A STRUCTURAL MODEL OF HOMOPHILY AND CLUSTERING IN SOCIAL NETWORKS

ANGELO MELE

JOHNS HOPKINS UNIVERSITY

This version: May 1, 2020. First version: August 20, 2017. Contact: angelo.mele@jhu.edu. I thank the editor, the associate editor and two extremely competent and helpful referees for suggestions that substantially improved the paper. I am grateful to Eric Auerbach, Federico Bandi, Stephane Bonhomme, Vincent Boucher, Aureo dePaula, Patrick Kline, Michael Konig, Michael Leung, Bryan Graham, Eleonora Patacchini, Demian Pouzo, Lingjiong Zhu, Shuyang Sheng, Elie Tamer, Roger Moon, Zhongjun Qu, Xiaodong Liu and Hiro Kaido and participants to many seminars and conferences for comments and suggestions. Special thanks to Michael Schweinberger for helping with the use of his package `hergm`. All remaining errors are mine.

This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining Data Files from Add Health should contact Add Health, The University of North Carolina at Chapel Hill, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (addhealth@unc.edu). No direct support was received from grant P01-HD31921 for this analysis.

Abstract. I develop and estimate a structural model of network formation with heterogeneous players and latent community structure, whose equilibrium homophily and clustering levels match those usually observed in real-world social networks. Players belong to communities unobserved by the econometrician and have community-specific payoffs, allowing preferences to have a bias for similar people. Players meet sequentially and decide whether to form bilateral links, after receiving a random matching shock. The model converges to a hierarchical exponential family random graph. Using school friendship network data from Add Health, I estimate the posterior distribution of parameters and unobserved heterogeneity, detecting high levels of racial homophily and payoff heterogeneity across communities. The posterior predictions of sufficient statistics show that the model is able to replicate the homophily levels and the aggregate clustering of the observed network, in contrast with standard exponential family network models without community structure.

JEL CODES: C13, C31, C57, D85

Keywords: Social networks, exponential random graphs, homophily, clustering

1. Introduction

Social networks are important determinants of economic success in many contexts: health, education, crime, investment, politics, employment, new product adoption (Jackson, 2008; DePaula, 2017; Chandrasekhar, 2016; Acemoglu et al., 2011; Golub and Jackson, 2011; Fafchamps and Gubert, 2007; Laschever, 2009; Topa, 2001; Conley and Udry, 2010; Echenique and Fryer, 2007; Nakajima, 2007; De Giorgi et al., 2010; Goldsmith-Pinkham and Imbens, 2013; Calvo-Armengol et al., 2009). Therefore it is crucial to understand how these relationships are formed, what network architectures are optimal, and what policies can affect their shape.

The analysis developed in this paper builds on the empirical observation that social networks tend to organize in clusters of densely connected individuals, with fewer links across clusters. There are at least two ways to rationalize this stylized fact. First, the clustering

may be the realization of preferences for similar individuals. This is what in the networks literature is referred to as homophily, that is the propensity to form links to similar people. Homophily has been reported for observable characteristics like race, gender, age, income, education and other demographics (Jackson, 2008; Currarini et al., 2009; Moody, 2001; Mele, 2017; DePaula et al., 2018; Boucher, 2015). Individuals may as well have biased preferences for similar unobserved characteristics (Graham, 2017; Dzemski, 2017; Boucher and Mourifie, 2017; Leung, 2014). Second, the clustering may be the result of preference for transitivity, that is the willingness to link of two individuals is increasing in the number of common neighbors in the network. Indeed, it is usually observed that if two people have a link to a common neighbor, it is very likely that they are also connected to each other (Jackson, 2008; Jackson and Rogers, 2007). This rationalization is related to strategic interactions, as the decision to form a link depends on decisions to form links of other individuals as well. As argued in Graham (2017), it is important to distinguish between these two explanations – homophily and preference for transitivity – for policy purposes. If the clustering is mostly explained by strategic interactions, then a policy that affects one link will have a cascade effect on many other links. On the other hand, if the observed clustering mostly reflects homophily in preferences, a policy targeting one link will not have any additional effect on the rest of the network.

I propose a structural model of network formation with heterogeneous players and latent community structure, to estimate preferences for homophily and transitivity in equilibrium. The model's equilibrium belongs to the class of discrete exponential family random graphs (Moody, 2001; Snijders, 2002; Caimo and Friel, 2011; Boucher and Mourifie, 2017; Mele, 2017). The point of departure from previous work is the introduction of a latent block structure to model unobserved heterogeneity in preferences to form links. These latent communities modify the dependence among linking decisions of the agents, thus the model's equilibrium converges to a hierarchical exponential random graph model (Schweinberger and Handcock, 2015).

The network is formed sequentially: in each period two players are randomly selected from the population and meet. Upon meeting, players decide whether to update their link, by maximizing the sum of their current utility. In the absence of any shock to the preferences, this process of network formation is consistent with pairwise stability with transfers, a common equilibrium notion used in the network formation literature (Jackson, 2008). Conditional on the latent community structure, the network formation process can be characterized as a potential game and in the long-run the sequence of link updates converges to a stationary distribution over networks, which is a discrete exponential family with intractable normalizing constant (Geyer and Thompson, 1992; Mele, 2017). This implies that in the long run, the observed networks are pairwise stable (with transfers) with very high probability (Monderer and Shapley, 1996; Butts, 2009; Jackson and Watts, 2001; Badev, 2013; Hsieh and Lee, 2012). This result leverages the microeconomic foundations developed in Mele (2017).

Players are partitioned into non-overlapping communities. Upon birth, each player is randomly assigned to a community. Preferences are defined over networks, covariates and community structure. The players' payoffs depend on the composition of direct connections, but also on the number of common friends. Preferences also depend on the unobserved heterogeneity: members of different communities are allowed to have different costs and benefits of forming links and different payoffs from common friends.

I estimate the posterior distribution of the structural preference parameters using a MCMC algorithm that sequentially samples community structure, parameters and simulates networks to approximate both normalizing constants of likelihood and posterior (Murray et al., 2006; Mele, 2017; Schweinberger and Handcock, 2015). A practical challenge to estimation is that the likelihood of the model is invariant to the labeling of the communities. To alleviate this problem, I impose a nonparametric prior that is not invariant to permutations of the labels (Ishwaran and James, 2001; Schweinberger and Handcock, 2015), and I use established methods to relabel the posterior simulation output (Stephens, 2000; McLachlan and Peel, 2000).

I estimate the model for different specifications of the payoff functions and different numbers of unobserved communities. The net benefits of a link are allowed to vary by unobserved type. I consider three specifications for the taste for transitivity: 1) no transitivity (i.e. stochastic blockmodels, Airoldi et al. (2008)); 2) homogeneous taste for transitivity, that does not vary by community as in standard ERGMs (Snijders, 2002); 3) local transitivity, so that the preference for transitivity is zero for links across communities, but nonzero for links within types (Schweinberger and Handcock, 2015).

The dataset used in estimation contains the network of friendships in a US high school, extracted from *The National Longitudinal Study of Adolescent Health* (Add Health). The data also contains race, gender, grade and parental income of each student.

As a practical matter, there is no established procedure to choose the model specification or the number of communities for this class of models. Standard procedures used in the stochastic blockmodels literature do not trivially extend to this more complex model. Therefore, I use an heuristic approach, estimating each specification with an increasing number of communities; then I check whether the posterior predictions of the estimated models are able to fit the properties of the observed network, using standard goodness of fit tests from the ERGM literature (Caimo and Friel, 2011; Schweinberger and Handcock, 2015). For example, I check whether the estimated model can replicate the number of links and triangles of the observed network.

