Using Potential Games to Parameterize ERG Models

Carter T. Butts

Department of Sociology and

Institute for Mathematical Behavioral Sciences

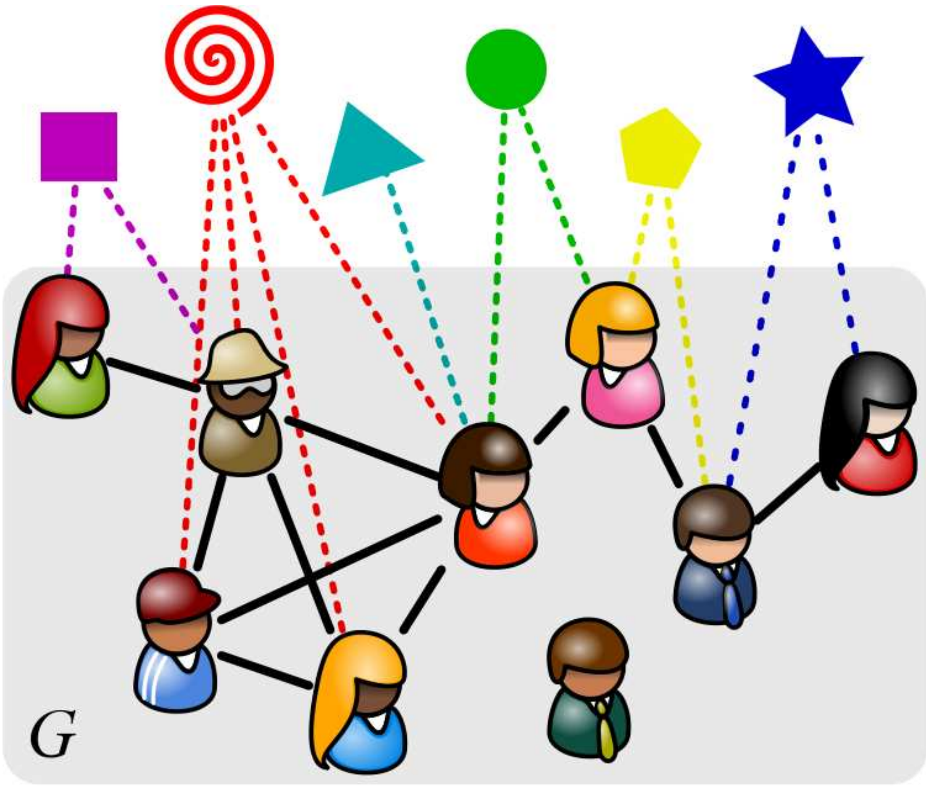
University of California, Irvine

buttsc@uci.edu

This work was supported in part by NSF award CMS-0624257 and ONR award N00014-08-1-1015.

The Problem of Complex Dependence

- Many human systems exhibit complex patterns of dependence
  - Nontrivial coupling among system elements
  - Particularly true within relational systems (i.e., social networks)
- A methodological and theoretical challenge
  - How to capture dependence without losing inferential tractability?
  - Not a new problem: also faced, e.g., by researchers in statistical physics



Challenge: Modeling Reality without Sacrificing Data

- How do we work with models which have non-trivial dependence?
- Can compare behavior of dependent-process models against stylized facts, but this has limits....
  - Not all models lead to clean/simple conditional or marginal relationships
  - Often impossible to disentangle nonlinearly interacting mechanisms on this basis
  - Very data inefficient: throws away much of the information content
  - Often need (very) large data sets to get sufficient power (which may not exist)
    - Collection of massive data sets often prohibitively costly
    - Many systems of interest *are* size-limited; studying only large systems leads to sampling bias
- Ideally, would like a framework which allows principled inference/model comparison without sacrificing (much) data

Our Focus: Stochastic Models for Social (and Other) Networks

- General problem: need to model graphs with varying properties
- Many *ad hoc* approaches:
  - Conditional uniform graphs (Erdös and Rényi, 1960)
  - Bernoulli/independent dyad models (Holland and Leinhardt, 1981)
  - Biased nets (Rapoport, 1949a;b; 1950)
  - Preferential attachment models (Simon, 1955; Barabási and Albert, 1999)
  - Geometric random graphs (Hoff et al., 2002)
  - Agent-based/behavioral models (Carley (1991); Hummon and Fararo (1995))
- A more general scheme: discrete exponential family models (ERGs)
  - General, powerful, leverages existing statistical theory (e.g., Barndorff-Nielsen (1978); Brown (1986); Strauss (1986))
  - (Fairly) well-developed simulation, inferential methods (e.g., Snijders (2002); Hunter and Handcock (2006))
- Today's focus – model parameterization

