Micro-Level Interpretation of Exponential Random Graph Models with Application to Estuary Networks

Bruce A. Desmarais
Department of Political Science
University of Massachusetts Amherst
desmarais@polsci.umass.edu

Skyler J. Cranmer
Department of Political Science
University of North Carolina at Chapel Hill
skyler@unc.edu

Abstract

A comprehensive understanding of the relational processes that constitute policy networks requires both an analysis of exogenous (i.e. covariates) determinants of the link structures in the network and a study of how the relationships that form the network depend upon each other, the endogenous determinants of the network. The Exponential Random Graph Model (ERGM) is an increasingly popular method for the statistical analysis of networks that can be used for both of these inferential tasks. Conventional interpretation of ERGM results is conducted at the network level, such that effects are related to overall frequencies of network structures (e.g., the number of closed triangles in a network). This limits the utility of the ERGM because there is often interest, particularly in the social sciences, in network dynamics at the actor or relationship levels. We present a comprehensive framework for interpretation of the ERGM at these lower levels, which casts network formation as block-wise updating of a network. These blocks can represent, for example, each potential link, each dyad, and the out or in-going ties of each actor. We contrast this interpretive framework with the popular actor-oriented model of network dynamics. The alternative models we discuss and the interpretation methods we propose are illustrated on estuary network data introduced in Berardo and Scholz (2010).

1 Introduction

The study of collaboration, conflict, and influence among relevant actors is essential to building a comprehensive understanding of the policymaking process. The set of relationships among policymaking entities forms a complex network whereby the organizations represent nodes (also called vertices) and the relationships between them represent edges (also called links or ties). A characteristic of network data that is both scientifically interesting and statistically challenging is that the probability of an edge forming between any two nodes in the network often depends upon the structure of the edges throughout the rest of the network (Cranmer and Desmarais 2011). As such, appropriate methods for the analysis of network data permit researchers to test theories about the dependence of edge formation upon covariates that are exogenous to the network as well as endogenous interdependence among the components of the network.

One method that is flexible enough to accommodate a wide range of theories of network generation is the exponential random graph model (ERGM) (Holland and Leinhardt 1981; Wasserman and Pattison 1996; Cranmer and Desmarais 2011). The ERGM models a network as a single multivariate observation in which the components of the network depend on exogenous covariates as well as endogenous dependencies among the nodes and edges. An extension proposed by Hanneke, Fu and Xing (2010), the temporal ERGM (TERGM), extends the ERGM such that it may analyze a discrete time series of network observations. Taken together, the ERGM/TERGM provides a powerful analytical framework that can be applied almost ubiquitously to political and policy networks.[1]

Despite its power, the ERG class of models has not been used extensively by political and policy scientists, largely because of a focus in those disciplines on actor- and relationship-level

[1] The ERGM/TERGM framework we discuss here is limited by an inability to accommodate networks with valued edges (i.e. edges must be binary, either present or absent), but recent extensions by Wyatt, Choudhury and Bilmes (2010), Desmarais and Cranmer (2011) and Krivitsky (2011) may provide the key to overcoming this limitation.

effects and a common misconception that ERGMs can only provide inferences at the network level. The roots of this misconception are understandable: the ERGM's treatment of the network as a single, multivariate, observation and its mathematical derivation at the network level (e.g. (Park and Newman 2004)) naturally lead to interpretation at the network level. Typically, researchers whose substantive interest lie in actor- and relationship-level effects have shied away from the ERGM in favor of techniques such as the actor-oriented dynamic network (Snijders, van de Bunt and Steglich 2010); widely perceived as more appropriate for drawing actor-level inference. The standard approach to inference and hypothesis testing with the ERGM can be summarized by a three step process. First, develop network statistics that represent theorized network effects (e.g., reciprocity, transitivity, homophily with respect to some node covariate). Second, estimate an ERGM or TERGM parameterized with the respective statistics. Third, use hypothesis tests on the corresponding parameter estimates to assess whether the given relational tendencies are significant features of the process that generated the network (see, e.g., Lazer et al. (2010), Cranmer, Desmarais and Menninga (Forthcoming), Goodreau, Kitts and Morris (2009)). Though parsimonious and useful for network-level inference, this approach to interpretation only taps the surface of the information available from ERGM estimates.

We develop and explicate a framework for interpreting the ERGM at the actor (node) and relationship (edge) levels; a framework that, we hope, will increase the applied utility of the ERGM to political and policy scientists. We begin by considering how the results from an ERGM describe a complete joint distribution of all of the components in the network. This means that, with a little arithmetic, we can derive the conditional distribution of any block in the network given the rest of the network. For instance, we can derive the probability of a set of out-going edges from a particular node given the rest of the rest of the network. We show how, when combined with insight from the theory of Gibbs sampling, the form of the ERGM implies a block-wise dynamic updating process, which can be interpreted, for example, at the node or dyad (link) level. We contrast this dynamic interpretation of the ERGM with the assumptions underlying the popular actor-oriented model of network dynamics, which we refer to by the acronym for its software implementation – SIENA. Through a replication analysis of Berardo and Scholz's (2010) study of estuary networks, we illustrate these interpretation methods as well as a comparison of ERGM and TERGM to SIENA for the purposes of fitting dynamic network models.

2 Exponential Random Graph Models

Here we briefly review the form of the ERGM before considering its dynamic application.[2] Let Y represent a network of n actors, which we treat interchangeably as notation for the adjacency matrix, such that $Y_{ij} = 1$ if there is an edge (i.e., tie, link) from node (i.e., actor, vertex) i to node j and $Y_{ij} = 0$ otherwise. Note that $Y_{ij} = Y_{ji}$ if the relationship under study is undirected, meaning there is no sender and receiver of the relation, but rather two symmetric participants in a relationship. Let $\boldsymbol{\mathcal{Y}}$ be the set of all networks that have the same direction and number of nodes as Y. The set $\boldsymbol{\mathcal{Y}}$, commonly referred to as the 'support' of Y, contains the empty network, with all elements of Y equal to 0, the full network, with all elements of Y equal to 1, and every possible permutation in between. Let $\Gamma(Y)$ be a network statistic, which takes as arguments a network adjacency matrix and other fixed information (e.g., the value of covariates), and returns a scalar-valued network statistic. Lastly, let θ be a real-valued parameter of the ERGM. The ERGM has as many parameters as there are Γ's in the specification.

Specification of the ERGM proceeds by developing and selecting network statistics that represent relational tendencies hypothesized to be components of the process that generated Y. These statistics take two general forms: statistics that capture endogenous dependencies

[2] See Cranmer and Desmarais (2011) for a more comprehensive review.

among the nodes and edges in the network, and statistics that include node or edge level covariates. All statistics are calculated as sums over subgraphs (i.e. the sum of relationships computed on a subset of the network), but their forms can vary substantially. Analogous to the selection of independent variables in regression modeling, the development and specification of network statistics should be deduced from theory and informed by the relevant literature.