According to this procedure, the models without taste for transitivity (stochastic blockmodels) severely underestimate the number of triangles, that is the clustering level of the network. This is well understood in the literature (Chandrasekhar, 2016; DePaula, 2017; Chandrasekhar and Jackson, 2016; Snijders, 2002; Schweinberger and Handcock, 2015; Mele, 2017). Models with transitivity provide better fit of the data.

There are few important results. First, a model without community structure performs much worse than a model with multiple communities. Indeed, the model with one block

cannot match the number of edges or triangles of the data, providing mostly posterior predictions that are degenerate (either almost complete or almost empty networks). Second, the models with local transitivity seem to perform slightly better than models with global transitivity. In Monte Carlo simulations the model with local transitivity performs as well as a stochastic blockmodel in recovering the block structure, according to the Yule's coefficient of block membership recovery. Finally, when estimated with real-world data, the posterior predictions of the model with local transitivity put most probability mass around the observed sufficient staistics of the network, such as number of links, triangles, stars and homophily levels.

The empirical estimates show high levels of homophily: preferences are biased towards links with students of the same racial group, grade and (parental) income levels. Furthermore, the estimated payoffs vary by community. The latter result proves that it is important to include unobserved heterogeneity in empirical network formation models, as there could be unobserved characteristics that make agents more or less social on average.

This paper contributes to the literature on empirical network formation models by developing a structural model that is able to estimate preferences for homophily in observed and unobserved characteristics, as well as taste for transitivity, in a strategic equilibrium. Therefore, this work provides a game-theoretical equilibrium counterpart of Schweinberger and Handcock (2015)'s statistical work, leveraging the strategic approach developed in the economics literature on networks (Jackson, 2008; Jackson and Wolinsky, 1996; Jackson and Watts, 2001; Bala and Goyal, 2000; Galeotti, 2006). In fact, my structural model mantains the flexible specification of the exponential family random graphs, while incorporating the strategic and equilibrium microfoundations introduced in Mele (2017).

Unobserved heterogeneity has been introduced in empirical network models in different ways. Graham (2016) considers a model with conditionally independent links, where estimation is performed using repeated observations over time of the network. Breza et al. (2017) and McCormick and Zheng (2015) consider hierarchical Bayes parametric community

structure estimation using summary statistics of a partially observed nework from a random sample of nodes. Graham (2017); Dzemski (2017); Charbonneau (2017); Jochmans (2017) and the literature on the β-model in statistics, model unobserved heterogeneity without including transitivity, estimating the model using only a network observation. My paper complements these approaches by considering the case in which the researchers has access to one network observation only and the model includes transitivity.

Structural models of network formation usually include biases in preferences or meetings to generate homophily in equilibrium (Currarini et al., 2009; Boucher, 2015; Mayer and Puller, 2008); some also include payoffs from common neighbors (DePaula et al., 2018; Menzel, 2015; Sheng, 2012; Ridder and Sheng, 2015). However, disentangling homophily, clustering and unobserved heterogeneity using only one network observation is extremely challenging (Graham, 2017; Chandrasekhar and Jackson, 2016). My model contributes to this literature by providing a Bayesian approach to estimating payoffs for homophily and transitivity in a strategic model.

The estimation of the model is computationally intensive, because the community structure is unobserved, limiting the Bayesian estimation approach to networks with a few hundred nodes. However, one could pre-process the data and use an algorithm to identify and estimate community memberships before the estimation of the network model, like in Bonhomme et al. (2017). This would speed up computations and estimation, because the simulation of the community structure is a major computational bottleneck.

The rest of the paper is organized as follows. Section 2 presents the structural model and the equilibrium characterization. Section 3 develops the estimation strategy and the asymptotic behavior of the sufficient statistics. Section 4 focuses on the empirical results, showing that friendship school networks exhibit high level of homophily and clustering, with a moderate level of unobserved heterogeneity in preferences. Section 5 concludes. The appendices contain additional theoretical results, proofs and details about the estimation.

2. A structural model of network formation

The economy consists of n players, each characterized by an M-dimensional vector $x_i = \{x_{i,1}, x_{i,2}, ..., x_{i,M}\}$ of observable characteristics. For example, x_i may include the gender, racial group, income, education levels and age of each individual in the economy.

Each player belongs to a *community*: this membership is observed by the players, but not by the econometrician. The community models unobserved heterogeneity in a very specific form: there exists some unobserved characteristics that separates individuals into different types. For example, some individuals are more social than others, a personality trait that is difficult to observe; some people care a lot about having a tightly-knit group of friends, others do not care as much. Type can represent geographic proximity, as well as social proximity. A player's type affects her preferences as well as the probability of meeting other people, as explained in detail below. Formally, the community structure is a partition of the n players in K subsets $\{\mathcal{C}_1, \mathcal{C}_2, ..., \mathcal{C}_K\}$. The K-dimensional vector of binary indicators $z_i = \{z_{i,1}, z_{i,2}, ..., z_{i,K}\}$ contains the membership information for player i. That is, player i belongs to community k if $z_{i,k} = 1$ and $z_{i,l} = 0$ for all the $l \neq k$. I will consider a model in which individuals can be member of one community only. The researcher cannot observe the memberships vectors z_i, nor the number of communities K.

Before the network formation game, nature chooses who belongs to each community (the community structure), according to a multinomial distribution. This is a relatively standard assumption in the stochastic blockmodels literature (Airoldi et al., 2008; Schweinberger and Handcock, 2015). The main advantage of the multinomial specification is that priors can be specified in a nonparametric way to speed up the computations of the posterior. I assume that the community assignment to each player is i.i.d.

$$Z_i | \eta_1, ..., \eta_K \overset{iid}{\sim} Multinomial\,(1; \eta_1, ...\eta_K) \text{ for } i = 1, ..., n \tag{1}$$

The $n \times M$ matrix x contains all the vectors of observable characteristics, and the $n \times K$ matrix z contains all the vectors of membership indicators.

The network of interactions is represented by a $n \times n$ matrix g, the adjacency matrix of the network, whose generic element $g_{ij} = 1$ if there is a link between i and j, and $g_{ij} = 0$ otherwise. I will consider an *undirected* network, with a symmetric matrix g. Some of the results below can be easily extended to directed networks (Mele, 2017; Badev, 2013; Schweinberger and Handcock, 2015).

The network formation process works as follows. In period 0, nature randomly chooses the type for each player i. Conditional, on the types $Z = z$, the network is formed sequentially as in Mele (2017) and Mele and Zhu (2020): in each period, two random players, i and j meet with probability $\rho(g, z_i, z_j, x_i, x_j)$. This probability can depend on the existing network g: for example, two people may meet more often if they have some friends in common. The function ρ also depends on the unobservable communities indicators z_i and z_j: people belonging to the same community may have more opportunities to meet than people in different communities, for example. Finally, it can also depend on the observable characteristics x_i and x_j: for example, people with similar observable characteristics may meet more often. Similar biased meetings are used in Currarini et al. (2009) and Currarini et al. (2010). Mele (2017), Badev (2013) and Chandrasekhar and Jackson (2014) also consider variants of this meeting technology. Furthermore, rho could be a function of the size of the network n: it could be easier to meet people in small networks than in large networks, for example.

Upon meeting, i and j decide whether to update their link g_{ij}: if the link does not exists, they decide whether to form it; if the link already exists, they choose whether to delete it. When updating the link, players behave myopically, and do not consider how their decision affects the future evolution of the network. This modeling approach is similar to Nakajima (2007), Mele (2017), Mele and Zhu (2020), Badev (2013), Bala and Goyal (2000), Jackson and Watts (2001) among others. I assume that players i and j maximize their joint payoff, when updating a link; this decision rule is compatible with pairwise stability with transfers,

one of the most common equilibrium notions used in the network literature (Jackson (2008), Mele and Zhu (2020), Chandrasekhar and Jackson (2014)).