Basic Notation

- Assume $G = (V, E)$ to be the graph formed by edge set E on vertex set V
  - Here, we take $|V| = N$ to be fixed, and assume elements of V to be uniquely identified
  - If $E \subseteq \left\{\{v, v'\} : v, v' \in V\right\}$, G is said to be *undirected*; G is *directed* iff $E \subseteq \left\{(v, v') : v, v' \in V\right\}$
  - $\{v, v\}$ or (v, v) edges are known as *loops*; if G is defined per the above and contains no loops, G is said to be *simple*
    - Note that multiple edges are already banned, unless E is allowed to be a multiset
- Other useful bits
  - E may be random, in which case $G = (V, E)$ is a *random graph*
  - Adjacency matrix $\mathbf{Y} \in \{0, 1\}^{N \times N}$ (may also be random); for G random, will usually use notation $\mathbf{y}$ for adjacency matrix of realization g of G
  - $\mathbf{y}^+_{ij}$ is used to denote the matrix $\mathbf{y}$ with the i, j entry forced to 1; $\mathbf{y}^-_{ij}$ is the same matrix with the i, j entry forced to 0

Exponential Families for Random Graphs

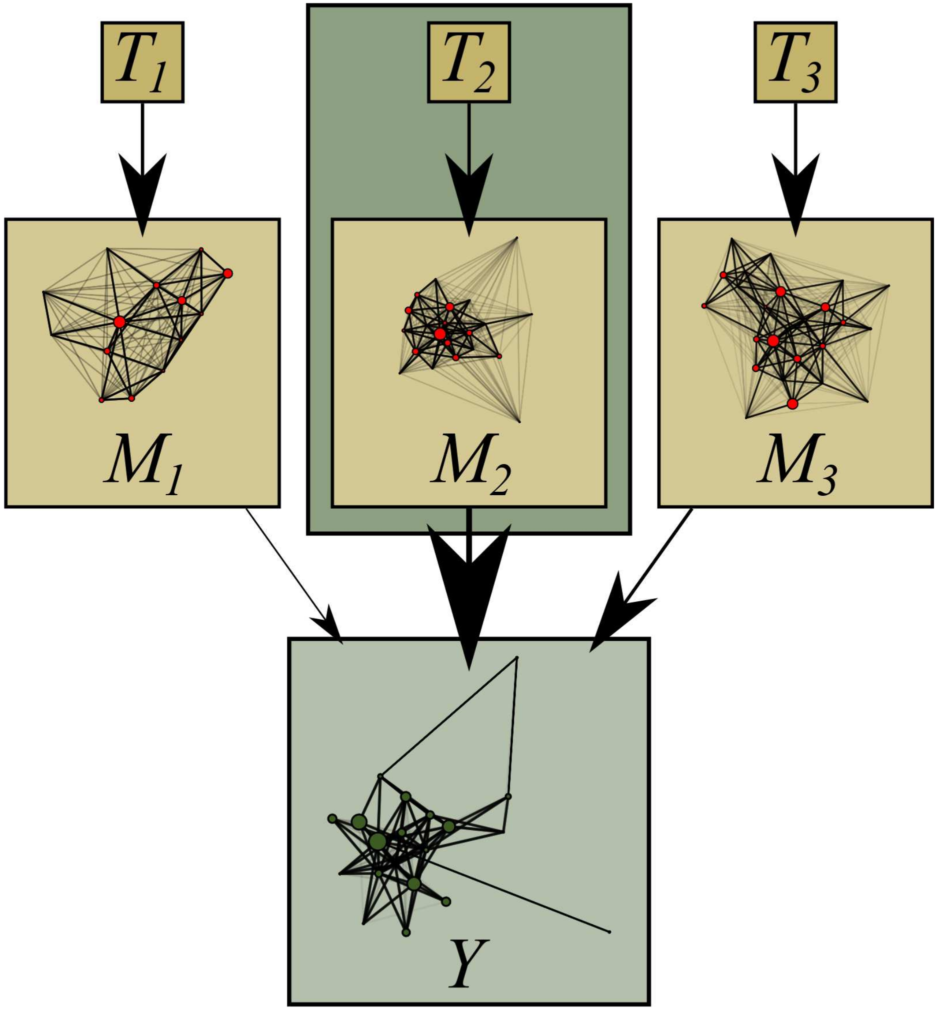
- For random graph G w/countable support $\mathcal{G}$, pmf is given in ERG form by

$$\Pr(G = g|\theta) = \frac{\exp\left(\theta^T \mathbf{t}(g)\right)}{\sum_{g' \in \mathcal{G}} \exp\left(\theta^T \mathbf{t}(g')\right)} I_{\mathcal{G}}(g) \tag{1}$$

