A Local Structure Graph Model: Formation of Network Edges as a Function of Other Edges

Olga V. Chyzh

Departments of Political Science and Statistics, Iowa State University

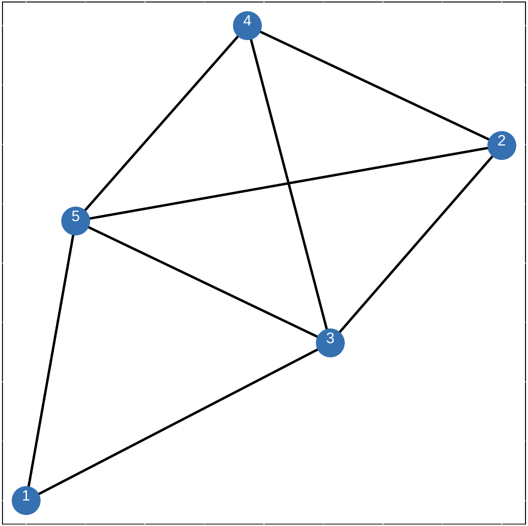
(with Mark S. Kaiser)

November 17, 2017

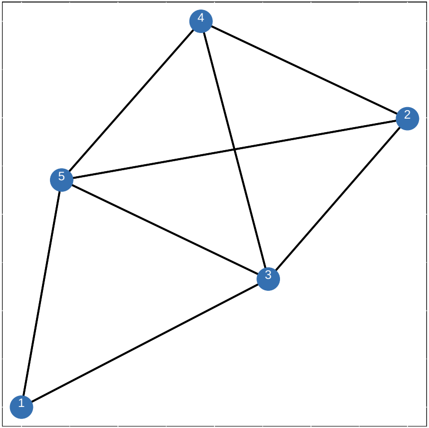
Objectives

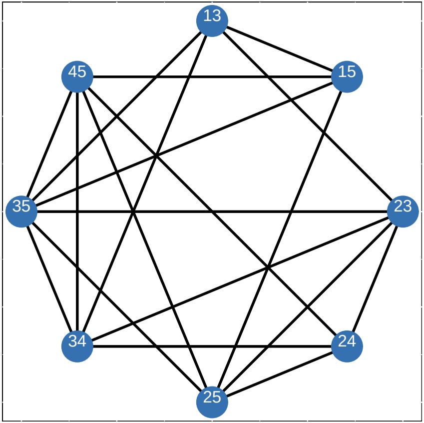
1. Expand network theory to a new class of problems:
   - E.g.: coalition-building, diffusion, tipping-point processes;
   - Beyond network of edges among nodes to networks of *edges among edges*;
2. Demonstrate a statistical way to model such processes—a local structure graph model (LSGM);
   - Monte Carlo results;
   - Two empirical applications.

Relationships Among Nodes



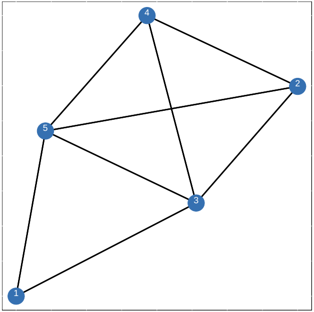
Degrees of Dependence

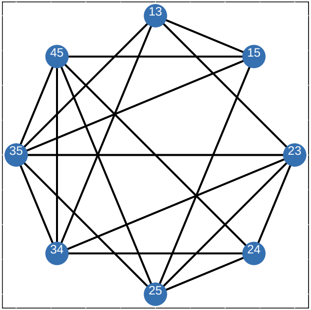




Network of Nodes ⇒ Network of Edges

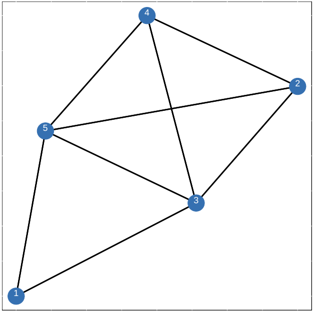
Degrees of Dependence

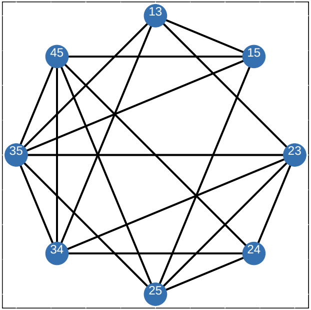




- Edges are connected if they share a common node;