2.1. **Meeting technology.** The probability $\rho(g, z_i, z_j, x_i, x_j)$ governs the opportunity to create and delete links. To obtain a closed-form solution for the likelihood of a network, I impose the following assumptions on the function ρ. Let g_{-ij} denote the network g with the exclusion of link g_{ij}.

ASSUMPTION 1. *Conditional on the unobserved community structure z, the meetings are i.i.d. over time and the probability that i and j meet is*

$$\rho(g, z_i, z_j, x_i, x_j) = \rho(g_{-ij}, z_i, z_j, x_i, x_j) > 0 \tag{2}$$

I also assume that the sum of these probabilities over all possible pairs of players is one.

This assumption guarantees that any two pair of agents in the network can meet with positive probability. Additionally, the long-run stationary distribution of the network will be independent of ρ, because the meeting probability $\rho(g, z_i, z_j, x_i, x_j) = \rho(g_{-ij}, z_i, z_j, x_i, x_j)$ does not depedend on the existence of a link between i and j. The latter property is sufficient condition for detailed balance.

2.2. **Preferences.** Players' preferences are defined over networks, observable characteristics and unobservable types. Players receive payoffs from their direct connections, and externalities from common friends.

Let $U_i(g, x, z; \theta)$ denote the utility of player i from network g, observable characteristics x, community structure z and parameters $\theta = (\theta_u, \theta_v)$. Preferences are described by

$$U_i(g, x, z; \theta) = \sum_{j=1}^{n} g_{ij} \left[u_{ij}(\theta_u) + \sum_{r \neq i,j}^{n} g_{jr} g_{ri} v_{ijr}(\theta_v) \right] \tag{3}$$

where $u_{ij}(\theta_u) := u(x_i, x_j, z_i, z_j; \theta_u)$ and $v_{ij}(\theta_v) := v(x_i, x_j, x_r, z_i, z_j, z_r; \theta_v)$ are the payoffs of direct connections and common friends, respectively.

Player i receives a direct payoff $u(x_i, x_j, z_i, z_j; \theta_u)$ for each link she creates (when $g_{ij} = 1$). This payoff may depend on the unobservable type z_i and z_j: for example, a person may have a bias for people of the same type; it may also depend on the observable characteristics x_i and x_j: for example, preferences may be a biased in favor of links with people of the same race, gender, income level, etc. The payoff $u(x_i, x_j, z_i, z_j; \theta_u)$ includes both costs and benefits of direct connections, so it should be interpreted as *net benefit* of forming a link. We assume as in Jackson and Wolinsky (1996) that players pay a cost for direct links, but not indirect connections.

Player i receives an additional payoff $v(x_i, x_j, x_r, z_i, z_j, z_r; \theta_v)$ for each friend in common with j. In general, this payoff depends on observables and can vary across types.

Throughout the paper, the payoffs functional forms are restricted for tractability and identification purposes according to the following assumption.

ASSUMPTION 2. *Consider a triple of players i, j, and r. Let $\phi(ijr)$ be a permutation of ijr. The payoff from common neighbors $v_{ijr}(\theta_v)$ satisfies*

$$v_{ijr}(\theta_v) = v_{\phi(ijr)}(\theta_v) \tag{4}$$

for any permutation $\phi(ijr)$ of ijr.

This assumption does not mean that each triple of players receives the same payoff from common friends. Indeed, the assumption restricts the payoffs such that $v_{123} = v_{213} = v_{132} = v_{231} = v_{321} = v_{312}$. However, it allows $v_{123} \neq v_{124}$, for example. So different triples $\{123\}$ and $\{124\}$ are allowed to have different preferences for common friends.

2.3. **Preference shocks.** Players receive a joint matching shock to the preferences before choosing whether to update a link. The random shock models idiosynchratic reasons that

could affect the decision to link: for example, in a particular period, a player i could be in a bad mood and reject a link to another player j that would have generated positive surplus.

ASSUMPTION 3. *Players receive a logistic matching shock before updating their links, which is i.i.d. over time and across pairs.*

Assumption 3 is standard in discrete choice models (Heckman (1978)), and it is crucial to obtain the likelihood in closed-form (Mele (2017), Mele and Zhu (2020), Chandrasekhar and Jackson (2014)), to perform maximum likelihood or Bayesian estimation.

2.4. **Equilibrium.** The structure imposed by Assumptions 1, 2 and 3 implies that network formation is a potential game (Monderer and Shapley (1996), Mele (2017)). There exists a potential function that summarizes the deterministic incentives of all the players. Upon meeting, the incentives of the players to form a particular link g_{ij} are given by the surplus generated by forming the link and the matching shock

$$U_i(g,x,z;\theta) + U_j(g,x,z;\theta) - [U_i(g',x,z;\theta) + U_j(g',x,z;\theta)] + \varepsilon_{ij} \tag{5}$$

where g is a network where i and j have a link, that is $g_{ij} = 1$; and g' is the same network g, excluding the link between i and j, that is $g'_{ij} = 0$ and $g'_{-ij} = g_{-ij}$; and ε_{ij} is the preference shock. We can show that there exists an aggregate function of the network that summarizes the incentives of forming links, that is a function Q with property

$$Q(g,x,z;\theta) - Q(g',x,z;\theta) = U_i(g,x,z;\theta) + U_j(g,x,z;\theta) - [U_i(g',x,z;\theta) + U_j(g',x,z;\theta)] \tag{6}$$

The following proposition shows that such function exists and has a specific functional form.

PROPOSITION 1. *Conditional on the community structure z, the network formation game is a potential game and there exists an aggregate potential function Q that summarizes*

the incentives of the players to form links upon meeting

$$Q(g,x,z;\theta) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}u_{ij}(\theta_u) + \frac{2}{3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r\neq i,j}^{n} g_{ij}g_{jr}g_{ri}v_{ijr}(\theta_v) \tag{7}$$

Proof. See Appendix A. □

The results in Proposition 1 implies that when computing equilibria of the network formation game, all the relevant information for computing profitable deviations is encoded in the potential. Furthermore, in a model without stochastic matching shocks, the profitable deviations of i and j can be computed by using the difference in utility (the right-hand side of equation (6)) or equivalently the difference in potential (the left-hand side of equation (6)). The main consequence is that all the pairwise stable networks (with transfers) can be found as local maxima of the potential (Monderer and Shapley (1996), Mele (2017), Jackson and Watts (2001)).

The existence of a potential function is important because it guarantees existence of at least one equilibrium (Monderer and Shapley (1996), Jackson (2008), Mele (2017)). An additional practical advantage is that one can simulate the network formation process without keeping track of each player profitable deviations: all that information is already incorporated in the potental function, which is a scalar.

The network formation model is a finite state space Markov chain, because the number of networks is finite. The chain is irreducible and aperiodic, because any pair of players can meet and update their link, and there is always a positive probability to update, given the assumptions on the preference shock and the meetings. Therefore network sequence generated by the model converges to a unique stationary distribution in the long run.

PROPOSITION 2. *Under Assumptions 1-3 and conditional on the community structure z, the sequence of networks generated by the model is a Markov chain with unique stationary*

distribution $\pi(g, x, z; \theta)$*:*

$$\pi(g, x, z; \theta) = \frac{\exp\left[Q(g, x, z; \theta)\right]}{c(\theta, x, z)} \tag{8}$$

where $c(\theta, x, z) = \sum_{g' \in \mathcal{G}} Q(g', x, z; \theta)$ *is a normalizing constant, summing over the set* $\mathcal{G}$ *of all possible networks with* n *players.*

Proof. The proof is analogous and follows the same steps of Theorem 1 in Mele (2017), therefore it is omitted for brevity. □

In the long run, the Markov chain of networks will spend most time in network configurations that have high potential. The result in Proposition 1 shows that the local maxima of the potential are pairwise stable (with transfers); therefore in the long run the networks that are most likely to be observed are pairwise stable.