- $\theta^T \mathbf{t}$: linear predictor
  - $\mathbf{t} : \mathcal{G} \to \mathbb{R}^m$: vector of sufficient statistics
  - $\theta \in \mathbb{R}^m$: vector of parameters
  - $\sum_{g' \in \mathcal{G}} \exp\left(\theta^T \mathbf{t}(g')\right)$: normalizing factor (aka partition function, Z)
- Intuition: ERG places more/less weight on structures with certain features, as determined by $\mathbf{t}$ and θ
  - Model is complete for pmfs on $\mathcal{G}$, few constraints on $\mathbf{t}$

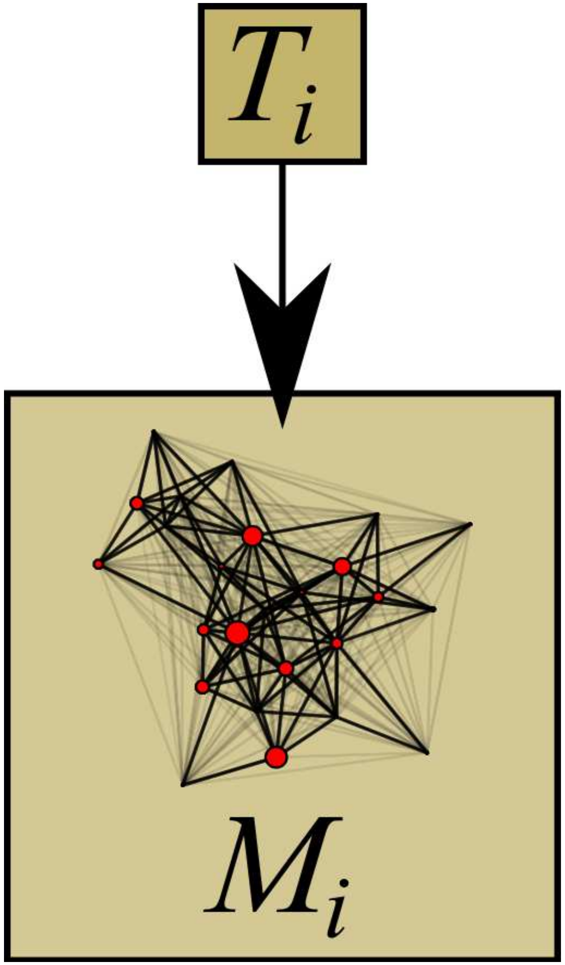
Inference with ERGs

- Important feature of ERGs is availability of inferential theory
  - Need to discriminate among competing theories
  - May need to assess quantitative variation in effect strengths, etc.
- Basic logic
  - Derive ERG parameterization from prior theory
  - Assess fit to observed data
  - Select model/interpret parameters
  - Update theory and/or seek low-order approximating models
  - Repeat as necessary



Parameterizing ERGs

- The ERG form is a way of representing distributions on $\mathcal{G}$, ***not*** a model in and of itself!
- Critical task: derive model statistics from prior theory
- Several approaches – here we introduce a new one....



A New Direction: Potential Games

- Most prior parameterization work has used *dependence hypotheses*
  - Define the conditions under which one relationship could affect another, and hope that this is sufficiently reductive
  - Complete agnosticism regarding underlying mechanisms – could be social dynamics, unobserved heterogeneity, or secret closet monsters
- A choice-theoretic alternative?
  - In some cases, reasonable to posit actors with some control over edges (e.g., out-ties)
  - Existing theory often suggests general form for utility
  - Reasonable behavioral models available (e.g., multinomial choice)
- The link between choice models and ERGs: *potential games*
  - Increasingly wide use in economics, engineering
  - Equilibrium behavior provides an alternative way to parameterize ERGs