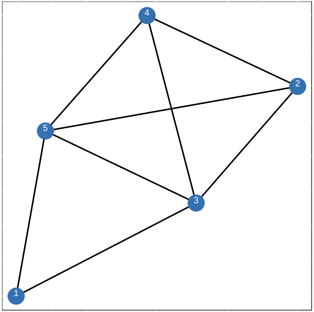
Degrees of Dependence

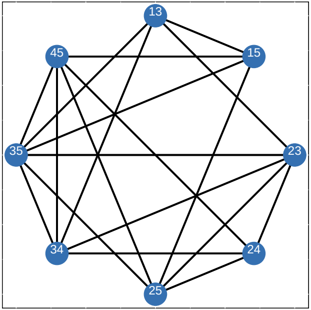




- Edges are connected if they share a common node;
- Edges among edges may represent other types of relationships among edges;

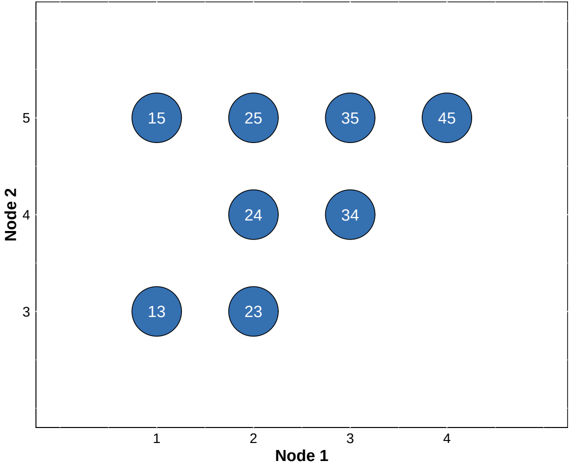
Degrees of Dependence



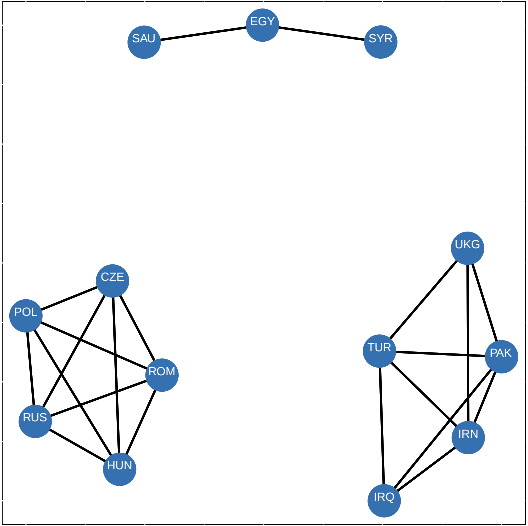


- Edges are connected if they share a common node;
- Edges among edges may represent other types of relationships among edges;
- Edges may be connected if they both connect the two nodes of the same color or two odd-numbered nodes.

