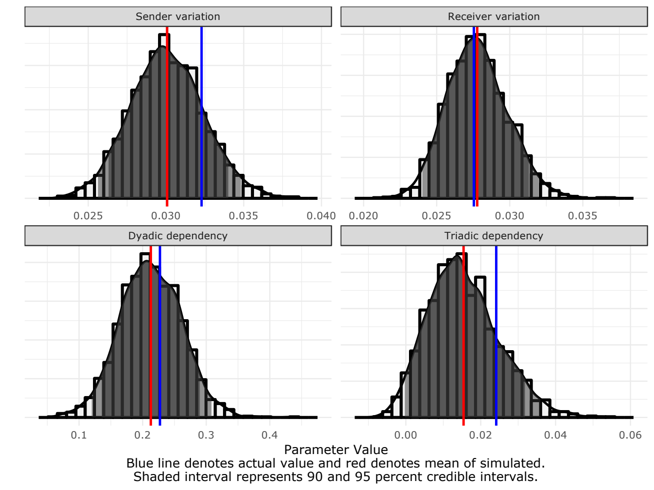


Out of Sample Forecast



Thanks.

Network GOF



Sender variation
0.025
0.030
0.035
0.040
Receiver variation
0.020
0.025
0.030
0.035
Dyadic dependency
0.1
0.2
0.3
0.4
Triadic dependency
0.00
0.02
0.04
0.06
Parameter Value
Blue line denotes actual value and red denotes mean of simulated.
Shaded interval represents 90 and 95 percent credible intervals.