I assume that the network in the data is a draw from the stationary equilibrium of the model, therefore the distribution (8) is the likelihood of observing a particular network, conditional on the community structure.

3. Estimation

3.1. **Model specification.** In estimating the model, I consider several alternative specifications that parametrize the payoffs. The utility from direct links is parameterized as

$$u(x_i, x_j, z_i, z_j; \alpha, \beta) = \alpha_{z_i z_j} + \sum_{p=1}^{P} \beta_p f_p(x_i, x_j) \tag{9}$$

The first part of the utility is parameter $\alpha_{z_i z_j}$, which is allowed to vary by community. The second part is a function of covariates and does not vary across blocks. Example of possible

functions f_p's are:

$$
\begin{aligned}
f_p(x_i, x_j) &= |x_i - x_j|; \qquad f_p(x_i, x_j) = \mathbf{1}_{\{x_i = x_j\}} \\
f_p(x_i, x_j) &= \mathbf{1}_{\{x_i = x_j = a\}}; \qquad f_p(x_i, x_j) = x_i + x_j \\
f_p(x_i, x_j) &= x_i \cdot x_j
\end{aligned}
$$

This formulation of the net payoff from direct links assumes that the effect of observed and unobserved variables is additively separable.

I consider two main parameterizations for α in the empirical analysis:

(1) **Block-dependent** α: parameter $\alpha_{z_i z_j}$ is assumed to be $\alpha_{z_i z_j} = \alpha_k$ if $z_{ik} = 1$. This parameterization assumes that the net benefit of direct links varies with the unobserved type of the player.

(2) **Block-dependent** α**, restricted**: $\alpha_{z_i z_j} = \alpha_k$ if $z_{ik} = z_{jk} = 1$ and $\alpha_{z_i z_j} = \alpha_b$ otherwise. This alternative parameterization assumes that the net benefit of a direct link varies by the unobserved type of both individuals, in a restricted way. If both have the same type k, the parameter is α_k; otherwise, the parameter is α_b.

For the utility from transitivity v_{ijr}, I consider three alternatives:

(1) **No transitivity**: $v_{ijr} = 0$ for any ijr. This corresponds to a simple stochastic blockmodel (Airoldi et al., 2008; Mele et al., 2019).

(2) **Global transitivity**: the parameter γ does not change by community, $v_{ijr} = \gamma$ for all ijr. This is in line with the literature on ERGMs (Snijders, 2002; Mele, 2017; Butts, 2009; Chandrasekhar and Jackson, 2016).

(3) **Local transitivity**: $v_{ijr} = \gamma_k$ if $z_{ik} = z_{jk} = z_{rk} = 1$ and $v_{ijr} = 0$ otherwise. This parameterization is the same used in Schweinberger and Handcock (2015) and it assumes that players care about common neighbors only within the same community.

TABLE 1. Parameterizations used in the empirical analysis

	global transitivity	local transitivity	no transitivity (sbm)
block dependent	$\alpha_{z_i z_j} = \alpha_k$ if $z_{ik} = 1$ $v_{ijr} = \gamma$ for all ijr	$\alpha_{z_i z_j} = \alpha_k$ if $z_{ik} = 1$ $v_{ijr} = \begin{cases} \gamma_k & \text{if } z_{ik} = z_{jk} = z_{rk} = 1 \\ 0 & \text{otherwise} \end{cases}$	$\alpha_{z_i z_j} = \alpha_k$ if $z_{ik} = 1$ $v_{ijr} = 0$ for all ijr
restricted	$\alpha_{z_i z_j} = \begin{cases} \alpha_k & \text{if } z_{ik} = z_{jk} = 1 \\ \alpha_b & \text{otherwise} \end{cases}$ $v_{ijr} = \gamma$ for all ijr	$\alpha_{z_i z_j} = \begin{cases} \alpha_k & \text{if } z_{ik} = z_{jk} = 1 \\ \alpha_b & \text{otherwise} \end{cases}$ $v_{ijr} = \begin{cases} \gamma_k & \text{if } z_{ik} = z_{jk} = z_{rk} = 1 \\ 0 & \text{otherwise} \end{cases}$	$\alpha_{z_i z_j} = \begin{cases} \alpha_k & \text{if } z_{ik} = z_{jk} = 1 \\ \alpha_b & \text{otherwise} \end{cases}$ $v_{ijr} = 0$ for all ijr

Table 1 summarizes the combinations of these parameterizations.

The models include the following covariates: race, gender, grade and parental income. The final specification of the utility function is

$$
\begin{aligned}
u_(x_i, x_j, z_i, z_j; \alpha, \beta) &= \alpha_{z_i z_j} + \beta_{w,w}\mathbf{1}_{\{race_{ij}=white\}} + \beta_{b,b}\mathbf{1}_{\{race_{ij}=black\}} + \beta_{h,h}\mathbf{1}_{\{race_{ij}=hisp\}} \\
&+ \beta_{7,7}\mathbf{1}_{\{grade_{ij}=7\}} + \beta_{8,8}\mathbf{1}_{\{grade_{ij}=8\}} + \beta_{9,9}\mathbf{1}_{\{grade_{ij}=9\}} + \beta_{10,10}\mathbf{1}_{\{grade_{ij}=10\}} \\
&+ \beta_{11,11}\mathbf{1}_{\{grade_{ij}=11\}} + \beta_{12,12}\mathbf{1}_{\{grade_{ij}=12\}} \\
&+ \beta_{m,m}\mathbf{1}_{\{gender_{ij}=male\}} + \beta_{f,f}\mathbf{1}_{\{gender_{ij}=female\}} + \beta_{income}|income_i - income_j|
\end{aligned} \tag{10}
$$

where $race_{ij}$ is a variable that records the race of both i and j; $grade_{ij}$ is the grade of the pair ij and $gender_{ij}$ is the gender of i and j. Our specification allows for homophily in race, gender, grade and (parental) income. Homophily in unobservables is captured by $\alpha_{z_i z_j}$.

3.2. **Likelihood factorization with local transitivity.** The potential function $Q(g, x, z, \theta)$ can be decomposed in *within*- and *between*-community potentials when we adopt the specification with *local transitivity*. Let $g_{k,l}$ denote the subnetwork formed by individuals of communities $\mathcal{C}_k$ and $\mathcal{C}_l$. Let $x^{(k)}$ denote the covariates of players in community $\mathcal{C}_k$. The potential can be decomposed into the sum of sub-potentials for the sub-networks $g_{k,l}$'s and the likelihood can be written as a factorized distribution. Lemma 1 in Appendix, proves that the potential $Q(g, x, z, \theta)$ can be decomposed as the sum of subpotentials $Q_{k,l}(g_{k,l}, x^{(k)}, x^{(l)}, z; \theta)$,

separating the within-community and between-community contributions as follows:

$$Q(g,x,z;\theta) = \sum_{k=1}^{K} Q_{k,k}(g_{k,l}, x^{(k)}, z;\theta) + \sum_{k=1}^{K}\sum_{l>k}^{K} Q_{k,l}(g_{k,l}, x^{(k)}, x^{(l)}, z;\theta) \tag{11}$$

$$= \sum_{k=1}^{K}\left[\sum_{i\in\mathcal{C}_k}\sum_{j\in\mathcal{C}_k} g_{ij}u(x_i,x_j,z_i,z_j;\alpha,\beta) + \frac{2\gamma_k}{3}\sum_{i\in\mathcal{C}_k}\sum_{j\in\mathcal{C}_k}\sum_{r\in\mathcal{C}_k} g_{ij}g_{jr}g_{ri}\right] \tag{12}$$

$$+ \sum_{k=1}^{K}\sum_{l>k}^{K}\left[\sum_{i\in\mathcal{C}_k}\sum_{j\in\mathcal{C}_l} g_{ij}u(x_i,x_j,z_i,z_j;\alpha,\beta)\right] \tag{13}$$

