Embracing Network Interdependence in International Relations[1]

Michael D. Ward
Department of Political Science, Duke University

Prepared presentation at the Department of Political Science, Texas A&M University, November 5, 2017.

[1]Based upon joint works with Shahryar Minhas (Michigan State University), Peter Hoff (Duke University), Cassy L. Dorff (University of New Mexico), Max Gallop (University of Strathclyde), Juan Tellez (Duke University), Howard Liu (Duke University), & Margaret Foster (Duke University).

Abstract

An Abstract for a Talk?

- Much international relations scholarship concerns dyads. Dyadic observations do not typically satisfy the conditional independence criterion required for many statistical approaches. As a result, studies often produce results with biased coefficient estimates and poorly calibrated standard errors. These biases have profound consequences for evaluating parametric models.
- We present an alternative, regression-based, approach that accounts for the dependencies complicating this type of analysis, and present a simulation exercise highlighting the model's ability to account for the dyadic dependencies. We replicate five prominent studies, comparing the standard approach to our alternative.
- Conventional methods overstate the effect of key variables, underestimate the uncertainties, and often lead researchers to faulty conclusions about the substantive importance of their variables.
- Our approach dominates in terms of out-of-sample cross-validations, rendering it more useful in forecasting applications and in modeling the data generating process behind outcomes of interest.

Overview of Seminar

Please, ask questions when you have them.

Motivation

Relational data consists of

- a set of units or nodes
- a set of measurements, y_{ij}, specific to pairs of nodes (i, j)

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

$\longrightarrow$

	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

Relational data assumptions

GLM: $y_{ij} \sim \beta^T X_{ij} + e_{ij}$
Networks typically show evidence against independence of $e_{ij} : i \neq j$
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)	Beck et al. (1998)	Snijders (2011)
Frank & Strauss (1986)	Signorino (1999)	Erikson et al. (2014)
Kenny (1996)	Li & Loken (2002)	Aronow et al. (2015)
Krackhardt (1998)	Ho & Ward (2004)	Athey et al. (2016)

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)

We use this model to form the additive effects portion of AME

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

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

- μ baseline measure of network activity
- 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)

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

- 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)

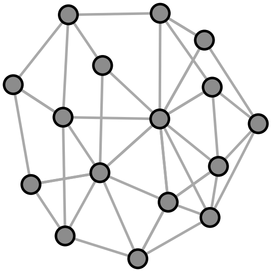
Social Relations Model (The "A" in AME)

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

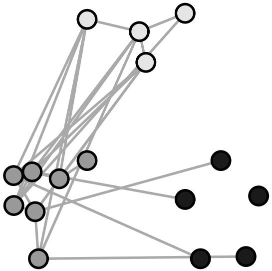
- ϵ_{ij} captures the within dyad effect
- Second-order dependencies are described by σ_{ϵ}^2
- Reciprocity, aka within dyad correlation, represented by ρ

Third-Order Dependencies

Homophily



Stochastic Equivalence



To account for these patterns we can build on what we have so far and find an expression for γ:

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

What Exactly Is Third-Order Dependency?

People Want to Know!

- If each node has individual characteristics, say x_i and x_j for two nodes. Then, $x_i \times x_j$ is the multiplicative product of their nodal characteristics and may be related to their relationship, y_{ij}. Let x be an encoding of whether a node is democratic or not. Then, $x_i \times x_j \equiv 1 \iff$ *both are democracies*. This approach is often used in standard analyses as a variable often known as *joint democracy*. In the context of networks, a positive effect from the interaction of x_i and x_j might affect some dyadic variable, such as not being in a militarized dispute against one another—often denoted the democratic peace. This positive effect is known in network terminology as *homophily*. Most studies just look at the correlation between aggregated variables, but the proposed approach can more directly estimate the degree of homophily.

What Exactly Is Third-Order Dependency?

People Want to Know!

- Homophily which stems from nodal characteristics results in certain types of network structures: 1) transitivity, 2) balance, and 3) clustering. For example, homophily tends to induce a large number of *triads*, since if i is similar to j and as a result has a link with j, for the same reason it will share a link with k if j and k are similar, thus creating a triad. Groups of triads tend to result in *clusters*. This can result in clusters in which many elements in the cluster are not only related to one another but share similar characteristics. This is known as *stochastic equivalence*.
- To capture this kind of stochastic equivalence—and other types of triadic dependence—we include a latent representation of nodal characteristics in the regression model. This allows the network representation in a matrix to be dependent upon how similar these latent characteristics are. These latent nodal characteristics are represented by a type of multiplicative effects that are expressed as $\mathbf{u}_i^\top \mathbf{v}_i$. This captures the third-order dependencies, i.e., the dependencies among triads.