Endogenous dependencies often include the processes of reciprocity (in directed networks), transitivity (i.e., the friend of a friend is a friend), popularity, and the like. Statistics to capture exogenous dependencies can take on many forms depending on what type of dependency they are designed to capture. For example, a common property of directed networks is that edge-formation is reciprocal, such that when $Y_{ij} = 1$ there is a higher probability that $Y_{ji} = 1$ than when $Y_{ij} = 0$. Reciprocity, or mutuality, is commonly included in ERGM specifications via the count of the number of mutual dyads in which $Y_{ij} = Y_{ji} = 1$ (Handcock et al. 2010). This reciprocity measure is

$$\Gamma_R(Y) = \sum_{i<j} Y_{ij} Y_{ji}. \tag{1}$$

A full review of every statistic that could be specified to capture an endogenous dependence is beyond the scope of the present discussion, but readers are referred to Snijders et al. (2006) and Handcock et al. (2010) for more extensive discussions.

Node and/or edge-level covariates may also be included with a simple and general network statistic. Let the covariate (X) be given as a matrix such that the element X_{ij} is the covariate value that corresponds to the ij element of Y. For example, X may be a matrix of dyadic covariate values capturing the relationships between each set of actors. The statistic used to include a covariate in the ERGM is

$$\Gamma_X(Y, X) = \sum_{i \neq j} X_{ij} Y_{ij}. \tag{2}$$

Though this may appear to limit covariates to those at the dyad-level because there is a value for each relationship in the network, a node-level covariate can also be represented in matrix form. For example, if interest is in measuring a sender effect of a given node-level variable (Z), with n values, then $X_{ij} = Z_i$.

The ERGM combines the network statistics and parameters to construct the probability of observing the network Y relative to the rest of the networks in $\boldsymbol{\mathcal{Y}}$. Specifically, the probability of observing Y is proportional to $\exp(\sum_{j=1}^{k} \theta_j \Gamma_j(Y))$. This means that the direction of the effect of the j^{th} statistic on the probability of observing an instance of the network is equal to the direction of the j^{th} parameter. For instance, if the parameter corresponding to the reciprocity/mutuality statistic is positive, then networks exhibiting a high degree of mutuality among edges will be more likely than those with low mutuality, all else equal. This proportionality is normalized to give a proper probability distribution for the network. The probability of observing an instance of the network given by an ERGM is

$$P(Y) = \frac{\exp(\sum_{j=1}^{k} \theta_j \Gamma_j(Y))}{\sum_{Y^* \in \boldsymbol{\mathcal{Y}}} \exp(\sum_{j=1}^{k} \theta_j \Gamma_j(Y^*))}. \tag{3}$$

The normalization constant in the denominator requires evaluation at every network in $\boldsymbol{\mathcal{Y}}$, the set of all possible permutations of Y, and thus can be a challenge for estimation in all but small networks. However, this this quantity can be approximated by simulation and, as such, only represents a barrier to estimation in very large networks where it cannot be easily approximated; see Desmarais and Cranmer (Forthcoming) for a discussion of this problem and potential solutions.

The ERGM makes two straightforward assumptions about the distribution from which the observed network or series of networks was drawn (Park and Newman 2004). First, given a set of statistics defined on the network, there is equal probability of observing any two networks that have the same value of those statistics. Second, the observed network exhibits the average value of those statistics over the possible networks that could have been observed. The first assumption implies that the model is completely and correctly specified. The second assumption serves to identify the parameters for estimation and is, in practice, no different from the assumption that the average relationships in a dataset are representative of the population. Notably, the second assumption is implicit to modeling with the linear model (i.e. OLS) and many variants of the generalized linear model (e.g., logit, Poisson regression).[3]

The flexibility of the ERGM approach lies in the fact that we have derived the probability distribution at the network level directly, with very few restrictions on the Γ that can be used to customize an ERGM analysis. In the next section, we develop a straightforward approach to manipulating ERGM results such that interpretation of network formation is possible at multiple, indeed arbitrary, subnetwork levels.

3 The ERGM as a Dynamic Network Formation Model

Conventional network-level interpretation of ERGM results may be insufficient for many political science and policy scholars because theories tend to be deduced from micro foundations

[3] This result follows from the multivariate canonical exponential family form of the generalized linear model. The generalized linear model (of which the linear model is a special case), parameterized with the canonical link function, is a multivariate canonical exponential family model (Nelder and Wedderburn 1972). Estimation of the parameter of the canonical multivariate exponential family by maximum likelihood is equivalent to finding the parameter values that imply expected sufficient statistics equal to the values observed in the data (Lehmann 1983). Thus, the assumption that the expected values of the sufficient statistics are equivalent to the values observed in the data is implicit to any empirical analysis that uses maximum likelihood estimation of the generalized linear model or least squares estimation of the linear model.

(e.g., Rapaport, Levi-Faur and Miodownik (2009); Berardo and Scholz (2010); Oakerson and Parks (2011)). The implication is that expectations for aggregate-level behavior/outcomes are derived from assumptions about the interacting preferences and behaviors of the component parts of the system. Network-level interpretations of the process under study will often not be able to directly address the theoretical process of interest and thus limit the utility of the ERGM to a broad class of researchers in the social sciences. To address this problem, we show how ERGM results can be manipulated to permit interpretation at any level of analysis.

3.1 Dynamic Generation of Exponential Random Graphs

In the social sciences, longitudinal network data are typically gathered at several discrete points in time. In most such datasets, the time intervals at which the data are gathered, even if regular, are 'coarse,' meaning that a substantively significant amount of time has passed between waves of data collection. In terms of measurement, the data appear as several cross-sectional recordings of a given network, tacitly implying that all edges within a given period are measured or change simultaneously. However, relational data are typically generated dynamically; relationships change sequentially in continuous time and changes are not typically clustered around the times of data collection. Moreover, it rarely makes theoretical sense to conceptualize edges as being dependent upon each other *and* generated simultaneously because dependence arises when individual components of the network are generated conditional upon the rest of the network (i.e., in sequence). Despite the fact that conventional interpretation of ERGM results focuses on the static, network-level, implications of the model, there are many sequential micro-level network generating processes that are consistent with a given ERGM model.

We draw upon the theory of Gibbs sampling (Geman and Geman 1984) to describe a useful class of sequential network generating processes that are consistent with any given ERGM model.[4] Let the adjacency matrix representing the network be partitioned into m disjoint groups or blocks. These blocks could be (1) each potential edge in the network, (2) each dyad in the network, (3) blocks of outgoing edges from each node in the network, (4) blocks of incoming edges to each node in the network, or any other conceivable partition of the network. Because we believe the actual data generating process to be sequential, we may assume there is some probabilistic process that determines an updating sequence for the m blocks of the partition ($S(m)$). For example, if there are three blocks in the partition, $S(m)$ probabilistically generates sequences from the possibilities in $\{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\}$.[5] The sequential updating process is analogous to a block Gibbs sampler, which produces sequential updates of several parameters (Gill 2007), because the conditional distributions of the blocks are properly deduced from the ERGM and the description of $S(m)$ is consistent with a stationary sequence. It follows then, given that the network has gone through a large number of updating cycles according to sequences drawn from $S(m)$ and ERGM-based conditional distributions, that Y will have an ERGM distribution. In other words, an ERGM appropriately describes a network that has been sequentially updated by probabilistically generating one block conditional upon the other blocks, whereby the 'conditioning' is based on conditional distributions derived from the ERGM.