Potential Games and Network Formation Games

- (Exact) Potential games (Monderer and Shapley, 1996)
  - Let X by a strategy set, u a vector utility functions, and V a set of players. Then (V, X, u) is said to be a *potential game* if $\exists\ \rho : X \mapsto \mathbb{R}$ such that, for all $i \in V$, $u_i\left(x'_i, x_{-i}\right) - u_i\left(x_i, x_{-i}\right) = \rho\left(x'_i, x_{-i}\right) - \rho\left(x_i, x_{-i}\right)$ for all $x, x' \in X$.
- Consider a simple family of *network formation games* (Jackson, 2006) on $\mathcal{Y}$:
  - Each i, j element of $\mathbf{Y}$ is controlled by a single player $k \in V$ with finite utility u_k; can choose $y_{ij} = 1$ or $y_{ij} = 0$ when given an "updating opportunity"
    - We will here assume that i controls $\mathbf{Y}_{i\cdot}$, but this is not necessary
  - Theorem: Let (i) $(V, \mathcal{Y}, u)$ in the above form a game with potential ρ; (ii) players choose actions via a logistic choice rule; and (iii) updating opportunities arise sequentially such that every (i, j) is selected with positive probability, and (i, j) is selected independently of the current state of $\mathbf{Y}$. Then $\mathbf{Y}$ forms a Markov chain with equilibrium distribution $\Pr(\mathbf{Y} = \mathbf{y}) \propto \exp(\rho(\mathbf{y}))$, in the limit of updating opportunities.
- One can thus obtain an ERG as the long-run behavior of a strategic process, and parameterize in terms of the hypothetical underlying utility functions

Proof of Potential Game Theorem

Assume an updating opportunity arises for y_{ij}, and assume that player k has control of y_{ij}. By the logistic choice assumption,

$$\Pr\left(\mathbf{Y} = \mathbf{y}_{ij}^{+} \,\middle|\, \mathbf{Y}_{ij}^{c} = \mathbf{y}_{ij}^{c}\right) = \frac{\exp\left(u_k\left(\mathbf{y}_{ij}^{+}\right)\right)}{\exp\left(u_k\left(\mathbf{y}_{ij}^{+}\right)\right) + \exp\left(u_k\left(\mathbf{y}_{ij}^{-}\right)\right)} \tag{2}$$

$$= \left[1 + \exp\left(u_k\left(\mathbf{y}_{ij}^{-}\right) - u_k\left(\mathbf{y}_{ij}^{+}\right)\right)\right]^{-1}. \tag{3}$$

Since $u, \mathcal{Y}$ form a potential game, $\exists\, \rho : \rho\left(\mathbf{y}_{ij}^{+}\right) - \rho\left(\mathbf{y}_{ij}^{-}\right) = u_k\left(\mathbf{y}_{ij}^{+}\right) - u_k\left(\mathbf{y}_{ij}^{-}\right) \;\forall\, k, (i,j), \mathbf{y}_{ij}^{c}$. Therefore, $\Pr\left(\mathbf{Y} = \mathbf{y}_{ij}^{+} \,\middle|\, \mathbf{Y}_{ij}^{c} = \mathbf{y}_{ij}^{c}\right) = \left[1 + \exp\left(\rho\left(\mathbf{y}_{ij}^{-}\right) - \rho\left(\mathbf{y}_{ij}^{+}\right)\right)\right]^{-1}$. Now assume that the updating opportunities for $\mathbf{Y}$ occur sequentially such that (i,j) is selected independently of $\mathbf{Y}$, with positive probability for all (i,j). Given arbitrary starting point $\mathbf{Y}^{(0)}$, denote the updated sequence of matrices by $\mathbf{Y}^{(0)}, \mathbf{Y}^{(1)}, \ldots$. This sequence clearly forms an irreducible and aperiodic Markov chain on $\mathcal{Y}$ (so long as ρ is finite); it is known that this chain is a Gibbs sampler on $\mathcal{Y}$ with equilibrium distribution $\Pr(\mathbf{Y} = \mathbf{y}) = \frac{\exp(\rho(\mathbf{y}))}{\sum_{\mathbf{y}' \in \mathcal{Y}} \exp(\rho(\mathbf{y}'))}$, which is an ERG with potential ρ. By the ergodic theorem, then $\mathbf{Y}^{(i)} \xrightarrow[i \to \infty]{} ERG(\rho(\mathbf{Y}))$. QED.