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

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

D is a $K \times K$ diagonal matrix

This can account for both stochastic equivalence and homophily, as well as clustering and balance.

- sample $\boldsymbol{\theta} \mid \boldsymbol{\beta}, \mathbf{X}, \boldsymbol{\theta}, \mathbf{a}, \mathbf{b}, \mathbf{U}, \mathbf{V}, \Sigma_{ab}, \rho,$ and σ_ϵ^2 (Normal)
- sample $\boldsymbol{\beta} \mid \mathbf{X}, \boldsymbol{\theta}, \mathbf{a}, \mathbf{b}, \mathbf{U}, \mathbf{V}, \Sigma_{ab}, \rho,$ and σ_ϵ^2 (Normal)
- sample $\mathbf{a}, \mathbf{b} \mid \boldsymbol{\beta}, \mathbf{X}, \boldsymbol{\theta}, \mathbf{U}, \mathbf{V}, \Sigma_{ab}, \rho,$ and σ_ϵ^2 (Normal)
- sample $\Sigma_{ab} \mid \boldsymbol{\beta}, \mathbf{X}, \boldsymbol{\theta}, \mathbf{a}, \mathbf{b}, \mathbf{U}, \mathbf{V}, \rho,$ and σ_ϵ^2 (Inverse-Wishart)
- update ρ using a Metropolis-Hastings step with proposal $p^*|p \sim$ truncated normal$_{[-1,1]}(\rho, \sigma_\epsilon^2)$
- sample $\sigma_\epsilon^2 \mid \boldsymbol{\beta}, \mathbf{X}, \boldsymbol{\theta}, \mathbf{a}, \mathbf{b}, \mathbf{U}, \mathbf{V}, \Sigma_{ab},$ and ρ (Inverse-Gamma)
- For each $k \in K$:
  - Sample $\mathbf{U}_{[,k]} \mid \boldsymbol{\beta}, \mathbf{X}, \boldsymbol{\theta}, \mathbf{a}, \mathbf{b}, \mathbf{U}_{[,-k]}, \mathbf{V}, \Sigma_{ab}, \rho,$ and σ_ϵ^2 (Normal)
  - Sample $\mathbf{V}_{[,k]} \mid \boldsymbol{\beta}, \mathbf{X}, \boldsymbol{\theta}, \mathbf{a}, \mathbf{b}, \mathbf{U}, \mathbf{V}_{[,-k]}, \Sigma_{ab}, \rho,$ and σ_ϵ^2 (Normal)
  - Sample $\mathbf{D}_{[k,k]} \mid \boldsymbol{\beta}, \mathbf{X}, \boldsymbol{\theta}, \mathbf{a}, \mathbf{b}, \mathbf{U}, \mathbf{V}, \Sigma_{ab}, \rho,$ and σ_ϵ^2 (Normal)[2]

[2] Subsequent to estimation, **D** matrix is absorbed into the calculation for **V** as we iterate through K.

Simulation Comparing GLM and LFM

We utilize a simulation study to highlight the utility of AME as an inferential tool for dyadic analysis. The goal of the simulation is to assess how well AME can provide unbiased and well-calibrated estimates of coefficient parameters in the presence of unobserved dependencies. Specifically, we are concerned with conducting inference on regression parameters of a linear model for a network in the case where there is an omitted variable. For instance, assume that the true data-generating process for a particular Y is given by:

$$y_{i,j} \sim \mu + \beta x_{i,j} + \gamma w_{i,j} + \epsilon_{i,j} \quad (1)$$

where $Y = \{y_{i,j}\} \in \mathbb{R}^{n \times n}$ is an observed sociomatrix, $X = \{x_{i,j}\} \in \mathbb{R}^{n \times n}$ is a matrix of observed dyad-specific characteristics, and $W = \{w_{i,j}\} \in \mathbb{R}^{n \times n}$ is a matrix of unobserved dyad-specific characteristics. Y can be thought of as a dyadic dependent variable, X and W are both dyadic covariates that are a part of the data-generating process for Y, but W is not observed.