In order to implement interpretation based on sequential generation, we need to compute the conditional probability of one block of the network given the rest of the network. This

[4] We are careful in our use of the phrase "consistent with." As with any statistical model, there are numerous processes that are consistent with the model. Take for instance, a simple linear regression model. The model implies a bivariate normal distribution, which holds if x causes y, y causes x or some mix of both. Except under strictly controlled experimental conditions, x is *assumed* to be exogenous. It is not possible for a model to imply or identify one and only one generating process, even if it is derived with one process in mind. Thus, it is the role of the theory and the researcher to determine which sequential process is most likely or interesting in a given application.

[5] When we say probabilistically, note that we do not imply a *uniform random* process. It is possible that some sequences are more prevalent than others.

computational process can be stated generally. Let $\boldsymbol{Y}_b$ be the b^{th} block of the network, $\boldsymbol{\mathcal{Y}}_b$ be the set of possible permutations of $\boldsymbol{Y}_b$, and $\boldsymbol{Y}_{-b}$ be the rest of the network not represented by the b^{th} block. The probability of any given block ($\boldsymbol{Y}_b$) given the rest of the network ($\boldsymbol{Y}_{-b}$) is

$$P(\boldsymbol{Y}_b|\boldsymbol{Y}_{-b}) = \frac{\exp(\sum_{j=1}^{k} \theta_j \Gamma_j(\boldsymbol{Y}_b \cap \boldsymbol{Y}_{-b}))}{\sum_{\boldsymbol{Y}_b^* \in \boldsymbol{\mathcal{Y}}_b} \exp(\sum_{j=1}^{k} \theta_j \Gamma_j(\boldsymbol{Y}_b^* \cap \boldsymbol{Y}_{-b}))}, \quad (4)$$

where the notation $\boldsymbol{Y}_b^* \cap \boldsymbol{Y}_{-b}$ stands for the complete network created by holding $\boldsymbol{Y}_{-b}$ constant and inserting $\boldsymbol{Y}_b^*$ into the b^{th} block of Y. The conditional probability of a block can be used to compute any number of conditional quantities, such as conditional means, variances, counts and correlations. The execution of this process and the computation of such quantities is made easy by our companion software described in the appendix.

3.2 Dynamic, Actor-Oriented Models of Networks

The ERGM is not the only well-developed method for modeling dynamic network formation. The actor-oriented model of network formation (Snijders, van de Bunt and Steglich 2010), commonly referred to as SIENA after the acronym of its software implementation, is a popular model among researchers interested in node-level effects. The SIENA model offers a dynamic approach to the study of networks with a focus and natural interpretation at the node level.[6]

The SIENA model assumes a very specific process by which the network is dynamically updated. SIENA begins by assuming that edges in the network change one at a time. An actor is selected according to an estimated rate of change function and the selected actor can change, at most, one out-going edge. Formally, let $Y^{(ij)}$ represent the network in which the ij^{th} element of the adjacency matrix is toggled (i.e. an edge is formed where none

[6]There are actually several options implemented for the process assumed by SIENA, but we limit our attention to the most common actor-oriented process and the one used by Berardo and Scholz (2010), whose replicated work we consider below.

existed, or an existing edge is dissolved). In SIENA, the probability that actor i, when the opportunity arises, changes its relationship with j is proportional to $\exp(\sum_{h=1}^{k}\theta_h\Gamma_{ih}(Y^{(ij)}))$, where $\sum_{h=1}^{k}\theta_h\Gamma_{ih}(Y^{(ij)})$ is referred to as the "objective function" (Snijders, van de Bunt and Steglich 2010) and includes parameters θ and network statistics $\Gamma()$, much like the ERGM. In other words, the objective function in a SIENA model is a linear combination of network statistics that change with actor i's out-going edges, weighted by real-valued parameters.[7]

The network formation processes in SIENA and the dynamically interpreted ERGM are very similar: in each model, changes in network structure favor, by construction, those that contribute to the linear combination of network statistics. The two models take different approaches to modeling the network generation process insofar as the SIENA model has parameters that fit the rate and sequence of decision-making, but the ERGM does not. Specifically, there is an additional equation in the SIENA model that determines the rate at which actors consider changing one of their out-going edges between observations of the network. These rate functions can be actor-specific or apply to the network as a whole. Whether this additional equation is warranted is application-specific. If theory suggests a change process similar to the one assumed by SIENA, then, at least initially, the additional structure is warranted. If there is no theoretical basis upon which to favor SIENA's conception of network change over the block-updating process consistent with ERGM, then adding the additional equation to the model may constitute over-fitting the data. As we illustrate below, it is both possible and advisable to compare multiple models of network dynamics.

[7]It is important to note that this objective function differs substantially from a conventional rational choice conception of a utility function. In the SIENA objective function, every node in the network shares the objective of increasing network statistics that correspond to positive parameter values and decreasing network statistics that correspond to negative parameter values. If interpreted as a utility function, this implies that every node in the network has the *same* utility function.

4 Empirical Application: Reanalysis of Berardo and Scholz (2010)

To illustrate the differences between the dynamically interpreted ERGM/TERGM and SIENA, we re-analyze data gathered by Schneider et al. (2003) and analyzed using the SIENA model by Berardo and Scholz (2010). The data measure relationships between 194 agencies involved in forming, implementing, and enforcing water usage and environmental policies in 10 estuaries. The data were gathered by surveying experts in each agency involved and asking them to identify three individuals, and the agencies they work for, on whom they have most heavily relied in the last year. Data were gathered in two waves: 1999 and 2001. The network is constructed by drawing a directed edge from the agency of the respondent to the agency of the reported contact. The networks for the 10 estuaries were represented in a single adjacency matrix using structural zeros to capture the fact that there are no cross-estuary edges. The result is two adjacency matrices to subject to statistical network analysis. For further details on data gathering and network coding, see Schneider et al. (2003) and Berardo and Scholz (2010).

Before proceeding our empirical analysis, which contrasts the dynamic interpretation of the ERGM to SIENA using the estuary data, we briefly consider three criticisms that Berardo and Scholz (2010) levied against the ERGM while defending their choice to use SIENA. First, Berardo and Scholz (2010, p. 638) criticized the ERGM as unable to account for the history of the network, whereas SIENA can control for previous realizations of the network. While true that the standard form of the ERGM, as presented in equation (3), addresses a single network, the temporal extension of the ERGM (the TERGM), originally proposed by Hanneke, Fu and Xing (2010) and now fully developed through subsequent studies (see e.g. Cranmer and Desmarais (2011), Desmarais and Cranmer (2010), and Desmarais and Cranmer (Forthcoming)), can easily account for previous realizations of the network. In a

TERGM, the network statistics (Γ) may include functions of earlier networks (see Cranmer and Desmarais (2011), Cranmer, Desmarais and Menninga (Forthcoming), and Desmarais and Cranmer (Accepted) for applied examples).