The main consequence of this decomposition is that the likelihood is factorized as

$$\pi(g,x,z;\theta) = \prod_{k=1}^{K}\frac{\exp\left[Q_{k,k}(g_{k,k},x^{(k)},z;\theta)\right]}{c_{k,k}(\mathcal{G}_{k,k},x^{(k)};\theta)}\left[\prod_{l>k}^{K}\frac{\exp\left[Q_{k,l}(g_{k,l},x^{(k)},x^{(l)},z;\theta)\right]}{c_{k,l}(\mathcal{G}_{k,l},x^{(k)},x^{(l)};\theta)}\right] \tag{14}$$

where the within-community normalizing constant is

$$c_{k,k}(\mathcal{G}_{k,k},x^{(k)};\theta) = \sum_{\omega_{k,k}\in\mathcal{G}_{k,k}} \exp\left[Q_{k,k}(\omega_{k,k},x^{(k)},z;\theta)\right] \tag{15}$$

Notice that since $Q_{k,l}(g_{k,l},x^{(k)},x^{(l)},z;\theta) = \sum_{i\in\mathcal{C}_k}\sum_{j\in\mathcal{C}_l} g_{ij}u(x_i,x_j,z_i,z_j;\alpha,\beta)$ the second part of the likelihood (14) can be written as the product of Bernoulli links,

$$\prod_{l>k}^{K}\frac{\exp\left[Q_{k,l}(g_{k,l},x^{(k)},x^{(l)},z;\theta)\right]}{c_{k,l}(\mathcal{G}_{k,l},x^{(k)},x^{(l)};\theta)} = \prod_{l>k}^{K}\prod_{i\in\mathcal{C}_k}\prod_{j\in\mathcal{C}_l}\frac{\exp\left[g_{ij}(u(x_i,x_j,z_i,z_j;\alpha,\beta)+u(x_j,x_i,z_j,z_i;\alpha,\beta))\right]}{1+\exp\left[u(x_i,x_j,z_i,z_j;\alpha,\beta)+u(x_j,x_i,z_j,z_i;\alpha,\beta)\right]}$$

To summarize, this parameterization of the model with local transitivity generates independence among between-commmunities links, while the links generated within-community may have strong dependence. The final result is a model that mantains the complex dependence structure of the exponential family random graphs locally, while allowing for weak dependence among links globally. This model is compatible with the exponential random graphs with local dependence developed in Schweinberger and Handcock (2015).

3.3. **Priors.** The main challenge in estimation of the model is that the community structure is unknown. The Bayesian approach is to impose a prior on the number of communities and use that to estimate the posterior. I follow Schweinberger and Handcock (2015) and use their nonparametric priors to model the communities. The likelihood of the model can be written as

$$L\left(g, Z; \theta, \eta, x\right) = \sum_{z \in \mathcal{Z}} P_\theta\left(G = g | X = x, Z = z\right) P_\eta\left(Z = z\right) \tag{16}$$

The first part is the likelihood of observing network g in the long run, given the covariates x and the community structure z. For the community structure, I follow a standard assumption in the stochastic blockmodels literature (Airoldi et al. (2008), Schweinberger and Handcock (2015)), and use a multinomial distribution with probability of membership $\eta_k \in [0, 1]$ for $k = 1, .., K$ and $\sum_{k=1}^{K} \eta_k = 1$,

$$Z_i | \eta_1, ..., \eta_K \overset{iid}{\sim} Multinomial\left(1; \eta_1, ...\eta_K\right) \text{ for } i = 1, ..., n \tag{17}$$

Following Ishwaran and James (2001) and Schweinberger and Handcock (2015), I use the following nonparametric prior for η_k, $k = 1, 2, 3, ..., K$,

$$\eta_1 = V_1 \tag{18}$$

$$\eta_k = V_k \prod_{j=1}^{k-1} \left(1 - V_j\right) \qquad k = 2, 3, 4, ... \tag{19}$$

$$V_k | \phi \overset{iid}{\sim} Beta(1, \phi) \qquad k = 1, 2, 3, ... \tag{20}$$

$$\phi > 0 \quad \text{and} \quad \sum_{k=1}^{K} \eta_k = 1 \ w.p.1 \tag{21}$$

As a practical matter, the estimation of the posterior is computationally expensive, especially when the number of communities K is large. The priors for the payoffs are multivariate normals, because there is no practical advantage from using conjugate priors that also depend

on normalizing constant. For the specification with *local transitivity* I use

$$(\alpha_k, \gamma_k)|\mu_w, \Sigma_w \sim MVN(\mu_w, \Sigma_w) \text{ for } k = 1, .., K_{max} \tag{22}$$

$$\beta|\mu_\beta, \Sigma_\beta \sim MVN(\mu_\beta, \Sigma_\beta) \tag{23}$$

while for the *global transitivity* specification I use

$$\alpha_k|\mu_w, \Sigma_w \sim MVN(\mu_w, \Sigma_w) \text{ for } k = 1, .., K_{max} \tag{24}$$

$$(\beta, \gamma)|\mu_\beta, \mu_\gamma, \Sigma \sim MVN\left(\begin{pmatrix}\mu_\beta \\ \mu_\gamma\end{pmatrix}, \begin{pmatrix}\Sigma_\beta & \Sigma_{\beta,\gamma} \\ \Sigma_{\beta,\gamma} & \Sigma_\gamma\end{pmatrix}\right) \tag{25}$$

In the restricted specification the prior for the between-block parameter α_b is

$$\alpha_b|\mu_b, \Sigma_b \sim MVN(\mu_b, \Sigma_b) \tag{26}$$

I adopt a hierarchical approach and specify hyper-priors for each parameter, following Schweinberger and Handcock (2015). For details see Appendix.

3.4. **Posterior and MCMC algorithm.** The complex form of the likelihood does not allow direct sampling from the posterior. I rely on the exchange MCMC method developed in Murray et al. (2006) and Liang (2010) and adapted to network models in Caimo and Friel (2011), Atchade and Wang (2014), Schweinberger and Handcock (2015) and Mele (2017). To be concrete, consider the specification with restricted block-dependent parameters and local transitivity. The posterior distribution can be written as follows

$$\begin{aligned} p(\phi, \mu_w, \Sigma_w, \mu_b, \Sigma_b, \mu_\beta, \Sigma_\beta, \eta, \alpha, \beta, \gamma, z|g, x) &\propto p(\phi, \mu_w, \Sigma_w, \mu_b, \Sigma_b, \mu_\beta, \Sigma_\beta, \eta, \alpha, \beta, \gamma) \\ &\times P_\eta(Z = z) P_\theta(G = g|X = x, Z = z) \end{aligned} \tag{27}$$

where $p(\phi, \mu_w, \Sigma_w, \mu_b, \Sigma_b, \mu_\beta, \Sigma_\beta, \eta, \alpha, \beta, \gamma)$ is the prior distribution. The prior is assumed to factorize in the following form

$$
\begin{aligned}
(28) \quad p(\phi, \mu_w, \Sigma_w, \mu_b, \Sigma_b, \mu_\beta, \Sigma_\beta, \eta, \alpha, \beta, \gamma) &= p(\phi)p(\mu_w)p(\Sigma_w)p(\mu_b)p(\Sigma_b)p(\mu_\beta)p(\Sigma_\beta) \\
&\times \; p(\eta|\phi)p(\alpha_b|\mu_b, \Sigma_b)p(\beta|\mu_\beta, \Sigma_\beta) \\
&\times \; \left[\prod_{k=1}^{K_{max}} p(\alpha_k, \gamma_k|\mu_w, \Sigma_w)\right]
\end{aligned}
$$

The details of the sampler are provided in Murray et al. (2006), Mele (2017), Caimo and Friel (2011), Liang (2010) and Schweinberger and Handcock (2015) and are implemented in the package `hergm` in `R` (Schweinberger and Luna, forthcoming).