Simulation Comparing GLM and LFM

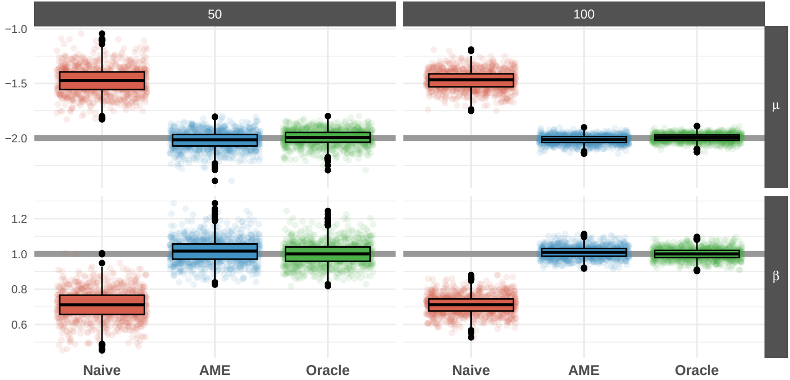
Three Alternatives:

- the standard international relations approach assuming independent errors (current state-of-the discipline);
- the AME approach with a unidimensional latent factor space ($K = 1$);
- an "oracle" regression model that assumes we have measured all sources of dependencies and thus includes both $x_{i,j}$ and $w_{i,j}$.

For the simulation we set the true value of μ (the intercept term) to -2 and β (the effect of X on Y) to 1. The value of γ is also set to 1, which corresponds to an example where W is associated with homophily.

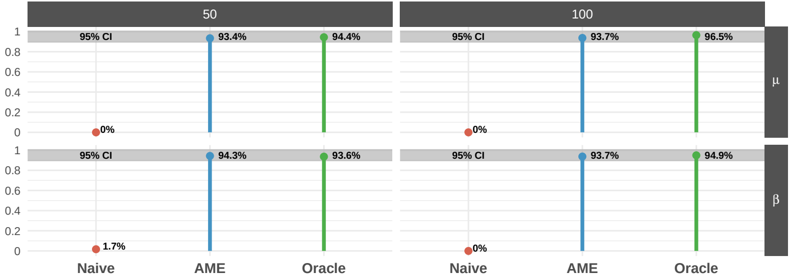
Regression Parameters from Simulations

Figure: Regression parameter estimates for the naive, AME, and oracle models from 1,000 simulations. Summary statistics are presented through a traditional box plot, and the estimates from each simulation are visualized as well as points.



Coverage in the Simulations

Figure: Proportion of times the true value fell within the estimated 95% confidence interval for the naive, AME, and oracle models from 1,000 simulations.



Does it Really Matter?

Let's Look at Five Recent Publications in IR

Table: Features of the Studies Re-estimated.

	Model	Date Range	N. Actors	N. Dyads	Dyads Type	Clustering $\sigma_{\hat{\beta}}$
Reiter & Stam (JCR; 2003)	Logit	1945–1995	193	753,456	Directed	Robust
McDonald (JCR; 2004)	Logit	1959–2002	198	92,354	Undirected	Robust
Rose (AER; 2004)	OLS	1948–1999	177	234,597	Directed	Robust
Weeks (APSR; 2012)	Logit	1946–1999	197	901,540	Directed	Robust
Gibler (APSR; 2017)	Logit	1816–2008	193	650,557	Undirected	none

Table: Here we provide a brief summary of the key variable in each of the five replications and a note about whether or not the finding is replicated when using our network-based approach.

Study	Central Finding	Replicates in a Network Model?
Reiter & Stam (2003)	Personalist Regimes Attack Democracies, Not Vice Versa	Does Not Replicate
McDonald (2004)	Lower Trade Barriers and Higher Trade Lead to Peace	Does Not Replicate
Rose (2004)	WTO Membership Does not Affect Trade	Partially Replicates
Weeks (2012)	Bosses, Juntas, and Strongmen are more Aggressive; Machines are Not	Does Not Replicate
Gibler (2017)	Power Parity at Time of Entry to International System Inceases Conflict	Does Not Replicate

Performance Statistics

In out-of-sample tests

Study	Statistic	GLM	AME
Reiter & Stam (2003)	Area Under ROC Curve, AUC-ROC	0.92	0.96
	Area Under PR Curve, AUC-PR	0.08	0.15
McDonald (2004)	AUC-ROC	0.92	0.99
	AUC-PR	0.13	0.28
Rose (2004)	RMSE	3.23	1.99
	RMDSE	2.01	1.06
Weeks (2012)	AUC-ROC	0.64	0.97
	AUC-PR	0.00	0.15
Gibler (2017)	AUC-ROC	0.52	0.91
	AUC-PR	0.00	0.08