Second, Berardo and Scholz (2010, p. 639) argue that the ERGM assumes that each relationship maximizes the utility of the relevant actor given the rest of the network. The implication is that the ERGM may carry a, sometimes unwanted, assumption that the edges in the network represent a static equilibrium given the actors' utility functions. However, both the ERGM and SIENA are probabilistic models that assume networks are generated in proportion to the frequency of a set of network statistics selected by the researcher. In terms of stationarity assumptions, the dynamically interpreted ERGM assumes the network under study is drawn from the stationary distribution. In contrast, SIENA assumes the network *changes* according to a stationary process. Neither assumption is more or less warranted *a priori*, and the difference is analogous to time series analysis with and without a unit root (Williams 1992). As we illustrate below, ultimately, model fit diagnostics can be used to assess which assumption is more appropriate.

Finally, Berardo and Scholz (2010, p. 640) point out that the ERGM may seem to be limited compared to SIENA because the SIENA model includes an optional "behavior" component that can model the coevolution of a network and an actor attribute. While true that the ERGM does not include such a feature, it is also true that SIENA cannot model the simultaneous generation of the network and behavior. The SIENA behavior function conditions behavioral change on extant network structure and network change on extant behavior (Snijders, van de Bunt and Steglich 2010). As such, there is no contemporaneous endogeneity (e.g., residual correlation) between the network and the behavior component of SIENA models. As we illustrate below, a similar "behavior" model can be fit without SIENA by simply regressing node behavior at time t on node-level network statistics measured at $t-1$.

In light of the fact that recent advances in ERGM technology make the *a priori* selection of SIENA over the ERGM an ambiguous choice, we reconsider the analysis of Berardo and Scholz (2010) using both methods and contrast their advantages.

4.1 Model Specification and Results

We estimate three models of network formation and two models of behavior based on the specification in Berardo and Scholz (2010): an exact replication of their SIENA model using the R package `RSiena`, a cross-sectional ERGM, a temporal ERGM (TERGM), and a least-squares model of behavior. Each of the specifications includes the effects used in Berardo and Scholz (2010).

There are three node-level independent variables of interest: *trust*, which measures the respondent's trust in other stakeholders on an 11-point scale; *government actor*, an indicator for whether the actor works for a government agency; and *prodevelopment*, which is a 7-category measure of how pro-development (7) or pro-environmental (1) an actor self-reports themselves to be. For each of these three variables, sender, receiver, and similarity effects are included in the model specification. Sender (out) effects are the sum of each actor's out-degree (i.e. the number of outgoing connections possessed by the node) multiplied by the sender's covariate value. Similarly, receiver effects are the sum of each actor's in-degree multiplied by the sender's covariate value. Lastly, the similarity effects are edge-wise covariates included using the form of equation (2); similarity is measured as the negated absolute difference between two nodes' covariate values.[8]

The other set of effects we include in each model are network effects. First, we include effects that account for density of the networks under study. Density is accounted for with out-degree in SIENA and, analogously, the count of the number of edges in the ERGM and

[8]In the SIENA model, the absolute difference is normalized by the maximum absolute difference across all dyads in the network.

TERGM. We also include effects that capture whether there is a popularity effect, meaning that edges are more likely to be sent to those with many edges. These include the in-degree of the alter (i.e. the node connected to the node of interest (ego)) in the SIENA model and in-two-stars[9] in the ERGM and TERGM. A reciprocity/mutuality effect equivalent to that described in equation (1) is also included in each model. Furthermore, we include a count of the number of transitive triangles, triangles in which node i sends an edge to node j, node j sends an edge to node k, and node i also sends an edge to node k, in each model to capture clustering effects.[10] Lastly, each model is estimated using the 2001 estuary networks conditioned on the realization of those networks in 1999. This conditioning accounts for the history of the network to the extent possible with only two time points. In SIENA, temporal conditioning is accomplished with a *change* variable. In the ERGM and TERGM models, temporal conditioning is accomplished with a dichotomous edge-wise covariate that equals 1 if there was an edge in 1999 and -1 otherwise.

The behavior specifications, which model the response on the *trust* scale in 2001, include four substantive effects. First, the Alters' Trust effect is the average value of the previous value of Trust for each node's neighbors in the network. The Trust 1999 effect is a node-level covariate, which accounts for autocorrelation in Trust. Also, the effects of prodevelopment beliefs and the government actor indicator are included in the behavior model.

There is an important difference between the ERGM and TERGM with respect to the formulation of dependence effects and the estimation of the model. In the ERGM, the dependence effects are simultaneous, meaning that the probability of observing a network at time t depends, for instance, upon the number of in-two-stars in the current network. This creates a well known challenge for maximum likelihood estimation in that it is not typically computationally feasible to perform the summation in the denominator of equation (3) (Geyer

[9]Two stars capture the popularity by measuring how often node i has connections to nodes j and k. The statistic is computed $\Gamma_{2S}(Y) = \sum_i \sum_{i \neq j} Y_{ji} Y_{ki}$.

[10]Specifically, the statistic is computed $\Gamma_T(Y) = \sum_i \sum_{i \neq j,k} Y_{ij} Y_{jk} Y_{ik}$.

and Thompson 1992; Snijders 2002; Desmarais and Cranmer Forthcoming). When there are simultaneous dependence terms in the ERGM, the parameters must be estimated using an approximation method, of which Markov Chain Monte Carlo maximum likelihood estimation (MCMC-MLE) (Snijders 2002) and maximum pseudolikelihood estimation (Strauss and Ikeda 1990) are the most common. We use MCMC-MLE to estimate the ERGM. Hanneke, Fu and Xing (2010) suggest a compromise on the simultaneity of the effects in the TERGM that makes estimation much more feasible. They recommend that, instead of including simultaneous subgraph counts, dependence terms be specified as the single-edge completion of subgraphs from one period to the next. For instance, to measure reciprocity, instead of using the number of reciprocal dyads in the current network, we use the number of reciprocal dyads that can be created by combining one edge from the current network and one edge from the previous network. Instead of counting the number of transitive triangles in the current network, transitivity is measured as the number of transitive triangles created by combining two edges from the previous network and a single edge from the current network. The maximum likelihood estimator of this subgraph completion TERGM is equivalent to logistic regression. The edge-wise covariate value for edge ij in the current network is equal to the number of cross-period subgraphs completed when that edge is present. This form of the TERGM represents a compromise in the data generating process represented by the model, which results in a substantially more developed and better understood estimation option. As we show below, the appropriateness of this compromise can be evaluated by comparing the explanatory power of the TERGM to SIENA and the ERGM. The results of our estimation are given in table 1.

[Table 1 about here.]

As can be seen in table 1, results vary across the three models. There are, however, a few effects that are similar across the three models. In each model, there are positive and statistically significant popularity and reciprocity (mutuality) effects. The change effect is not technically comparable between the ERGM/TERGM and SIENA, but in both the ERGM and TERGM, we see evidence for positive autocorrelation in edge values. Also, the receiver effect of the government actor variable is positive in each model and statistically significant in both the ERGM and TERGM. A number of substantive inferences differ, however, based on which model is used. For instance, the SIENA model finds a positive but statistically insignificant effect for Trust-In, an independent variable of primary interest to Berardo and Scholz (2010) because it captures the extent to which more trusting actors will have higher in-degree, while the ERGM and TERGM find statistically significant negative effects. In other words, we have the common problem of our substantive conclusions being model dependent (Ho, Imai and King 2007). Our models of behavior do not shed any consistent light on the dynamics underlying inter-organizational trust in the estuary networks. The one substantive effect that is statistically significant is that of Alters' Trust in the least squares model of behavior, but the effect is positive and insignificant in the behavior component of the SIENA model.[11] As we discussed above, there are subtle differences across these three models, and, as we see from the estimates, which model is chosen to evaluate a particular hypothesis can change the substantive inferences drawn from the analysis. As such, we propose that, in addition to careful consideration of the assumptions behind each model, attention to model fit is necessary to avoid the possibility of drawing false inferences.