The exchange algorithm for this model is slightly different from the original Murray et al. (2006) and Liang (2010)'s sampler used in Mele (2017). The main additional complication is that the communities are unknown, and therefore need to be treated as an additional parameter in the sampler. The algorithm augments the posterior parameters with auxiliary variables g^*, z^* and $\theta^* := (\alpha^*, \beta^*, \gamma^*)$, proceeding with the following steps at each iteration:

(1) Sample (θ^*, z^*) from auxiliary distribution $q(\theta^*, z^*|\eta, \theta, z, g)$
(2) Sample g^* from $\pi(\omega, x, z^*; \theta^*)$ using the Metropolis-Hastings sampler of Mele (2017);
(3) Propose to swap (θ, z) with (θ^*, z^*), accepting with probability $\min\{1, exch\}$, where $exch$ is

$$
\begin{aligned}
(29) \qquad exch = &\frac{P_\eta(Z = z^*)}{P_\eta(Z = z)} \frac{q(\theta, z|\eta, \theta^*, z^*, g)}{q(\theta^*, z^*|\eta, \theta, z, g)} \frac{\pi(g, x, z^*; \theta^*)}{\pi(g, x, z; \theta)} \frac{\pi(g^*, x, z; \theta)}{\pi(g^*, x, z^*; \theta^*)} \\
&\times \frac{\prod_{k=1}^{K_{max}} p(\alpha_k^*, \gamma_k^*|\mu_w, \Sigma_w)}{\prod_{k=1}^{K_{max}} p(\alpha_k, \gamma_k|\mu_w, \Sigma_w)}
\end{aligned}
$$

and $P_\eta(Z = z)$ is the multinomial distribution that generates the community structure, $\pi(g, x, z; \theta)$ is the stationary distribution of the model conditional on community structure, and $p(\alpha_k, \gamma_k|\mu_w, \Sigma_w)$ are the priors.

The practical implication of the formula for acceptance probability $exch$, is that the normalizing constants included in the discrete exponential distribution cancel out. Therefore the sampling using the exchange algorith is feasible, while sampling from the posterior using standard Metropolis or Gibbs sampler is infeasible. For a formal discussion see Mele (2017), Appendix B.

The auxiliary distribution $q(\theta^*, z^*|\eta, \theta, z, g)$ proposes θ^* that are Gaussians centered at θ and z^* that are generated from the full conditional distribution of the community memberships. The reason for such updates is that these local moves do not lead to a very high rejection rate of the exchange algorithm, as pointed out in Caimo and Friel (2011) and Mele (2017).

An additional challenge is that the likelihood of this model is invariant to permutations of the community labels. This problem is common in the literature on finite mixture models, where the likelihood is invariant to permutations of the labels of the mixture's components (Gelman et al. (2003), McLachlan and Peel (2000), Stephens (2000)). This complicates inference for the community-specific parameters, because the community labels may switch several times during the MCMC simulation. While the use of nonparametric priors partially alleviate this problem, as they make the posterior not invariant to the labeling, the issue may present itself in the MCMC simulations. Therefore I follow the Bayesian literature on finite mixtures and use a relabeling algorithm, first developed in Stephens (2000) and suggested by Schweinberger and Handcock (2015) for this class of models. This procedure makes the label of the blocks consistent through the MCMC simulation output and is convenientely implemented in R. Details are provided in Appendix.

3.5. **A note on identification.** While general results on parameters identification in this class of models are rare, there is some work on special cases of this statistical model. If the block structure is known or perfectly recovered, then the parameters that do not vary by block are identifiable by variation of the sufficient statistics across blocks. Additionally, if the

block structure is known, consistent estimation of block-dependent parameters is possible (Schweinberger and Stewart, 2020). However, (Schweinberger and Stewart, 2020) assumes parameters are restricted as $\alpha_k = \alpha \log(n_k)$, where n_k is the number of nodes in block k.

In our hierarchical Bayesian approach we can identify the expected value of the prior for block-dependent parameters. Using a relabeling algorithm and the nonparametric priors discussed above, the posterior will not suffer from invariance of the block labels. Therefore, we can also intepret the block-specific parameters.

Finally, Schweinberger (2020) shows that it is possible to estimate the block structure consistently under assumptions on smoothness and weak dependence. To obtain consistent estimates of the block structure, the model is restricted to the case where the exponential family graph is identified. In appendix, I use Monte Carlo simulations with small networks to show that block recovery for the model with local transitivity is possible and has similar performance as the more standard stochastic blockmodel, assuming the model is correctly specified.

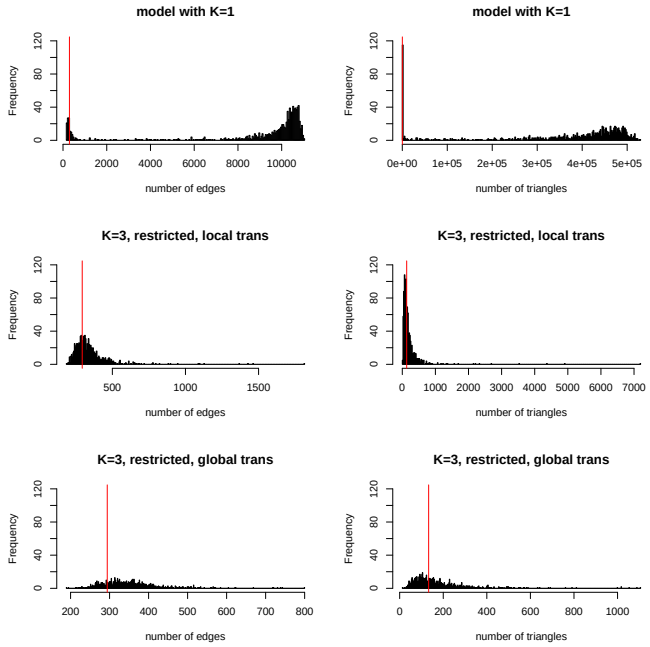
4. Empirical Results

I use data from school 28 and Wave I (1994) of in Add Health, which contains complete data on friendship networks and demographics of the students. This school contains 150 students with 58.7% females, and there are 294 friendship links. The school is racially heterogeneous: 42% Whites/Caucasians, 45.3% African-Americans, 0.7% Asians, 10.7% Hispanics and 1.3% Other race. The racial fragmentation index is 0.606. The school offers all grades from 7 to 12, with a relatively balanced population among the different age groups, respectively 17.3%, 17.3%, 20%, 16.7%, 14%, and 14.7%. The level of racial segregation, measured according to the Freeman (1972) index is 0.72 for Whites/Caucasians, 0.764 for African Americans, and 0.429 for Hispanics. The segregation by gender is 0.255.

The estimated model includes controls for homophily in race, gender, grade and parental income. These are some of the variables that are considered good predictors of friendships

during adolescence (Moody, 2001; Mayer and Puller, 2008; Boucher, 2015; Calvo-Armengol et al., 2009).

FIGURE 1. Posterior predictions of edges and triangles, comparison of three specifications.



The figure reports simulated posterior predictions for number of edges and triangles for three models: no block structure ($K = 1$) in the first row; local transitivity with $K = 3$ in the second row; global transitivity and $K = 3$ in the third row. The vertical red line is the observed value in the data. The models with unobserved heterogeneity put most of the posterior probability mass around the observed sufficient statistics, while the model with only one block has bi-modal distribution of the posterior predictions.

4.1. **Choosing the number of communities K and specification.** There is no established criterion to select a specification and the number of blocks K for this class of models. Therefore, I follow an heuristic approach and select specification and K using a mix of goodness-of-fit tests borrowed from the ERGM literature (Snijders, 2002; Koskinen, 2008). This is also the approach suggested in Schweinberger and Handcock (2015) for hierarchical ERGMs.