Some Potential Game Properties

- Game-theoretic properties
  - Local maxima of ρ correspond to Nash equilibria in pure strategies; global maxima of ρ correspond to stochastically stable Nash equilibria in pure strategies
    - At least one maximum must exist, since ρ is bounded above for any given θ
  - Fictitious play property; Nash equilibria compatible with best responses to mean strategy profile for population (interpreted as a mixed strategy)
- Implications for simulation, model behavior
  - Multiplying θ by a constant $\alpha \to \infty$ will drive the system to its SSNE
    - Likewise, best response dynamics (equivalent to conditional stepwise ascent) always leads to a NE
  - For degenerate models, "frozen" structures represent Nash equilibria in the associated potential game
    - Suggests a social interpretation of degeneracy in at least some cases: either correctly identifies robust social regimes, or points to incorrect preference structure

Building Potentials: Independent Edge Effects

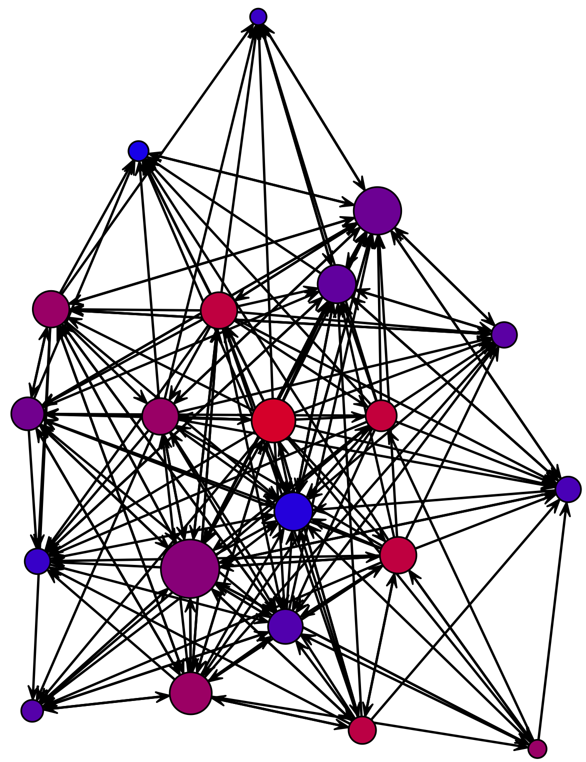
- General procedure
  - Identify utility for actor i
  - Determine difference in u_i for single edge change
  - Find ρ such that utility difference is equal to utility difference for all u_i
- Linear combinations of payoffs
  - If $u_i(\mathbf{y}) = \sum_j u_i^{(j)}(\mathbf{y})$, $\rho(\mathbf{y}) = \sum_j \rho_i^{(j)}(\mathbf{y})$
- Edge payoffs (homogeneous)
  - $u_i(\mathbf{y}) = \theta \sum_j y_{ij}$
  - $u_i\left(\mathbf{y}_{ij}^{+}\right) - u_i\left(\mathbf{y}_{ij}^{-}\right) = \theta$
  - $\rho(\mathbf{y}) = \theta \sum_i \sum_j y_{ij}$
  - Equivalence: p_1/Bernoulli density effect
- Edge payoffs (inhomogeneous)
  - $u_i(\mathbf{y}) = \theta_i \sum_j y_{ij}$
  - $u_i\left(\mathbf{y}_{ij}^{+}\right) - u_i\left(\mathbf{y}_{ij}^{-}\right) = \theta_i$
  - $\rho(\mathbf{y}) = \sum_i \theta_i \sum_j y_{ij}$
  - Equivalence: p_1 expansiveness effect
- Edge covariate payoffs
  - $u_i(\mathbf{y}) = \theta \sum_j y_{ij} x_{ij}$
  - $u_i\left(\mathbf{y}_{ij}^{+}\right) - u_i\left(\mathbf{y}_{ij}^{-}\right) = \theta x_{ij}$
  - $\rho(\mathbf{y}) = \theta \sum_i \sum_j y_{ij} x_{ij}$
  - Equivalence: Edgewise covariate effects (netlogit)