Reiter-Stam Performance Comparisons

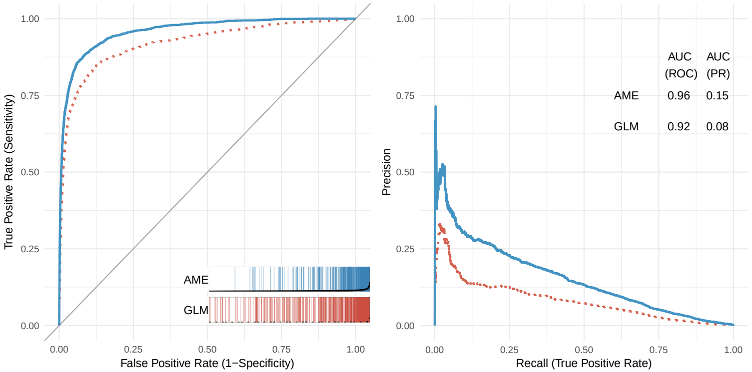


Figure: Reiter-Stam Performance

McDonald Performance Comparisons

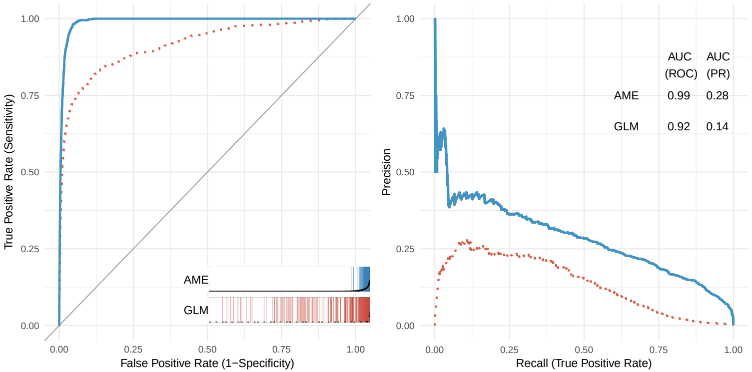


Figure: McDonald Performance

Rose Performance Comparisons

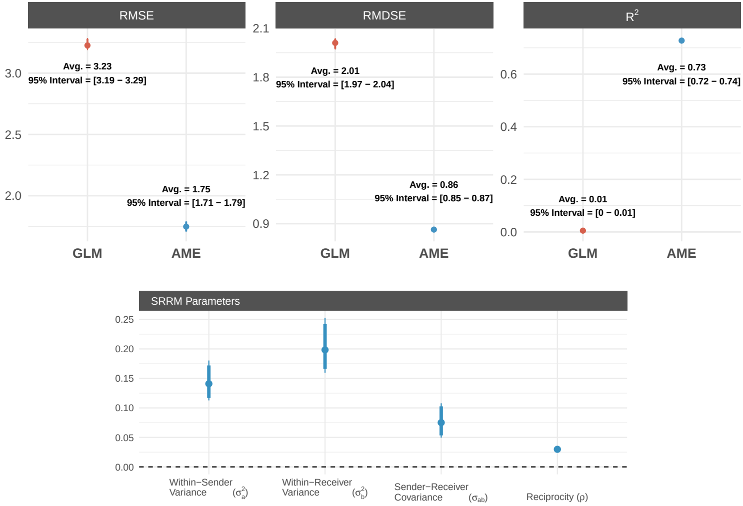


Figure: Rose Performance Statistics

Weeks Performance Comparisons

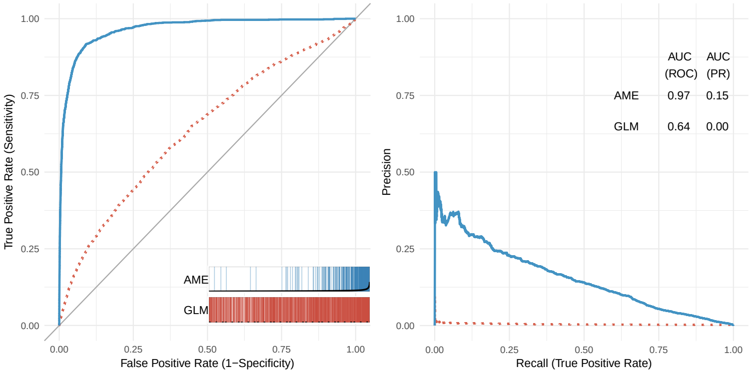


Figure: Weeks Performance; no single MID predicted

Gibler Performance Comparisons

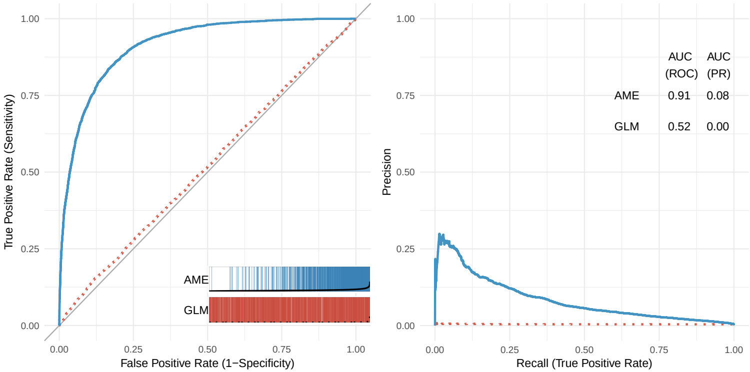


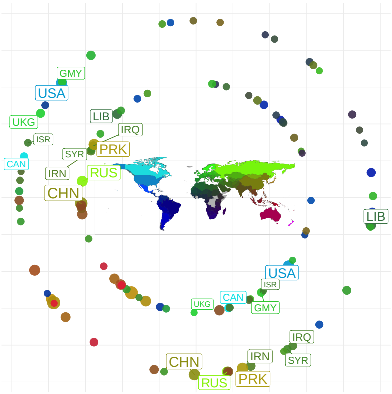
Figure: Gibler Performance; no single MID predicted

Some Lessons Learned

- Findings that do not take interdependencies into account lose their statistical significance when network effects are estimated via AME. In GLM approaches to the analysis of dyadic data, coefficents are biased, imprecisely measured, and have poorly calibrated standard errors. This means that significance testing (for better or worse) is compromised when network effects are ignored.
- New insights emerge from the additional information derived from AME. In particular, there is actual information about the dependencies so that clusters can be identified, the extent of reciprocity at the dyad level, and among dyads.
- The actual results—not the estimated coefficients and their covariances—which are generated by the models differ greatly in expectations.
- The AME approach dominates the GLM approaches in terms of performance. Not only it is better at correctly identifying cases in which the dependent variable takes a value of 0, but it also dominates at correctly identifying occurrences of 1.

Networks Really Matter

We miss a lot when we ignore them.



- clustering
- different patterns for senders and receivers
- similar nodal, and dyadic patterns
- gauge reciprocity (and orthogonality)

Thanks!

There's more, but I'll stop here for your questions and comments.

More work; copies available upon request:

- Shahryar Minhas, Peter D. Hoff, and Michael D. Ward. 2017. *Additive and Multiplicative Latent Factor Models for Network Influence*.
- Shahryar Minhas, Peter D. Hoff, and Michael D. Ward. 2017. *Influence Networks in International Relations*.