4.2 Simulation-based model comparison

Each model in our analysis is estimated using a different method: SIENA models are estimated by simulated method of moments, the ERGM is estimated by MCMC-MLE, and

[11] Our SIENA results differ from those of Berardo and Scholz (2010), despite the fact that we use the same data and specification. Possible reasons for this include (1) the inherent randomness in the simulation-based estimation, and (2) several new version releases of the software (i.e., we ran our models in `RSiena`, whereas Berardo and Scholz (2010) used the original, GUI-based, standalone software.

the TERGM is estimated by MLE. There are different assumptions underlying each model and also slight differences in the specification of the common effects. However, the objective with each is to fit a model that provides a parsimonious and accurate explanation of the data at hand, which can be extrapolated to other similar networks. As such, we can turn to out-of-sample predictive methods, which evaluate fit on data that were *not* used to estimate the model, to compare model fit while guarding against over-fitting (Ward, Greenhill and Bakke 2010). In order to compare the out-of-sample predictive performance of our three models, we divide the sample of ten networks into a training sample of five networks and a validation sample of five networks. We re-estimate the three models on the training sample and then, based on the initial networks and covariate values for the validation sample, we simulate 1,000 predicted networks for each of the validation networks using the parameter estimates derived from the training sample. As a baseline, we also include an Erdös-Rényi graph, which is an ERGM with only an edges (i.e., density) term estimated.

In terms of the network dynamics, we evaluate model fit based on the mean absolute difference between network statistics generated in the simulation and the network statistics computed on the validation networks. Lower values indicate better fit. Predictive R^2 is used to evaluate the fit of the behavior components; larger values indicate better fit. We consider several predictive metrics to judge model fit for the network: mutuality (reciprocity) and transitivity at the network level Butts (2010), the in- and out-degree of each node in the network, and the individual edge values. The results are depicted in figure 1 and do not support either the ERGM or SIENA models. The Erdös-Rényi graph, a simple Bernoulli random network, out-performs ERGM and SIENA according to every metric. The TERGM performs best out of all models considered. According to only one of the metrics, graph transitivity, does the Erdös-Rényi network outperform the TERGM. In other words, the TERGM provides the best predictions on four of the five metrics, the baseline Erdös-Rényi random graph provides the best predictions on one of the five metrics, and both the

ERGM an SIENA provide the best predictions on none of the five metrics. In terms of the behavior component, the least-squares model out-performs the SIENA model at predicting the values of Trust in 2001 for the held-out nodes. However, neither of the behavior models are particularly illuminating, so the choice between these two models is less consequential.

[Figure 1 about here.]

To be clear, our out-of-sample predictive results do not suggest that the network-analytic approach to modeling these estuary networks should be discarded. Rather, they support the use of a theoretically informed model that incorporates single-period lagged network dependences. Fortunately, the TERGM is substantially easier to estimate than either SIENA or the ERGM, so such a conclusion would not present an barrier to applied researchers.

4.3 Edge, Dyad and Node-Level ERGM Interpretation

We apply our proposed micro-level interpretation methods to the ERGM results.[12] For illustration, we render model interpretation at the edge, dyad, and node levels. In each instance, we compute the effects at 100 randomly selected positions in the network (i.e., 100 randomly selected edges, dyads or nodes). In order to minimize the effect of focusing our interpretation on the worse-fitting model, we limit our attention to effects that are fairly consistent across the three models.

The interpretation edge, dyad, and node level interpretations are presented graphically in figure 2. The edge-stability (i.e., change) effect is illustrated by the probability that an edge exists in the current period given the state of the edge in the previous period. It can be seen here that the presence of an edge in the previous period nearly triples the likelihood

[12] It may seem odd to focus on the ERGM, since it is not the best-fitting model in this application. However, the ERGM is much more common in the literature than the TERGM. For this reason, we see a focus on the ERGM results as more pedagogically useful for a general readership.

that an edge will be present in the current period, thus signaling a strong stability effect. The dyad-level effect illustrates the dyad distribution, which is regulated by the mutuality effect. In the first mutuality plot, the estimated dyad distribution is depicted. This does not convey the reciprocity in the networks. The second mutuality plot illustrates the ratio of the estimated dyad distribution to the dyad distribution when the mutuality effect is fixed at zero. From this second plot, we see that the probability of a mutual dyad in the estimated model is approximately twice as likely relative to a model with no mutuality effect.

Lastly, we render a node-level interpretation that we feel is particularly appropriate given this application. The edges in the network under study represent information-provision relationships. It is thus reasonable to interpret popular nodes as experts. The popularity effects in these models can be interpreted as a process by which nodes become recognized as experts. To interpret popularity, we randomly fix a number of in-going edges to a node, then evaluate the average probability of an in-going edge among the non-fixed edges to that node. We find that the number of in-coming edges to an organization substantially increases the likelihood of additional in-coming edges. When there are zero fixed in-going edges, there is roughly a 5% chance of an in-coming edge from a randomly selected sender. When there are six randomly fixed in-going edges, there roughly a 25% chance of an in-coming edge from a randomly selected non-fixed sender. Thus, organizations with many edges are likely to receive more edges; a process consistent with organizations becoming known experts or valuable partners.

[Figure 2 about here.]

5 Conclusion

We have shown that, despite the fact that it is often used for network-level inference, the ERGM and its temporal extension may be used to draw inference and test hypotheses about actor (node) and relationship (edge) level effects. The ERGM is fully consistent with a dynamic interpretation of network formation in which blocks of the network – consisting of a single node, single edge, or any conceivable network partition of interest – update their ties sequentially. As such, the ERGM can model the process by which any given block in the network changes conditional on the rest of the network; a powerful and general paradigm for modeling networks that change over time and interpreting effects of interest at the micro-level.

The dynamic interpretations of the ERGM and TERGM are quite similar the dynamic model implemented in SIENA; the SIENA model being the dominant model used by political and policy scientists to analyze dynamic networks with a substantive interested in actor- and relation-level effects. We have shown that, though similar, SIENA imposes some string assumptions about the mechanisms by which networks change over time, while the ERGM is somewhat more general. Depending on the application at hand, a more general (ERGM/TERGM) or more specific (SIENA) model may be theoretically preferable, but often the choice between models will not be obvious *a priori*.

We showed, through our reanalysis of Berardo and Scholz's (2010) work on estuary networks, that one can use out-of-sample predictive performance to judge which of the three candidate models (SIENA, ERGM, and TERGM) is more appropriate in situations where substantive theory does not guide one towards SIENA's specific change mechanism. In our application, we found that, not only do the substantive inferences one would draw from the three models vary, but that the TERGM consistently outperforms SIENA and the ERGM in predictive performance; often by a wide margin.