Building Potentials: Dependent Edge Effects

- Reciprocity payoffs
  - $u_i(\mathbf{y}) = \theta \sum_j y_{ij} y_{ji}$
  - $u_i\left(\mathbf{y}_{ij}^{+}\right) - u_i\left(\mathbf{y}_{ij}^{-}\right) = \theta y_{ji}$
  - $\rho(\mathbf{y}) = \theta \sum_i \sum_{j<i} y_{ij} y_{ji}$
  - Equivalence: p_1 reciprocity effect
- 3-Cycle payoffs
  - $u_i(\mathbf{y}) = \theta \sum_{j \neq i} \sum_{k \neq i,j} y_{ij} y_{jk} y_{ki}$
  - $u_i\left(\mathbf{y}_{ij}^{+}\right) - u_i\left(\mathbf{y}_{ij}^{-}\right) = \theta \sum_{k \neq i,j} y_{jk} y_{ki}$
  - $\rho(\mathbf{y}) = \frac{\theta}{3} \sum_i \sum_{j \neq i} \sum_{k \neq i,j} y_{ij} y_{jk} y_{ki}$
  - Equivalence: Cyclic triple effect
- Transitive completion payoffs
  - $u_i(\mathbf{y}) = \theta \sum_{j \neq i} \sum_{k \neq i,j} \left[\begin{array}{c} y_{ij} y_{ki} y_{kj} + y_{ij} y_{ik} y_{jk} \\ + y_{ij} y_{ik} y_{kj} \end{array} \right]$
  - $u_i\left(\mathbf{y}_{ij}^{+}\right) - u_i\left(\mathbf{y}_{ij}^{-}\right) = \theta \sum_{k \neq i,j} \left[y_{ki} y_{kj} + y_{ik} y_{jk} + y_{ik} y_{kj} \right]$
  - $\rho(\mathbf{y}) = \theta \sum_i \sum_{j \neq i} \sum_{k \neq i,j} y_{ij} y_{ik} y_{kj}$
  - Equivalence: Transitive triple effect

Empirical Example: Advice-Seeking Among Managers

- Sample empirical application from Krackhardt (1987): self-reported advice-seeking among 21 managers in a high-tech firm
  - Additional covariates: friendship, authority (reporting)
- Demonstration: selection of potential behavioral mechanisms via ERGs
  - Models parameterized using utility components
  - Model parameters estimated using maximum likelihood (Geyer-Thompson)
  - Model selection via AIC



Advice-Seeking ERG – Model Comparison

- First cut: models with independent dyads:

	Deviance	Model df	AIC	Rank
Edges	578.43	1	580.43	7
Edges+Sender	441.12	21	483.12	4
Edges+Covar	548.15	3	554.15	5
Edges+Recip	577.79	2	581.79	8
Edges+Sender+Covar	385.88	23	431.88	2
Edges+Sender+Recip	405.38	22	449.38	3
Edges+Covar+Recip	547.82	4	555.82	6
Edges+Sender+Covar+Recip	**378.95**	**24**	**426.95**	**1**

- Elaboration: models with triadic dependence:

	Deviance	Model df	AIC	Rank
Edges+Sender+Covar+Recip	378.95	24	426.95	4
Edges+Sender+Covar+Recip+CycTriple	361.61	25	411.61	2
Edges+Sender+Covar+Recip+TransTriple	368.81	25	418.81	3
Edges+Sender+Covar+Recip+CycTriple+TransTriple	**358.73**	**26**	**410.73**	**1**

- Verdict: data supplies evidence for heterogeneous edge formation preferences (w/covariates), with additional effects for reciprocated, cycle-completing, and transitive-completing edges.