In practice, I estimate each specification in Table 1, with an increasing $K = \{1, 2, 3, 4, 5, ...\}$. Using the estimated posterior of the parameters, I simulate many networks and check whether the estimated model is able to replicate some empirical properties of the network in the data.

Crucially, I check whether at a minimum, the model is able to replicate the number of links and triangles, and other features such as number of stars degree distribution and homophily levels. I choose the most parsimonious model, that is the one with the smallest K. More details, tables and figures are shown in Appendix.

The first comparison is with a model that has no block structure, that is $K = 1$. Such model is unable to replicate the aggregate features of the data. In Figure 1 I show the ability of three different models to fit the most important sufficient statistics of the data, number of edges and number of triangles. In the first row I report the histogram of posterior predictions of a model with no community structure ($K = 1$). The observed value in the data is the vertical red line. The posterior predictions are bi-modal, showing a pattern that many practitioners of the ERGM literature have encountered in empirical applications (Snijders, 2002; Chandrasekhar and Jackson, 2016; Mele, 2017). This model is unable to replicate the feature of the data, providing predictions that generate an almost empty or almost complete network in simulations. The second row and third row show the results of the structural model with unobserved heterogeneity and $K = 3$. By comparison these models' posterior predictions impose most probability mass around the observed value of the sufficient statistics. The model with local transitivity (second row) seems to perform slightly better (last row).

Since the model without unobserved block structure performs poorly, I select my specification among the models with $K > 1$. Following the procedure of Schweinberger and Handcock (2015) immediately rules out the stochastic blockmodel specifications with $v_{ijr} = 0$ for any ijr. Indeed stochastic blockmodels underestimate the number of triangles in each specification and for any number of blocks K (see Figure 2 in Appendix). The specification with unrestricted block dependence for α underestimate the number of 2-stars. Therefore the specifications that fit best the current data are the ones with restricted block-dependence for α.

Based on the RMSE's for posterior predictions of edges, two-stars and triangles in Tables 3, 4 and 5 in Appendix, the model with restricted block-dependence and global transitivity seems to provide a better RMSE for these data. The conclusion is the same if we consider using Median Absolute Deviation (MAD) instead of RMSE.

The chosen specification is a model with restricted block-dependence and local transitivity, with $K = 3$ blocks. In appendix B and the online appendix I report the results for all the estimated specifications.

4.2. **Estimated structural parameters.** The estimated structural parameters are shown in Table 2, where I report the specification with $K = 3$. I include mean, standard deviation, median and 2.5% and 97.5% quantiles of the posterior distribution.

The top panel reports the estimated prior mean and standard deviation for the block-specific parameters, $(\mu_{w,\alpha}, \sigma_{w,\alpha})$ for α_k with $k = 1, 2, 3$ and $(\mu_{w,\gamma}, \sigma_{w,\gamma})$ for γ_k's. This parameters are identified in the hierarchical Bayes approach, as well as the parameters in Panel B, that estimate homophily/heterophily in observed characteristics.

Most block-specific parameters are estimated precisely, with the exception of γ_3, the payoff from common friends in the third community; and the parameters relative to homophily by gender ($\beta_{male,male}$ and $\beta_{female,female}$).

In panel A, I show the marginal posterior for the community-specific payoffs. The estimated α's are negative; the net community-specific payoff of links across communities α_b is higher than the net community-specific payoff within communities $(\alpha_1, \alpha_2, \alpha_3)$ in absolute value. While net payoffs of forming links within communities do not differ significantly for students of type 1 and 2, individuals in community 3 seem relatively more social on average.

Panel B reports the estimates for terms including covariates. There is homophily by race, as the marginal utility of a link increases when players form a link with a student of the same racial group. The same effect is present for grade. The estimates for gender are close

to zero and very widely spread. I also estimate homophily by income, as the coefficient $\beta_{|income_i - income_j|}$ is negative.

Panel C shows the estimates for payoffs from common friends. An additional common friend is more valuable for students of type 2 than type 1. For student of type 3 the estimate is not very precise, but it is nonetheless positive on average.

I conclude that there is strong homophily by observable characteristics, especially race, grade and income. In addition, the estimated marginal posteriors show that there is some heterogeneity across communities in both net payoffs from direct links and payoffs from common friends. In particular there is a significant different in the payoff of forming links across communities and within community. Community 3 seems to be the most social: student in this community have either a lower cost or a higher benefit of forming links. Students in community 2 care more about common friends than the average student. However, these block-specific parameters have large posterior variability.

4.3. **Fit of the model.** The model fit is relatively good: our posterior predictions are able to match the observed links, triangles and homophily. In appendix B, I show more details about all the estimated models and specifications with additional goodness of fit figures. Figure 2 shows the histogram of posterior predictions for the number of links in the network and the number of triangles. I simulated 1000 realizations of the network, drawing from the posterior distribution of the parameters. The observed number of links is 294 and the posterior mean prediction is 346.5, with median prediction of 317.

It is well known that the number of triangles, is the most difficult statistics to match. Diaconis and Chatterjee (2013) and Mele (2017) show that in exponential-family random graphs the number of triangles tends to be degenerate, either very close to zero or very close to the maximum number. In this network, there are 133 triangles and the posterior mean prediction is 225.6, with a median prediction of 133. There are some extreme values in the posterior simulations but the general fit is good.

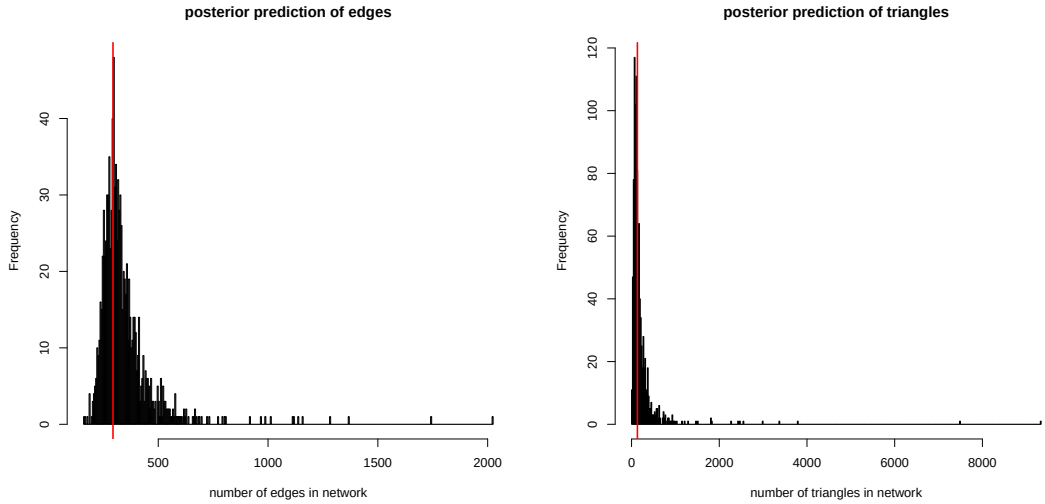
TABLE 2. Estimated posterior of the structural parameters ($K = 3$)