- Shahryar Minhas, Cassy L. Dorff, Margaret Foster, Max Gallop, Howard Liu, Juan Tellez, and Michael D. Ward. 2017. *Taking Dyads Seriously*.

Thank you!

Example $\mathcal{R}$ code

Using the AMEN function requires formatting data into a particular structure. The primary distinction in data formatting is whether the outcome of interest represents a directed or undirected network. Use this link `GitHub.com/s7minhas/amen`

If undirected, the AMEN function has three main inputs:

- Y: a T length **list** of $n \times n$ adjacency matrices, where $T =$ number of years in the dataset and $n =$ number of nodes in the network.
- Xdyad: a T length **list** of $n \times n \times p$ arrays, where $p =$ number of dyadic covariates in dataset.
- Xrow: a T length **list** of $n \times p$ matrices, where $p =$ number of monadic (nodal) covariates in dataset.

If directed, AMEN further requires:

- Xrow: a T length list of $n \times p$ matrices, where $p =$ number of sender (nodal) covariates in dataset.
- Xcol: a T length list of $n \times p$ matrices, where $p =$ number of receiver (nodal) covariates in dataset.

Beyond the data inputs, the AMEN function requires additional specification:

- model: how to model the outcome variable, e.g., 'logit'
- symmetric: whether the input network is symmetric
- intercept: whether to estimate an intercept
- nscan: number of iterations of the Markov chain
- burn: burn-in period
- odens: thinning interval
- R: dimension of the multiplicative effect (referred to as K in the paper)
- gof: whether to calculate goodness-of-fit statistics

Actual Code

```
# running in parallel varying k
imps = 10000 ; brn = 25000 ; ods = 10 ; latDims = 0:3
# Run amen in parallel library(doParallel) ; library(foreach) ;
cl=makeCluster(4)
registerDoParallel(cl)
foreach(ii=1:length(latDims), .packages=c("amen")) %dopar%
# load previous model run
load(prevModelFiles[ii])
# extract start vals
startVals0 = ameFit$'startVals'
# dump rest
rm(ameFit)
ameFit = ame_repL(
Y=yList,Xdyad=xDyadList,Xrow=NULL,Xcol=NULL,
model="bin",symmetric=FALSE,intercept=TRUE,R=latDims[ii],
nscan=imps, seed=1, burn=brn, odens=ods,
plot=FALSE, print=FALSE, gof=TRUE, startVals=startVals0,
periodicSave=TRUE )
save(ameFit, file=paste0('model_k', latDims[ii],'_v2.rda') )
```

backup slides

Go ahead, ask me about ERGM.