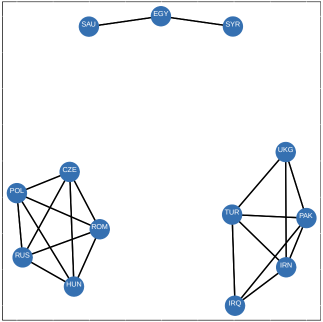
Continuous Edge Connectivities



Political Applications: Allies 1955

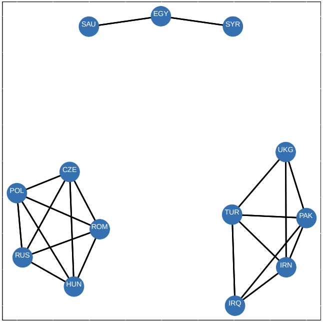


Relationships Among Nodes: Allies 1955



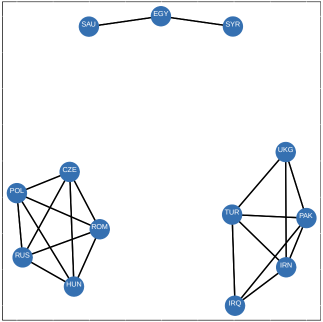
- This framework allows for modeling alliance formation as a function of nodal and edge-level covariates;

Relationships Among Nodes: Allies 1955



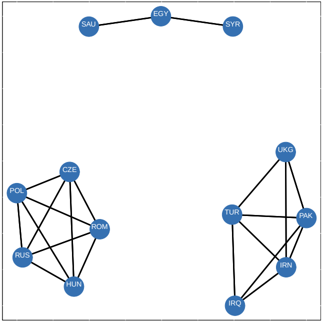
- This framework allows for modeling alliance formation as a function of nodal and edge-level covariates;
- Many theories suggest that alliance edges realize *in response to realization of other edges* (e.g., balancing against ideological threat, “birds-of-a-feather”);

Relationships Among Nodes: Allies 1955



- This framework allows for modeling alliance formation as a function of nodal and edge-level covariates;
- Many theories suggest that alliance edges realize *in response to realization of other edges* (e.g., balancing against ideological threat, "birds-of-a-feather");
- Need to treat alliances as nodes and measure relationships among them—a network of edges among edges;

Relationships Among Nodes: Allies 1955

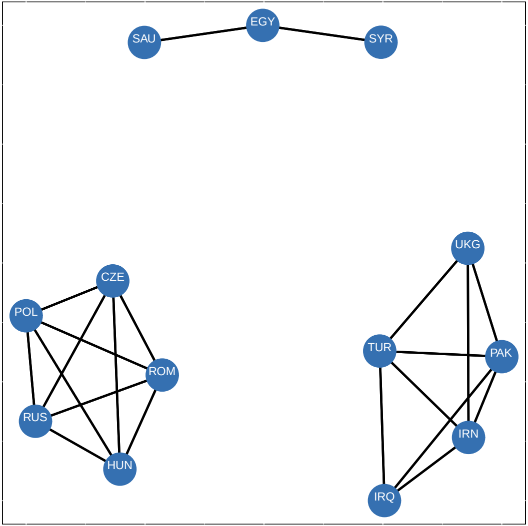


- This framework allows for modeling alliance formation as a function of nodal and edge-level covariates;
- Many theories suggest that alliance edges realize *in response to realization of other edges* (e.g., balancing against ideological threat, "birds-of-a-feather");
- Need to treat alliances as nodes and measure relationships among them—a network of edges among edges;
- Similar logic applies to modeling formation of legislative coalitions or advocacy groups.

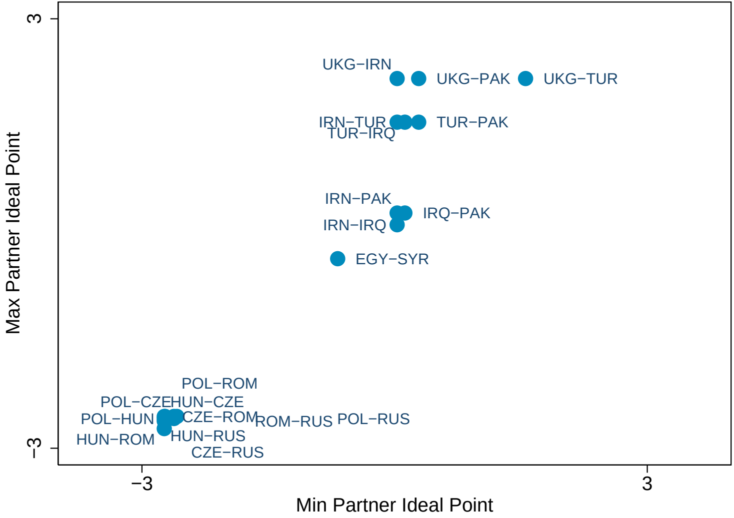
Placing Alliances within Ideational Space