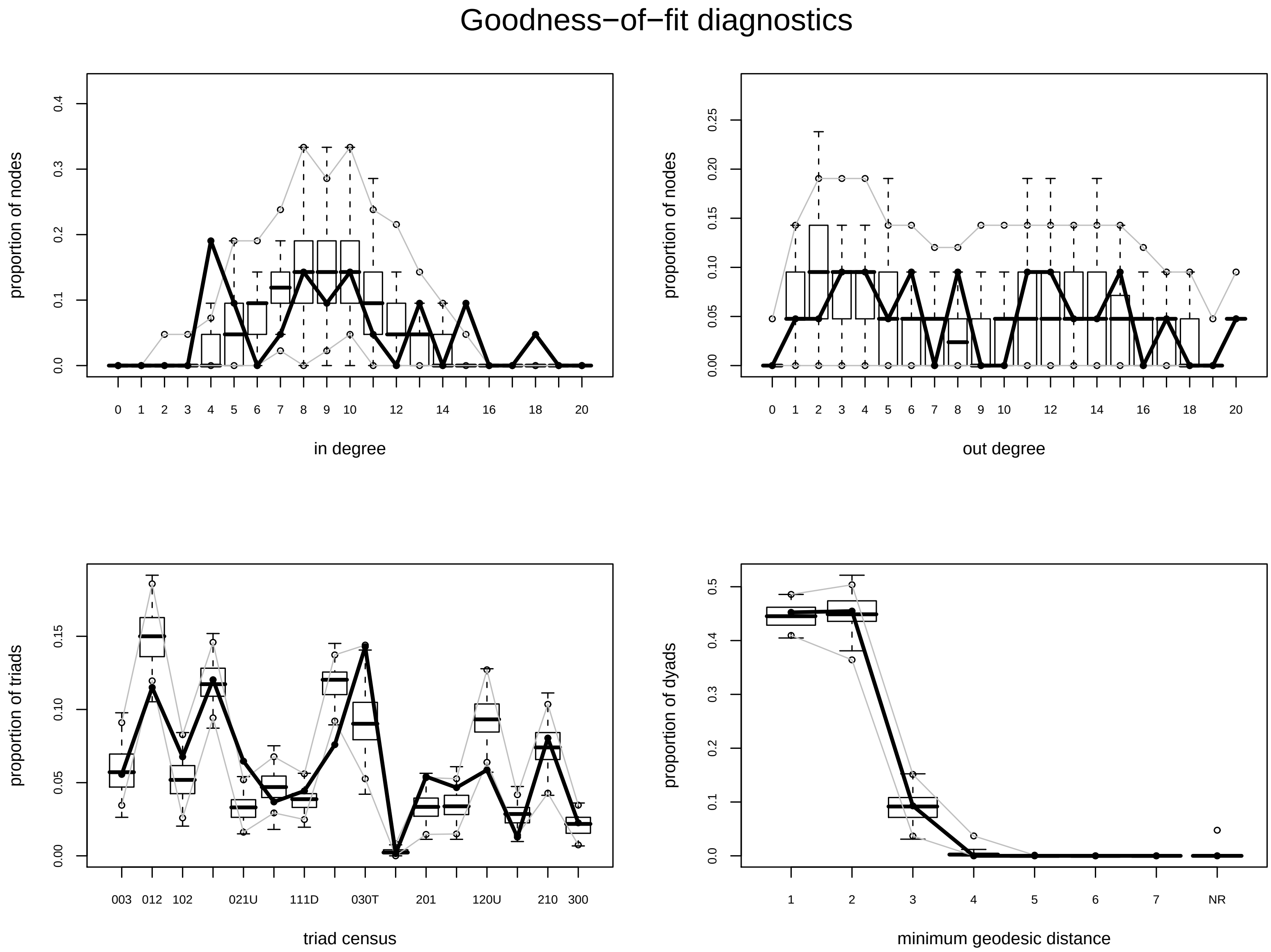
Advice-Seeking ERG – AIC Selected Model

Effect	$\hat{\theta}$	s.e.	$\Pr(>\|Z\|)$		Effect	$\hat{\theta}$	s.e.	$\Pr(>\|Z\|)$	
Edges	**−1.022**	**0.137**	**0.0000**	∗∗∗	**Sender14**	**−1.513**	**0.231**	**0.0000**	∗∗∗
Sender2	**−2.039**	**0.637**	**0.0014**	∗∗	**Sender15**	**16.605**	**0.336**	**0.0000**	∗∗∗
Sender3	0.690	0.466	0.1382		**Sender16**	**−1.472**	**0.232**	**0.0000**	∗∗∗
Sender4	−0.049	0.441	0.9112		**Sender17**	**−2.548**	**0.197**	**0.0000**	∗∗∗
Sender5	0.355	0.495	0.4734		**Sender18**	**1.383**	**0.214**	**0.0000**	∗∗∗
Sender6	**−4.654**	**1.540**	**0.0025**	∗∗	**Sender19**	**−0.601**	**0.190**	**0.0016**	∗∗
Sender7	−0.108	0.375	0.7726		Sender20	0.136	0.161	0.3986	
Sender8	−0.449	0.479	0.3486		Sender21	0.105	0.210	0.6157	
Sender9	0.393	0.496	0.4281		**Reciprocity**	**0.885**	**0.081**	**0.0000**	∗∗∗
Sender10	0.023	0.555	0.9662		**Edgecov (Reporting)**	**5.178**	**0.947**	**0.0000**	∗∗∗
Sender11	**−2.864**	**0.721**	**0.0001**	∗∗∗	**Edgecov (Friendship)**	**1.642**	**0.132**	**0.0000**	∗∗∗
Sender12	**−2.736**	**0.331**	**0.0000**	∗∗∗	**CycTriple**	**−0.216**	**0.013**	**0.0000**	∗∗∗
Sender13	**−0.986**	**0.194**	**0.0000**	∗∗∗	**TransTriple**	**0.090**	**0.003**	**0.0000**	∗∗∗