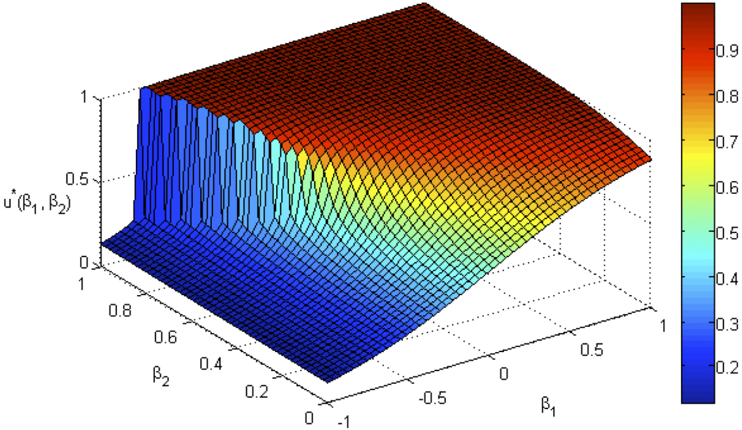
There is a problem with ERGM.
It is a bug, not a feature.

It is a bug, not a feature. If you don't believe us, believe these scholars:

1. Schweinberger, M. (2011). Instability, Sensitivity, and Degeneracy of Discrete Exponential Families. **Journal of the American Statistical Association**, 106(496):13611370.
2. Schweinberger, M. and Handcock, M. S. (2015). Local Dependence in Random Graph Models: Characterization, Properties and Statistical Inference. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, 77(3):647676.
3. Chatterjee, S. and Diaconis, P. (2013). Estimating and Understanding Exponential Random Graph Models. **The Annals of Statistics**, 41(5):24282461.
4. Rastelli, R., Friel, N., and Raftery, A. E. (2016). Properties of Latent Variable Network Models. **Network Science**, 4(4):126.

What Is Really Wrong?

You may assume this is your likelihood surface. Courtesy of Chatterjee and Diaconis, who concluded that sufficient statistics are not sufficient to solve this problem. Estimation of ERGMs contains this flaw, with pseudolikelihood, maximum likelihood, or MCMC.



What Does This Mean?

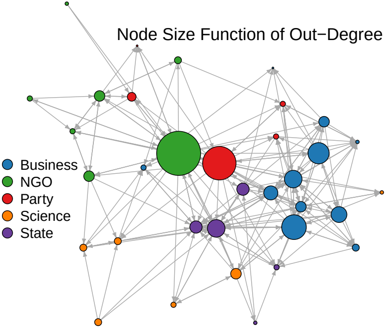
1. Probabilistic ERGM models place almost all of the probability on networks that are either nearly empty (degenerate) with no linkages or nearly saturated with all nodes being interconnected.
2. The likelihood surface contains steep or discontinuous gradients that render it impossible to solve numerically (or analytically). Even (especially) for very small networks this is problematic.
3. Find a different way to estimate network data.

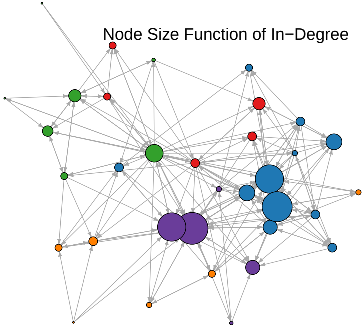
Extant Political Science Examples

Extant Political Science Examples Using Latent Distance Models Are Wrong.

1. There is a problem with the MCMC estimator in `latentnet`. Don't use it until it is fixed.
2. Latent distance models are hard to correctly interpret, in prominent evaluations in the literature it is done incorrectly. This is explained in their introduction a decade ago.
3. The advice in Cranmer et al is wrong. Latent network models are not biased and when correctly implemented out-perform ERGM models.

Cross-sectional network measuring whether an actor indicated that they collaborated with another during the policy design of the Swiss CO_2 act (Ingold 2008)