Looking forward, we hope that our explication of the dynamic process consistent with the ERGM and TERGM, as well as our approach to diagnostic testing of model fit, will make the ERGM and TERGM more useful to applied researchers interested in effects at the actor and relationship levels; not just for network-level inference. To aid this process, we have developed companion software, described in the appendix, that makes out-of-sample prediction with network models and micro-level interpretation of the ERGM or TERGM simple to implement.

6 Software Appendix

We have implemented methods that interface with the `ergm` package in R to compute the conditional probabilities of individual edges, dyads, outgoing edges coming from a node and ingoing edges going to a node. They are available at `http://.../InterpretationFunctions.R`. The functions `pNode`, `pDyad`, and `pEdge` compute probabilities at the node, dyad, and edge levels, respectively. The functions we provide take estimates from an ERGM model as arguments for the conditioned and conditioning network blocks. Below we illustrate how this interpretation method, implemented using these functions in R, can be used to produce lower-level conditional interpretations of ERGM results.

We begin by considering the edge-level function, as it is a single predicted probability.

```
# Read in the ERGM Library
library(ergm)
# Read in the SNA library
library(sna)


# Read in the Interpretation Functions
```

```
source("InterpretationFunctions.R")

# Set the random number seed
set.seed(5)

# Create an artificial network (see "?rgraph" for more details) of 20 nodes
art_net <- network(rgraph(20,tprob=0.25))

# Estimate an ERGM with edges, out and in two stars, and reciprocity
gest <- ergm(art_net ~ edges + istar(2)+ostar(2) + mutual)
summary(gest)

# Compute the probability of an edge from node one to node two
prob <- pEdge(ergm_formula = "net ~ edges + istar(2)+ostar(2) + mutual",theta = coef
(gest),i=1,j=2,net=art_net)

# Print prob in the R console
> prob
[1,] 0.1702965
```

The `pEdge` function takes several arguments. The first argument, 'ergm_formula' is the formula used to estimate the ERGM model given in quotation marks (i.e., as a character object), with 'net' substituted for the name of the network. The second argument, 'theta,' is the vector of ERGM parameter estimates. These can be obtained by wrapping `coef()` around the ERGM object (i.e., 'gest'). The arguments 'i' and 'j' indicate the sender and receiver nodes for which the probability of an edge is to be computed, respectively. Lastly, 'net' is the network for which 'net' was substituted in the first argument. The function returns the probability that there is an edge from i to j, given the ERGM specification, parameter estimates, and the rest of the network.

Now we turn to the function, `pDyad`, which computes the probability distribution of a dyad (i.e., the joint distribution of the two edges in a dyad), given the rest of the network.

```
# Compute the dyad distribution
prob_dyad <- pDyad(ergm_formula = "net ~ edges + istar(2)+ostar(2) + mutual",
theta = coef(gest),i=1,j=2,net=art_net)


# Print it in the R console
> prob_dyad
          j->i = 0  j->i = 1
i->j = 0 0.6295625 0.1384381
i->j = 1 0.1292176 0.1027818
```

The arguments to the `pDyad` function are the same as those to `pEdge`, except the 'i' and 'j' represent both edges in the dyad rather than a single edge. The matrix returned gives the probability of all four possible combinations of the edge from 'i' to 'j' and the edge from 'j' to 'i', given the model and the rest of the network. The row and column labels indicate the state of each edge in the respective dyad outcome. For instance, looking at the bottom-right cell, we see that there is a 0.1027818 probability that there is an edge from i to j and one from j to i.

Lastly, we illustrate the `pNode` function, which computes the joint probability of a group of out-going or in-coming ties from/to a node.

```
# Compute the joint distribution of three out-going edges
```

```
prob_out <- pNode(ergm_formula = "net ~ edges + istar(2)+ostar(2) + mutual",
theta = coef(gest),node=4,others=c(3,1,8),nodeSend=T,net=art_net)
```

```
> prob_out
            prob Receiver3 Receiver1 Receiver8
     0.584567587         0         0         0
     0.005390058         1         1         1
veci 0.114907302         1         0         0
veci 0.109148871         0         1         0
veci 0.114907302         0         0         1
veci 0.023283500         1         1         0
veci 0.024511881         1         0         1
veci 0.023283500         0         1         1
```

The arguments to `pNode` differ from those to the two previous functions. There are three actor-related arguments. The argument 'node' indicates the single sender or receiver in the block. The 'others' argument is a vector of recipient nodes in the case that 'node' is a sender, and sender nodes in the case that 'node' is a receiver. The argument 'nodeSend' is a logical indicator of whether 'node' is a sender (T) or receiver (F). The matrix returned has a number of columns equal to the length of 'others' plus one. The first is the probability of the block-wise outcome given the model and the rest of the network, and the other columns indicate the state of the edges between 'node' and 'others'. For example, from the first row, we see that there is a 0.584567587 probability that there are no edges from node 4 to nodes 3, 1 and 8.

The quantities derived with these functions can be used to compute the interpretation metrics given in Figure 2, as well as many other measures. For instance, to compute the 'Edge Stability' metric, `pEdge` is used to compute the probability of an edge given an edge in the previous period versus no edge in the previous period. Confidence intervals are derived by computing the probabilities over 100 randomly selected directed dyads. Also, the 'Popularity Effect' is computed by (a) randomly selecting a target node, (b) randomly fixing a number of in-going ties corresponding to the x-axis in the 'Popularity Effect' plot, and (c) using `pNode` to compute the average probability that the non-fixed in-going ties to that node exist.

References

Berardo, Ramiro and John T. Scholz. 2010. "Self-Organizing Policy Networks: Risk, Partner Selection, and Cooperation in Estuaries." *American Journal of Political Science* 54(3):632–649.

Butts, Carter T. 2010. *sna: Tools for Social Network Analysis*. R package version 2.2-0. **URL:** *http://CRAN.R-project.org/package=sna*

Cranmer, Skyler J. and Bruce A. Desmarais. 2011. "Inferential Network Analysis with Exponential Random Graph Models." *Political Analysis* 19(1):66–86.

Cranmer, Skyler J., Bruce A. Desmarais and Elizabeth Menninga. Forthcoming. "Complex Dependencies in the Alliance Network." *Conflict Management and Peace Science* .

Desmarais, Bruce A. and Skyler J. Cranmer. 2010. "Consistent Confidence Intervals for Maximum Pseudolikelihood Estimators." *Neural Information Processing Systems: Workshop on Computational Social Science and the Wisdom of Crowds* .

Desmarais, Bruce A. and Skyler J. Cranmer. 2011. "Statistical Inference for Valued-Edge Networks: The Generalized Exponential Random Graph Model." *PLoS-one (revise and resubmit)* .

Desmarais, Bruce A. and Skyler J. Cranmer. Accepted. Forecasting the Locational Dynamics of Transnational Terrorism: A Network Analytic Approach. In *Proceedings of the European Intelligence and Security Informatics Conference (EISIC) 2011*. IEEE Computer Society.

Desmarais, Bruce A. and Skyler J. Cranmer. Forthcoming. "Statistical Mechanics of Networks: Estimation and Uncertainty." *Physica A* .