Null Dev 582.24; Res Dev 358.73 on 394 df

- Some observations...
  - Arbitrary edges are costly for most actors
  - Edges to friends and superiors are “cheaper” (or even positive payoff)
  - Reciprocating edges, edges with transitive completion are cheaper...
  - ...but edges which create (in)cycles are more expensive; a sign of hierarchy?

Model Adequacy Check



Where Would One Go Next?

- Model refinement
  - Goodness-of-fit is not unreasonable, but some improvement is clearly possible
  - Could refine existing model (e.g., by adding covariates) or propose more alternatives
- Replication on new cases
  - Given a smaller set of candidates, would replicate on new organizations
  - May lead to further refinement/reformulation
- Simplification
  - Given a model family that works well, can it be simplified w/out losing too much?
  - Seek the smallest model which captures essential properties of optimal model; general behavior can then be characterized (hopefully)

Summary

- Models for non-trivial networks pose non-trivial problems
  - Many ways to describe dependence among elements
  - Once one leaves simple cases, not always clear where to begin
- Potential games for ERG parameterization
  - Allow us to derive random cross-sectional behavior from strategic interaction
  - Provide sufficient conditions for ERG parameters to be interpreted in terms of preferences
  - Allows for testing of competing behavioral models (assuming scope conditions are met!)
- Approach seems promising, but many questions remain
  - Can we characterize utilities which lead to identifiable models?
  - How can we leverage other properties of potential games?

1 References

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 206:509–512.

Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory*. John Wiley and Sons, New York.

Brown, L. D. (1986). *Fundamentals of Statistical Exponential Families, with Applications in Statistical Decision Theory*. Institute of Mathematical Statistics, Hayward, CA.

Carley, K. M. (1991). A theory of group stability. *American Sociological Review*, 56(3):331–354.

Erdös, P. and Rényi, A. (1960). On the evolution of random graphs. *Public Mathematical Institute of Hungary Academy of Sciences*, 5:17–61.

Harary, F. (1953). On the notion of balance of a signed graph. *Michigan Mathematical Journal*, 3:37–41.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. John Wiley and Sons, New York.

Hoff, P. D., Raftery, A. E., and Handcock, M. S. (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association*, 97(460):1090–1098.

Holland, P. W. and Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs (with discussion). *Journal of the American Statistical Association*, 76(373):33–50.

Hummon, N. P. and Fararo, T. J. (1995). Assessing hierarchy and balance in dynamic network models. *Journal of Mathematical Sociology*, 20:145–159.

Hunter, D. R. and Handcock, M. S. (2006). Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics*, 15:565–583.

Jackson, M. (2006). A survey of models of network formation: Stability and efficiency. In Demange, G. and Wooders, M., editors, *Group Formation Economics: Networks, Clubs, and Coalitions*. Cambridge University Press, Cambridge.

Krackhardt, D. (1987). Cognitive social structures. *Social Networks*, 9(2):109–134.

Monderer, D. and Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, 14:124–143.

Rapoport, A. (1949a). Outline of a probabilistic approach to animal sociology I. *Bulletin of Mathematical Biophysics*, 11:183–196.

Rapoport, A. (1949b). Outline of a probabilistic approach to animal sociology II. *Bulletin of Mathematical Biophysics*, 11:273–281.

Rapoport, A. (1950). Outline of a probabilistic approach to animal sociology III. *Bulletin of Mathematical Biophysics*, 12:7–17.

Simon, H. A. (1955). On a class of skew distribution functions. *Biometrika*, 42:425–440.

Snijders, T. A. B. (2002). Markov Chain Monte Carlo estimation of exponential random graph models. *Journal of Social Structure*, 3(2).

Strauss, D. (1986). On a General Class of Models for Interaction. *SIAM Review*, 28(4):513–527.