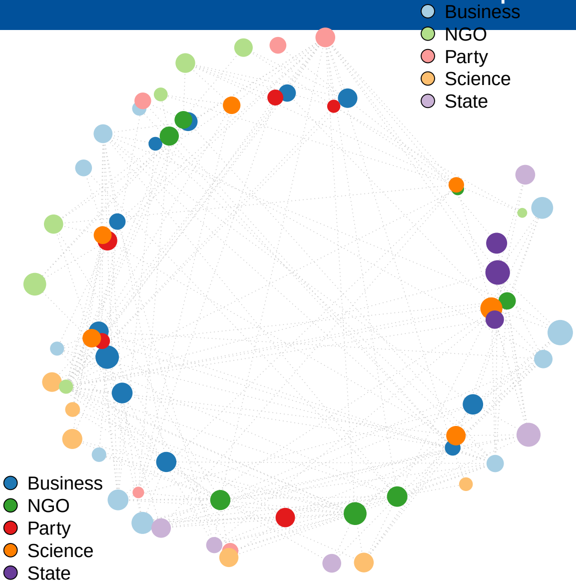




Parameter Estimates for Swiss Example

	Expected Effect	Logit	MRQAP	LSM	ERGM	AME
Conflicting policy preferences						
Business vs. NGO	$-$	-0.86	-0.87*	-1.37*	-1.11*	-1.37*
Opposition/alliance	$+$	1.21*	1.14*	0.00	1.22*	1.08*
Preference dissimilarity	$-$	-0.07	-0.60	-1.76*	-0.44	-0.79*
Transaction costs						
Joint forum participation	$+$	0.88*	0.75*	1.51*	0.90*	0.92*
Influence						
Influence attribution	$+$	1.20*	1.29*	0.08	1.00*	1.09*
Alter's influence indegree	$+$	0.10*	0.11*	0.01	0.21*	0.11*
Influence absolute diff.	$-$	-0.03*	-0.06*	0.04	-0.05*	-0.07*
Alter = Government actor	$+$	0.63*	0.68	-0.46	1.04*	0.55
Functional requirements						
Ego = Environmental NGO	$+$	0.88*	0.99	-0.60	0.79*	0.67
Same actor type	$+$	0.74*	1.12*	1.17*	0.99*	1.04*

Latent Factor Visualization for Swiss Example

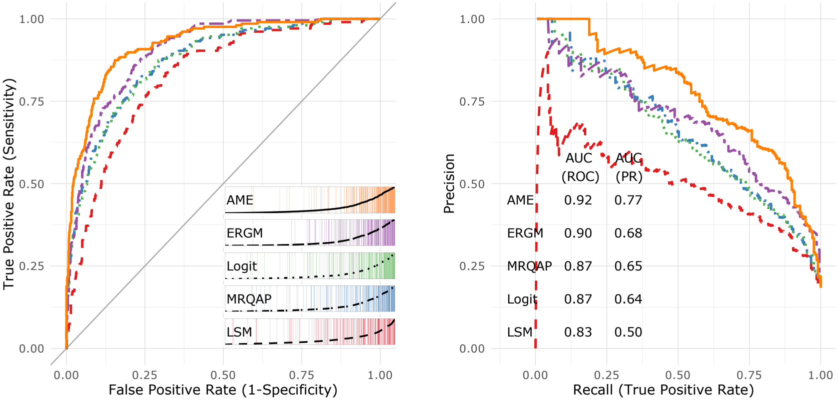


Out-of-Sample Performance Assessment: Strategy

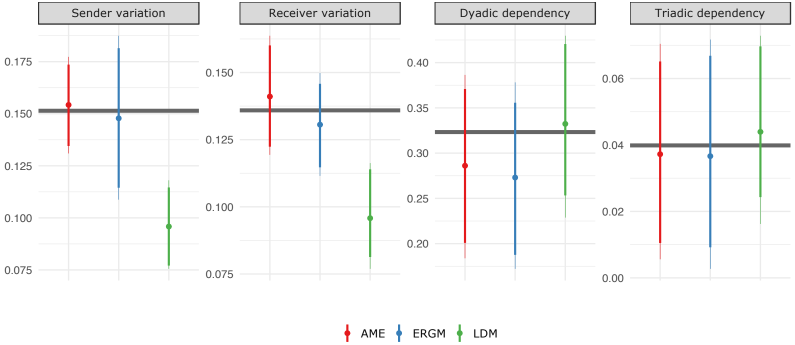
- Randomly divide the $n \times (n-1)$ data points into S sets of roughly equal size, letting s_{ij} be the set to which pair $\{ij\}$ is assigned.
- For each $s \in \{1, \ldots, S\}$:
    - Obtain estimates of the model parameters conditional on $\{y_{ij} : s_{ij} \neq s\}$, the data on pairs not in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

This procedure generates a sociomatrix of out-of-sample predictions of the observed data