Geman, Stuart and Donald Geman. 1984. "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* PAMI-6(6):721 –741.

Geyer, Charles J. and Elizabeth A. Thompson. 1992. "Constrained Monte Carlo maximum likelihood for dependent data." *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 54(3):657–699.

Gill, Jeff. 2007. *Bayesian methods: A social and behavioral sciences approach.* New York: Chapman and Hall/CRC.

Goodreau, Steven .M., James A. Kitts and Martina Morris. 2009. "Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks." *Demography* 46(1):103–25.

Handcock, Mark S., David R. Hunter, Carter T. Butts, Steven M. Goodreau, Martina Morris and Pavel Krivitsky. 2010. *ergm: a package to fit, simulate and diagnose exponential-family models for networks.* Seattle, WA: . Version 2.2-2. Project home page at url-http://statnetproject.org.

URL: *http://CRAN.R-project.org/package=ergm*

Hanneke, Steve, Wenjie Fu and Eric P. Xing. 2010. "Discrete Temporal Models of Social Networks." *The Electronic Journal of Statistics* 4:585–605.

Ho, Daniel, Kosuke Imai and Gary King. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15(3):199–236.

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

Krivitsky, Pavel N. 2011. “Exponential-Family Random Graph Models for Valued Networks.”.

Lazer, David, Brian Rubineau, Carol Chetkovich, Nancy Katz and Michael Neblo. 2010. “The coevolution of networks and political attitudes.” *Political Communication* 27(3):248–274.

Lehmann, Erich Leo. 1983. *Theory of Point Estimation*. New York: Wiley & Sons.

Nelder, J.A. and R.W.M. Wedderburn. 1972. “Generalized linear models.” *Journal of the Royal Statistical Society. Series A (General)* 135(3):370–384.

Oakerson, Ronald J. and Roger B. Parks. 2011. “The Study of Local Public Economies: Multi-organizational, Multi-level Institutional Analysis and Development.” *Policy Studies Journal* 39(1):147–167.

Park, Juyong. and M.E.J. Newman. 2004. “Statistical mechanics of networks.” *Physical Review E* 70(6):66117–66130.

Rapaport, Orit, David Levi-Faur and Dan Miodownik. 2009. “The Puzzle of the Diffusion of Central-Bank Independence Reforms: Insights from an Agent-Based Simulation.” *Policy Studies Journal* 37(4):695–716.

Schneider, Mark, John Scholz, Mark Lubell, Denisa Mindruta and Matthew Edwardsen. 2003. “Building Consensual Institutions: Networks and the National Estuary Program.” *American Journal of Political Science* 47(1):143–158.

Snijders, Tom A. B., Philippa E. Pattison, Garry L. Robins and Mark S. Handcock. 2006. “New specifications for exponential random graph models.” *Sociological Methodology* 36(1):99–153.

Snijders, Tom A.B. 2002. “Markov Chain Monte Carlo Estimation of Exponential Random Graph Models.” *Journal of Social Structure* 3(2):1–40.

Snijders, Tom A.B., Gerhard G. van de Bunt and Christian E.G. Steglich. 2010. "Introduction to stochastic actor-based models for network dynamics." *Social Networks* 32(1):44 – 60.

Strauss, David and Michael Ikeda. 1990. "Pseudolikelihood estimation for social networks." *Journal of the American Statistical Association* 85(409):204–212.

Ward, Michael D., Brian D. Greenhill and Kristin M. Bakke. 2010. "The Perils of Policy by p-value: Predicting Civil Conflicts." *Journal of Peace Research* 47(4):363–375.

Wasserman, Stanley and Philippa Pattison. 1996. "Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and p^*." *Psychometrika* 61(3):401–425.

Williams, John T. 1992. "What Goes Around Comes Around: Unit Root Tests and Cointegration." *Political Analysis* 4(1):229–235.

Wyatt, D., T. Choudhury and J. Bilmes. 2010. Discovering long range properties of social networks with multi-valued time-inhomogeneous models. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*. pp. 630–636.

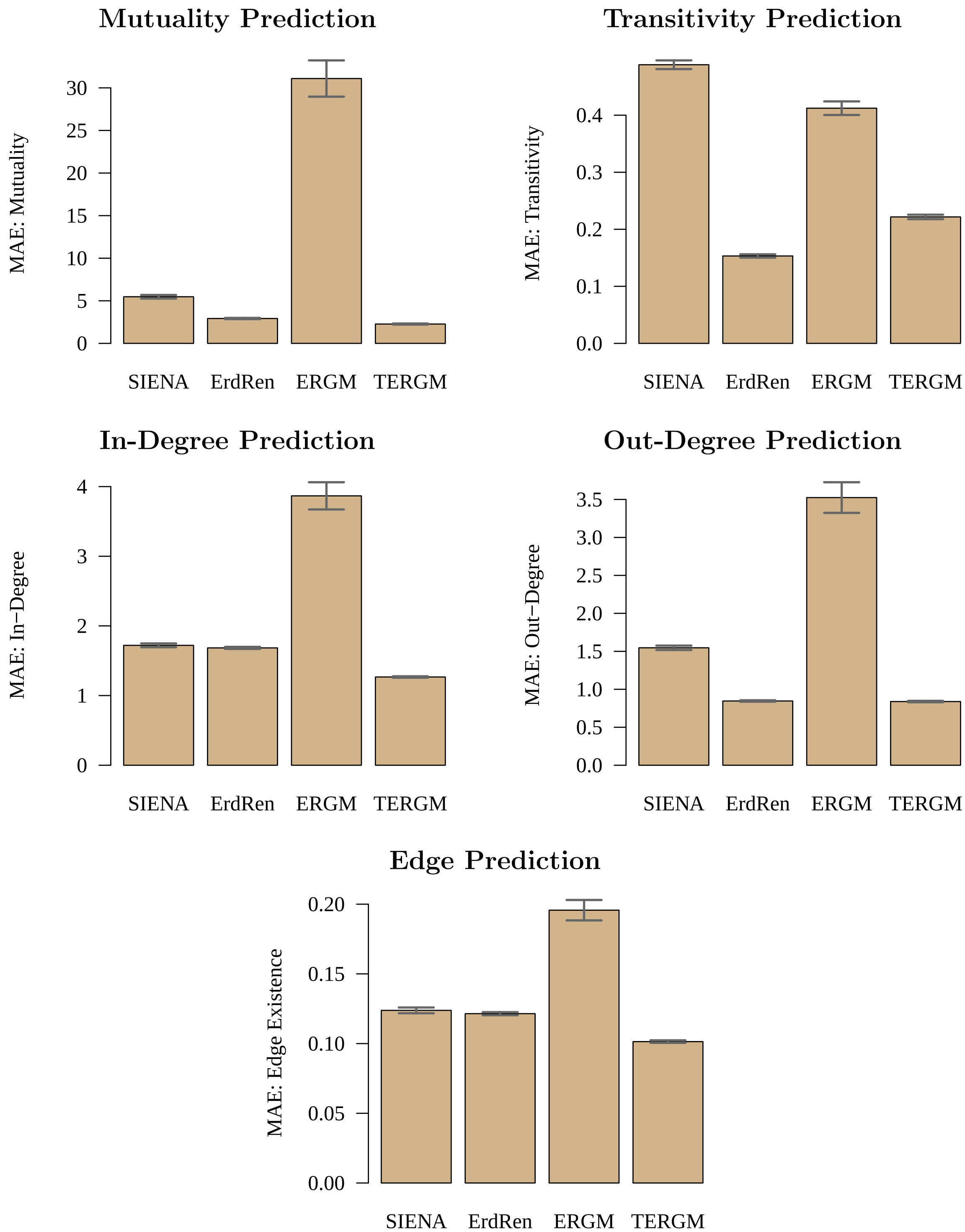

Figure 1: Out-of-sample predictive performance. Barplots depict mean absolute error (MAE) in predicting the respective quantities in the observed validation networks using networks simulated from the respective models estimated using the training networks. Error bars depict 95% confidence intervals over 1,000 simulations.