- Ideal Point scores based on UNGA voting (Bailey, Strezhnev, & Voeten, 2016):
  - All states, between 1946-2007;
  - Range between ± 3, standard normal distribution;
  - US scores range between 1.18 and 2.06;
  - Russia/Soviet Union–between -2.74 and 1.12;
- Use each alliance partner score as (x, y) in Cartesian space.
- Distance between alliances measures ideational distance.

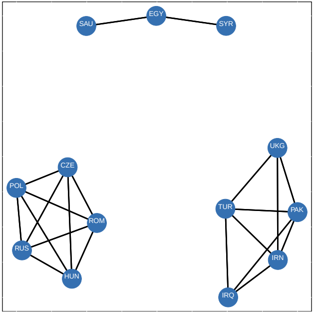
Political Applications: Allies 1955

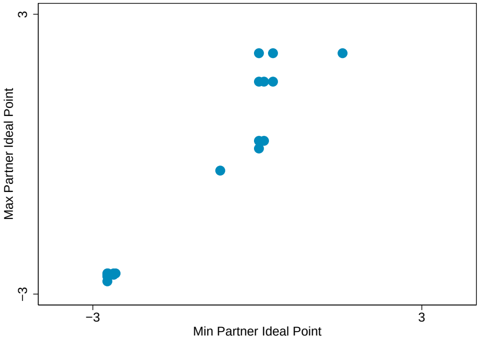


Ideational Distance Among Alliances: 1955

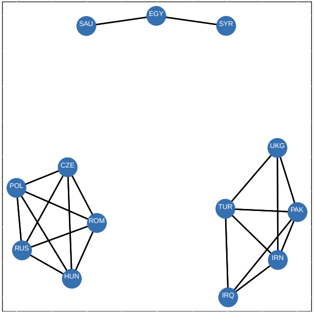


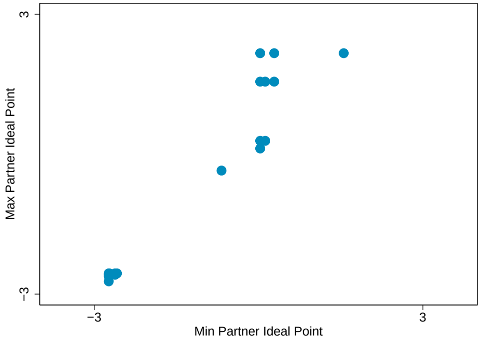
Alliances: 1955





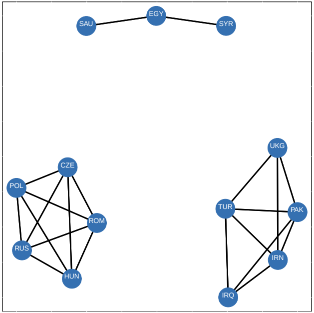
Alliances: 1955

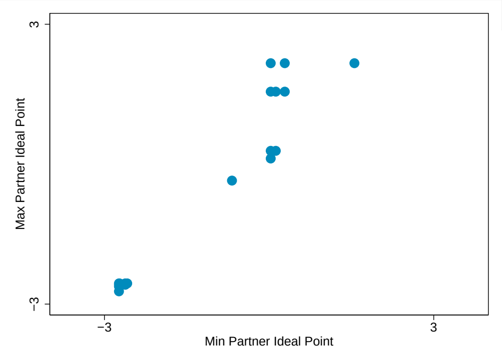




- RUS–POL–HUN–ROM–CZE bloc is much more ideationally cohesive than UKG–TUR–PAK–IRN–IRQ.