Performance Comparison for Swiss Example



Network Dependencies: Swiss Example

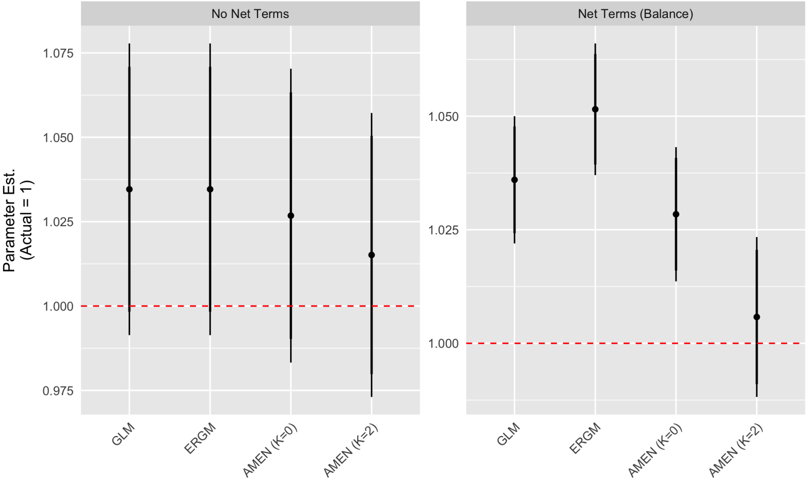


Much current advice is wrong.

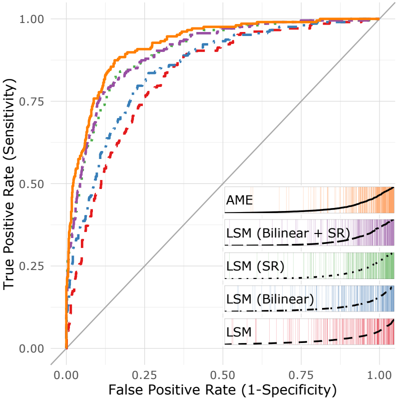
Unfortunately some misleading and incorrect advice is in the current literature. Maybe it will be corrected.

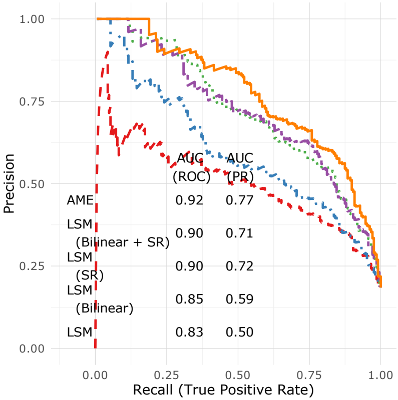
1. Cranmer, Leifeld, McClurg, Rolfe (2017 AJPS) stated that latent space models are only preferable if there are few isolates, enough observations, and the interdependence is theoretically uninteresting.
2. All of this is wrong. Not only do they incorrectly assume all latent models to be based on Euclidean distance but their statistical comparisons are based on a program that contains known errors. Moreover, using BIC to choose among different network models is itself questionable.
3. It is not true that "ERGM offers the most preferable combination of model fit and number of parameters," despite their assertions.

Simulation Comparison

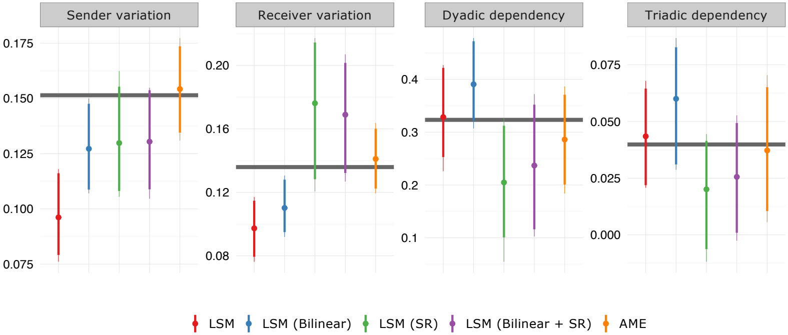


AMEN v LSM Performance

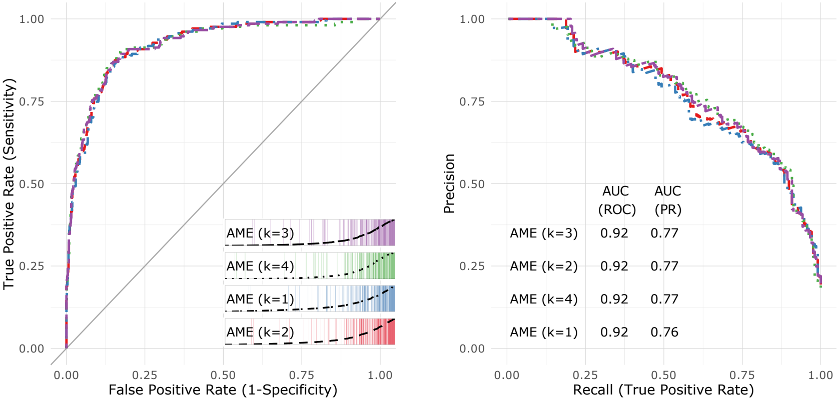




AMEN versus LSM Net Dependence



AMEN varying K Performance



AMEN varying K Net Dependence