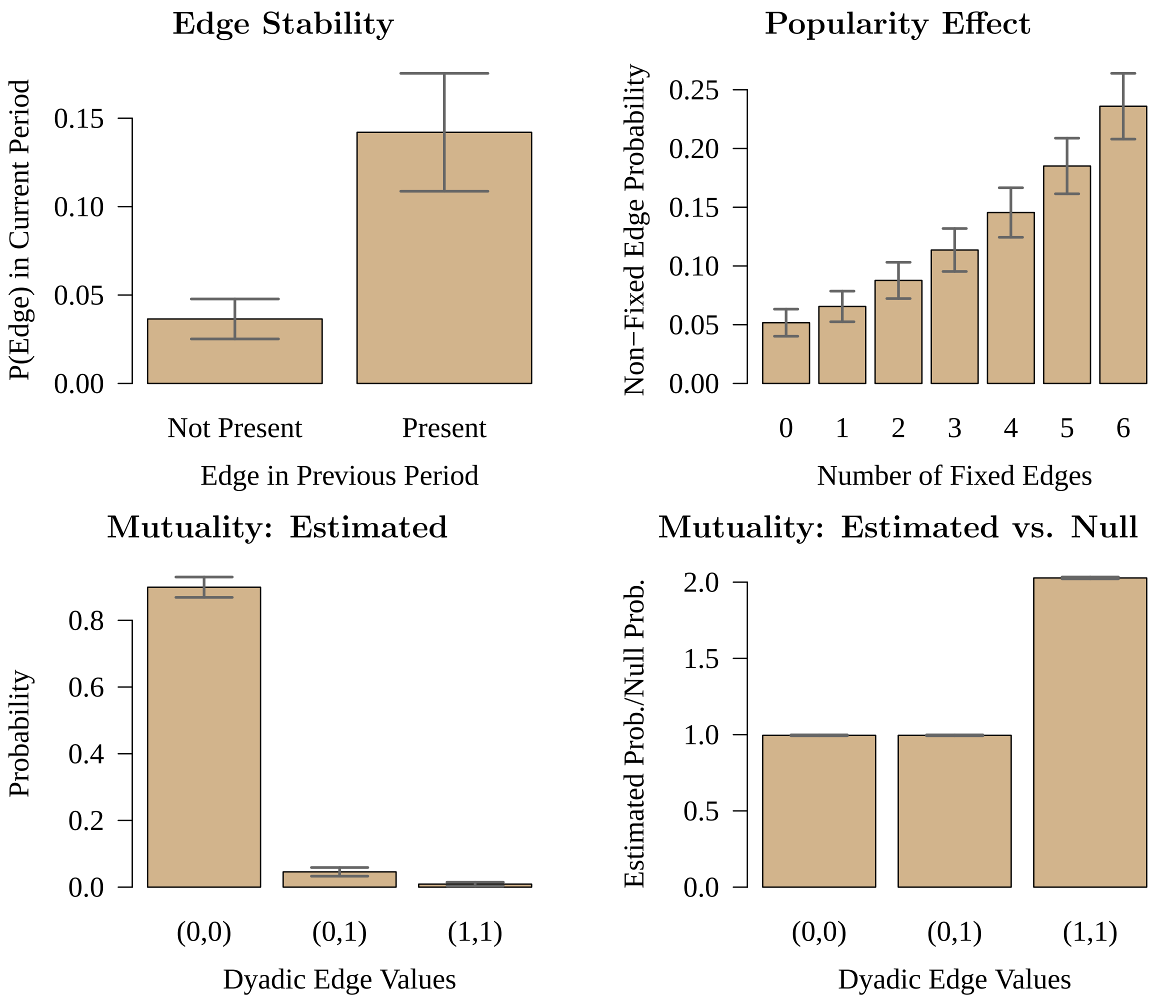

Figure 2: Interpretation at different levels. Barplots depict the mean of respective quantities computed from the full ERGM estimates in table 1. Error bars depict 95% confidence intervals over 100 randomly selected subgraphs. Note, in the mutuality plots (0,1) represents the average over the two asymmetric configurations in the dyad.

	SIENA		ERGM		TERGM	
Network Dynamics						
Density	**-3.181**	**(0.19)**	**-2.694**	**(0.72)**	0.527	(3.90)
Popularity	**0.99**	**(0.11)**	**0.332**	**(0.04)**	**0.268**	**(0.03)**
Mutuality	**0.722**	**(0.24)**	**0.711**	**(0.09)**	**0.699**	**(0.21)**
Transitivity	0.093	(0.14)	**0.15**	**(0.07)**	**-0.013**	**(0.00)**
Change	**4.999**	**(0.51)**	**0.874**	**(0.03)**	**12.234**	**(1.24)**
Trust-Out	-0.048	(0.13)	**0.007**	**(0.00)**	-0.02	(0.04)
Trust-In	0.003	(0.06)	**-0.038**	**(0.00)**	**-0.098**	**(0.03)**
Trust-Sim	0.967	(1.89)	0.002	(0.00)	0.014	(0.04)
Prodev-Out	0.274	(0.42)	**0.37**	**(0.01)**	0.471	(0.50)
Prodev-In	0.229	(0.33)	**0.171**	**(0.01)**	0.464	(0.46)
Prodev-Sim	0.288	(0.68)	0.036	(0.04)	0.291	(0.51)
Gov-Out	-0.001	(0.18)	**0.049**	**(0.01)**	0.06	(0.17)
Gov-In	0.207	(0.12)	**0.12**	**(0.01)**	**0.402**	**(0.17)**
Gov-Sim	0.183	(0.15)	**0.328**	**(0.01)**	0.279	(0.16)
Behavior Dynamics						
Intercept (Rate)	14.12	(74.79)	**9.09**	**(1.73)**	–	–
Trust Tendency	0.27	(0.99)	–	–	–	–
Alters' Trust	0.72	(2.88)	**-0.163**	**(0.064)**	–	–
Trust 1999	-0.0614	(0.660)	0.161	(0.146)	–	–
Prodev	-0.116	(2.83)	-1.32	(1.81)	–	–
Gov.	-0.175	(0.180)	0.208	0.616	–	–
Predictive R^2	1.07e-5		0.042			

Table 1: Coefficient estimates, with standard errors in parentheses, from the SIENA, ERGM, and TERGM models of estuary networks respectively. Estimates in **bold** are statistically significant at the traditional 0.05 level.