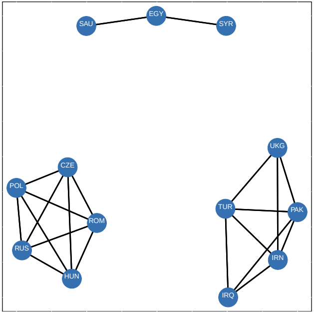
Alliances: 1955

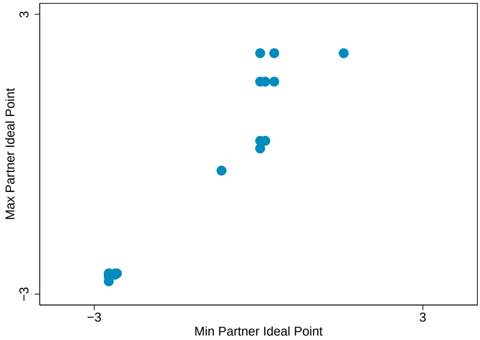




- RUS–POL–HUN–ROM–CZE bloc is much more ideationally cohesive than UKG–TUR–PAK–IRN–IRQ.
- Two blocs are located roughly in opposite parts of the ideational spectrum—polarization, "birds-of-a-feather" theory?

Alliances: 1955

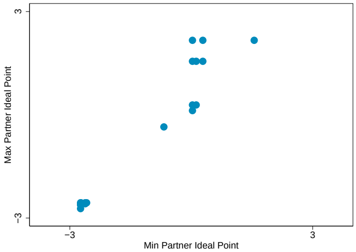




- RUS–POL–HUN–ROM–CZE bloc is much more ideationally cohesive than UKG–TUR–PAK–IRN–IRQ.
- Two blocs are located roughly in opposite parts of the ideational spectrum—polarization, "birds-of-a-feather" theory?
- Alliances tend to form among ideationally similar states—ideological bandwagoning?

The Estimator

- Estimate a model of edges that form in response to other edges;
- Use a local structure graph model (LSGM) (Casleton, Nordman, Kaiser 2016, Besag 1974);
- Treat edges as observations and model local dependence in edge formation by specifying a source of connectivity:



The Estimator: Set-up

- Suppose i is an potential edge in a network of potential edges (realized and unrealized),

The Estimator: Set-up

- Suppose i is an potential edge in a network of potential edges (realized and unrealized),
- Then $s_i = (u_i, v_i)$ is i's location in Cartesian space.

The Estimator: Set-up

- Suppose i is an potential edge in a network of potential edges (realized and unrealized),
- Then $s_i = (u_i, v_i)$ is i's location in Cartesian space.
- Denote the binary random variable, $y(s_i) = y_i$, so that:

$$y_i = \begin{cases} 1 & \text{if an edge is realized} \\ 0 & \text{otherwise} \end{cases}$$

The Estimator: Set-up

- Suppose i is an potential edge in a network of potential edges (realized and unrealized),
- Then $s_i = (u_i, v_i)$ is i's location in Cartesian space.
- Denote the binary random variable, $y(s_i) = y_i$, so that:
  $y_i = \begin{cases} 1 & \text{if an edge is realized} \\ 0 & \text{otherwise} \end{cases}$
- Define i's neighbors as $N_i = \{s_j : s_j \text{ is a neighbor of } s_i\}$.

The Estimator: Set-up

- Suppose i is an potential edge in a network of potential edges (realized and unrealized),
- Then $s_i = (u_i, v_i)$ is i's location in Cartesian space.
- Denote the binary random variable, $y(s_i) = y_i$, so that:
  $y_i = \begin{cases} 1 & \text{if an edge is realized} \\ 0 & \text{otherwise} \end{cases}$
- Define i's neighbors as $N_i = \{s_j : s_j \text{ is a neighbor of } s_i\}$.
- Make a Markov assumption of conditional spatial independence:

$$f(y(s_i)|\mathbf{y}(s_j) : s_j \neq s_i) = f(y(s_i)|\mathbf{y}(N_i))$$

The Estimator: Set-up

- Suppose i is an potential edge in a network of potential edges (realized and unrealized),
- Then $s_i = (u_i, v_i)$ is i's location in Cartesian space.
- Denote the binary random variable, $y(s_i) = y_i$, so that:
  $y_i = \begin{cases} 1 & \text{if an edge is realized} \\ 0 & \text{otherwise} \end{cases}$
- Define i's neighbors as $N_i = \{s_j : s_j \text{ is a neighbor of } s_i\}$.
- Make a Markov assumption of conditional spatial independence:

$$f(y(s_i)|\mathbf{y}(s_j) : s_j \neq s_i) = f(y(s_i)|\mathbf{y}(N_i))$$

- If connectivities between edges are continuous, then the Markov assumption is redundant.

The Binary Conditional Distribution

$$\mathrm{P}(Y_i = y_i | \boldsymbol{y}(N_i)) = \exp\left[A_i(\boldsymbol{y}(N_i))y_i - B_i(\boldsymbol{y}(N_i))\right], \tag{1}$$

where A_i is a natural parameter function and $B_i = \log[1 + \exp(A_i(y(N_i)))]$, and $\mathbf{y}\,(N_i)$ is a vector of values of the binary random variables (edges) of i's neighbors.

$$A_i(\mathbf{y}(N_i)) = \log\left(\frac{\kappa_i}{1-\kappa_i}\right) + \eta \sum_{j \in N_i} w_{ij}(y_j - \kappa_j), \qquad (2)$$

where $\log\left(\frac{\kappa_i}{1-\kappa_i}\right) = \boldsymbol{X}_i^T \beta$, $\mathbf{X}_i$ is a column vector of k exogenous covariates, β is a k by 1 vector of estimation parameters, w_{ij} is the ij^{th} element of a matrix of connectivities among edges $\mathbf{W}$, and η is its parameter.

$$A_i(\mathbf{y}(N_i)) = \log\left(\frac{\kappa_i}{1-\kappa_i}\right) + \eta \sum_{j \in N_i} w_{ij}(y_j - \kappa_j), \tag{2}$$

where $\log\left(\frac{\kappa_i}{1-\kappa_i}\right) = \boldsymbol{X}_i^T \beta$, $\mathbf{X}_i$ is a column vector of k exogenous covariates, β is a k by 1 vector of estimation parameters, w_{ij} is the ij^{th} element of a matrix of connectivities among edges $\mathbf{W}$, and η is its parameter.

- When $y_j > \kappa_j$, then the dependence term makes a positive contribution to $A_i(\mathbf{y}(N_i))$—complementary processes;

$$A_i(\mathbf{y}(N_i)) = \log\left(\frac{\kappa_i}{1-\kappa_i}\right) + \eta \sum_{j \in N_i} w_{ij}(y_j - \kappa_j), \qquad (2)$$

where $\log\left(\frac{\kappa_i}{1-\kappa_i}\right) = \boldsymbol{X}_i^T \beta$, $\mathbf{X}_i$ is a column vector of k exogenous covariates, β is a k by 1 vector of estimation parameters, w_{ij} is the ij^{th} element of a matrix of connectivities among edges $\mathbf{W}$, and η is its parameter.

- When $y_j > \kappa_j$, then the dependence term makes a positive contribution to $A_i(\mathbf{y}(N_i))$—complementary processes;
- When $y_j < \kappa_j$, then the dependence term makes a negative contribution to $A_i(\mathbf{y}(N_i))$—substitution-type processes;

$$A_i(\mathbf{y}(N_i)) = \log\left(\frac{\kappa_i}{1-\kappa_i}\right) + \eta \sum_{j \in N_i} w_{ij}(y_j - \kappa_j), \tag{2}$$

where $\log\left(\frac{\kappa_i}{1-\kappa_i}\right) = \boldsymbol{X}_i^T \beta$, $\mathbf{X}_i$ is a column vector of k exogenous covariates, β is a k by 1 vector of estimation parameters, w_{ij} is the ij^{th} element of a matrix of connectivities among edges $\mathbf{W}$, and η is its parameter.

- When $y_j > \kappa_j$, then the dependence term makes a positive contribution to $A_i(\mathbf{y}(N_i))$—complementary processes;
- When $y_j < \kappa_j$, then the dependence term makes a negative contribution to $A_i(\mathbf{y}(N_i))$—substitution-type processes;
- Key condition: $w_{ij} = w_{ji}$.

$$A_i(\mathbf{y}(N_i)) = \log\left(\frac{\kappa_i}{1-\kappa_i}\right) + \eta \sum_{j \in N_i} w_{ij}(y_j - \kappa_j), \quad (2)$$

where $\log\left(\frac{\kappa_i}{1-\kappa_i}\right) = \boldsymbol{X}_i^T \boldsymbol{\beta}$, $\mathbf{X}_i$ is a column vector of k exogenous covariates, $\boldsymbol{\beta}$ is a k by 1 vector of estimation parameters, w_{ij} is the ij^{th} element of a matrix of connectivities among edges $\mathbf{W}$, and η is its parameter.

- When $y_j > \kappa_j$, then the dependence term makes a positive contribution to $A_i(\mathbf{y}(N_i))$—complementary processes;
- When $y_j < \kappa_j$, then the dependence term makes a negative contribution to $A_i(\mathbf{y}(N_i))$—substitution-type processes;
- Key condition: $w_{ij} = w_{ji}$.
- Model does not require (prohibits) row-standardization of $\mathbf{w}$.

Estimation

$$\log PL = \sum_i \{ y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \}, \quad (3)$$

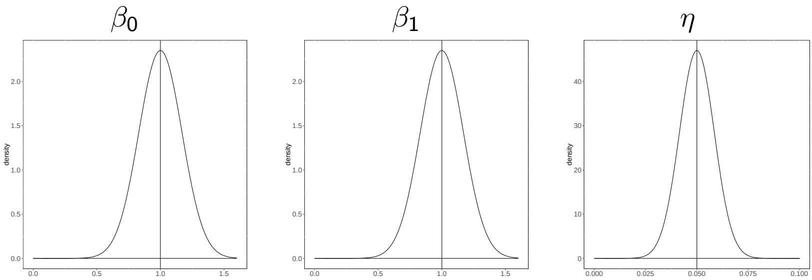
where:

$$p_i = \frac{\exp[A_i(y(N_i))]}{1 + \exp[A_i(y(N_i))]} \quad (4)$$

Monte Carlo Simulations

- Generate a list of $i = 1, 2, \ldots, 100$ units with characteristics captured by variable $\mathbf{x}$, drawn from a standard normal distribution;
- Convert to dyadic data ($n = c(100, 2) = 4950$);
- To generate a meaningful dependence matrix, $\mathbf{W}_{100x100}$, we placed each unit on an evenly spaced ten-by-ten grid and calculated the Euclidean distance between the two units in each dyad.
- Use a Gibbs sampler to generate random variable, Y:
  1. Use a vector $\boldsymbol{y}_0 = \{y_{01}, y_{02}, \ldots, y_{0n}\}$ drawn from a binomial distribution as starting values.
  2. Simulate $y_{11} = f(y|y_{02}, y_{03} \ldots, y_{0n})$.
  3. Simulate $y_{12} = f(y|y_{11}, y_{02}, y_{03}, \ldots, y_{0n})$.
  4. Simulate $y_{13} = f(y|y_{11}, y_{12}, y_{03}, y_{04}, \ldots, y_{0n})$.
  5. Continue until simulate a complete network y_1, then iterate steps (2)-(5) using $\boldsymbol{y}_1$ as starting values;
  6. Discard the first 100 networks for burnin; record every 50^{th} network.

Monte Carlo Results



Note: Given randomly initialized values for all edges, the Gibbs sampler was run with a burn-in of 100 complete graph iterations after which sample graphs were retained from 100,000 subsequent rounds with 50 iterations for thinning. True parameter values are denoted by vertical lines.

Empirical Application 1: International Alliance Network

- Ideological Balancing Hypothesis: We should observe alliance formation in different parts of the ideational space—positive coefficient on the dependence term (Schweller 2004).
- Ideational Clustering: We should observe alliance clustering in ideational space—negative coefficient on the dependence term (Lai and Reiter 2000).

Empirical Application 1: International Alliances

- Data on international alliances between 1946–2007 (Gibler 2009);
- Treat alliances as network edges;
- Use ideational distance among alliances as $\mathbf{W}$;
- Control for military power ratio, trade, joint democracy.

International Alliance Network, 1947-2000

Edge Connectivity:		
Ideational Distance	0.016*	(0.001)
Military Power Ratio	-2.363*	(0.073)
Dyadic Trade	0.015*	(0.005)
Joint Democracy	0.884*	(0.024)
Constant	0.094	(0.072)

Note: $^{*}p < 0.05$ Standard errors are obtained using a parametric bootstrap (1100 simulations of complete networks, 100 burnin and 50 iterations for thinning).

Empirical Application 2: Senate Cosponsorships

- Ideological Balancing Hypothesis: We should observe cosponsorship clusters in the opposite parts of the ideological space—positive coefficient on the dependence term.
- Ideational Clustering: We should observe consponsorship clustering in ideational space—negative coefficient on the dependence term.

Empirical Application 2: Senate Cosponsorships

- Data on cosponsorships of labor-related legislation (Senate of the 107^{th} US Congress);
- Treat all potential cosponsorships as edges;
- Use DWNominate scores (first dimension) to measure ideological distance in the connectivity matrix $\mathbf{W}$;
- Control for same party, labor committee, and minimum seniority.

Cosponsorships on Labor Bills, Senate of the 107^{th} US Congress

Edge Connectivity:		
Ideological Distance	-1.235*	(0.519)
Same Party	0.704*	(0.051)
Labor Committee	0.149*	(0.044)
Minimum Seniority	-0.047*	(0.010)
Constant	0.387*	(0.089)

Notes: Standard errors were obtained using a parametric bootstrap via a Gibbs sampler of 300 complete simulations (50 for burnin and thinning.)

Conclusion

- Many political science applications require conceptualizing networks as dependencies among *edges* rather than nodes.
- Introduce LSGM as a statistical tool for modeling many political processes involving dependence among network edges;
- Applied to modeling formation of international alliance network and legislative cosponsorships;
- Other applications: lobbying groups, parties joining to share ballot lines, multilateral cooperation (sanctions), diffusion, tipping-point processes.

LSGM vs. SAR

- SAR: models feedback loops: by construction, Y_i is a function of outcomes in its neighbors, **AND** the neighbors' neighbors.
- LSGM (CAR): may specify the connectivity matrix or include additional dependence terms to model the effect of neighbors' neighbors, but only first-order effects are modeled "by default";
- Besag (1974) demonstrated that CAR may be estimated by maximizing a pseudo-likelihood—under some conditions results in substantial gains in speed of estimation.
- Standard errors: Gibbs sampler, "Godambe" information matrix.
- The trade-off: LSGM (CAR) requires that the connectivity matrix be symmetric (most applications I've seen have a symmetric W matrix);
- LSGM naturally extends to other functions in the exponential family, e.g. Poisson.

Natural Exponential Family Functions

- $f(y|\theta) = \exp[y\theta - b(\theta) + c(y)]$, where θ is the natural parameter;
- For a binary dependent variable: $f(y|p) = p^y(1-p)^{(1-y)}$;
- Take a natural log and exponentiate:

$$f(y|p) = \exp[y\log(p) + (1-y)\log(1-p)] =$$
$$\exp[y\{\log(p) - \log(1-p)\} + \log(1-p)] =$$
$$\exp[y\theta - b(\theta)],$$

where $\theta = \log\frac{p}{1-p}$, and $b(\theta) = \log(1+e^{\theta})$