Parameter	**Post. mean**	**Post. s.d.**	**Posterior quantiles** 2.5%	50%	97.5%
ϕ			0.117	0.686	2.130
$\mu_{w,\alpha}$			-3.729	-2.501	-1.114
$\mu_{w,\gamma}$			-0.252	0.874	2.035
$\sigma_{w,\alpha}$			0.406	0.849	1.585
$\sigma_{w,\gamma}$			0.512	1.021	1.767
A. Community-specific payoff					
α_1	-4.070	0.464	-4.888	-4.086	-3.091
α_2	-3.854	0.587	-4.883	-3.895	-2.589
α_3	-2.527	1.049	-4.385	-2.609	-0.316
α_b	-5.754	0.455	-6.636	-5.763	-4.837
B. Payoff from covariates					
$\beta_{white,white}$	1.002	0.246	0.500	1.017	1.420
$\beta_{black,black}$	0.923	0.252	0.424	0.938	1.364
$\beta_{hisp,hisp}$	1.965	0.628	0.789	1.920	3.128
$\beta_{grade7,grade7}$	1.371	0.290	0.685	1.409	1.831
$\beta_{grade8,grade8}$	1.321	0.311	0.627	1.327	1.892
$\beta_{grade9,grade9}$	1.203	0.332	0.568	1.172	1.883
$\beta_{grade10,grade10}$	1.140	0.446	0.207	1.127	1.929
$\beta_{grade11,grade11}$	1.241	0.433	0.249	1.291	1.973
$\beta_{grade12,grade12}$	1.029	0.281	0.435	1.033	1.562
$\beta_{male,male}$	-0.061	0.297	-0.689	-0.029	0.450
$\beta_{female,female}$	-0.170	0.254	-0.725	-0.135	0.294
$\beta_{\lvert income_i - income_j \rvert}$	-0.588	0.278	-1.208	-0.568	-0.136
C. Payoff from common friends					
γ_1	0.969	0.149	0.644	0.977	1.244
γ_2	1.573	0.562	0.508	1.561	2.738
γ_3	0.995	0.948	-0.889	0.969	2.920

The estimates are obtained from a run of 100,000 steps of the exchange algorithm, collecting a posterior sample of 8000 draws. Panel A shows the estimates of community-specific net payoffs; Panel B shows the estimates of the homophily terms; Panel C shows the estimates for the common friends' payoffs. I report mean, standard deviation, median, the 2.5% and 97.5% quantiles.

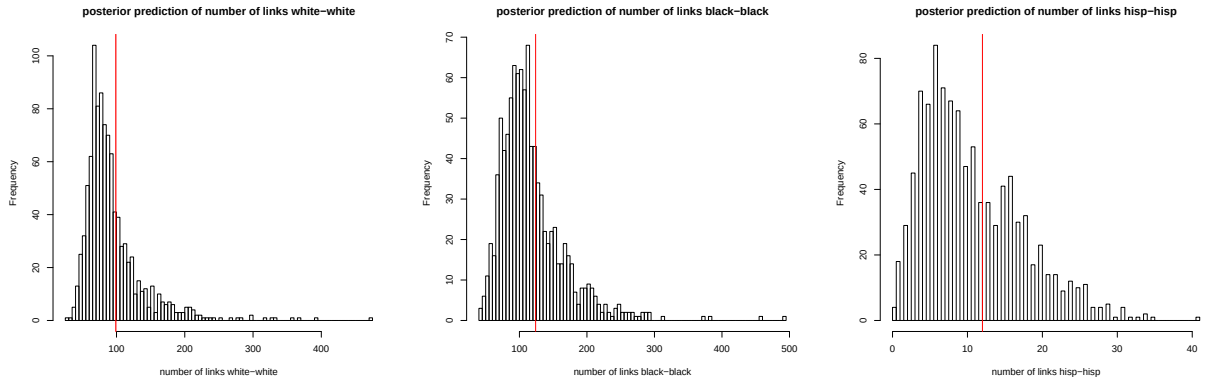
Figure 3 shows that the model is able to replicate also the homophily by race. I report the histograms of posterior predictions for the observed number of friendships among whites, blacks and hispanics in the school. The red vertical line is the observed value. As in the figures above, while there are some extreme values, the fit of the model is very good.

FIGURE 2. Posterior predictions of number of links and triangles in the network



The posterior predictions are obtained by a 1000 simulations from the posterior estimated in Table 2. The red line is the value observed in the data

FIGURE 3. Posterior predictions for racial homophily



The posterior predictions are obtained by a 1000 simulations from the posterior estimated in Table 2. The red line is the value observed in the data.

The observed number of friendships among whites, african-american and hispanic students are 99, 124 and 12, respectively. The predicted posterior means are 94.9, 118.4 and 10.8; the predicted posterior medians are 83, 108 and 9 respectively.

I conclude that the estimated model is able to replicate the empirical aggregate features of the network, both in terms of triadic closure and homophily.

5. Conclusion

This paper has developed a structural model of network formation with observed and unobserved player heterogeneity, generating homophily and clustering in equilibrium. Players belongs to different unobserved communities that affect their cost and benefits of linking. Agents care about the composition of their links and the number of common friends (transitivity), and update their links upon meeting. A sequential meeting technology generates a sequence of meetings and link updates that in the long-run converges to a finite mixture of discrete exponential family random graphs.

The model is fairly general and allows many different specifications of the payoffs. In particular, the payoff from common friends could be modeled as a local (only matters within blocks, as in (Schweinberger and Handcock, 2015)) or a global parameter (it is the same within and across blocks, as in the ERGM literature (Mele, 2017; Snijders, 2002)). I also estimate models without transitivity payoff, which correspond to a stochastic blockmodel (Airoldi et al., 2008).

I show that the model with community structure is able to match the empirical features of Add Health school friendship data, while also providing insights on the economics of network formation. A model with no community structure cannot fit the number of triangles in the observed network, while a model with few communities makes a great job at matching the aggregate number of triangles and other sufficient statistics.

My empirical results suggest that students in different unobserved communities have different payoffs of direct links and have different payoffs from common friends. This adds to the literature that shows how unobserved characteristics may affect network formation as much as observables (Graham, 2017; Schweinberger and Handcock, 2015; Mele et al., 2019). Therefore it is important to develop models of empirical network formation that can deal with unobserved heterogeneity.

A practical limitation of the approach is the computationl burden of estimation. My Bayesian approach via exchange algorithm is not easily scalable to large networks. While some of these issues can be partially attenuated by more efficient algorithms that exploit the parallel nature of the MCMC samplers, an alternative approach to estimation relies on approximate likelihood methods and recently developed two-step methods for unobserved heterogeneity (Schweinberger, 2020; Bonhomme et al., 2017). I do not explore such methods in this work, but my empirical and simulation results suggest that this modeling strategy has clear benefits when compared to models without unobserved heterogeneity.

References

Acemoglu, D., M. Dahleh, I. Lobel and A. Ozdaglar (2011), 'Bayesian learning in social networks', *Review of Economic Studies* **78**(4), 1201–1236.

Airoldi, Edoardo M., David Blei, Stephen E. Fienberg and Eric P. Xing (2008), 'Mixed membership stochastic blockmodels', *Journal of Machine Learning* **9**, 1981–2014.

Atchade, Yves and Jing Wang (2014), 'Bayesian inference of exponential random graph models for large social networks', *Communications in Statistics - Simulation and Computation* **43**(2), 359–377.

Babkin, Sergei, Jonathan Stewart, Xiaochen Long and Michael Schweinberger (2020), Large-scale estimation of random graph models with local dependence. working paper, https://arxiv.org/pdf/1703.09301.pdf.

Badev, Anton (2013), Discrete games in endogenous networks: Theory and policy.

Bala, Venkatesh and Sanjeev Goyal (2000), 'A noncooperative model of network formation', *Econometrica* **68**(5), 1181–1229.

Bonhomme, Stephane, Thibaut Lamadon and Elena Manresa (2017), Discretizing unobserved heterogeneity. Working Paper.

Boucher, Vincent (2015), 'Structural homophily', *The International Economic Review* **56**(1), 235–264.

Boucher, Vincent and Ismael Mourifie (2017), 'My friend far far away: A random field approach to exponential random graph models', *Econometrics Journal* **20**(3), S14–S46.

Breza, Emily, Arun Chandrasekhar, Tyler Mccormik and Mengjie Pan (2017), Using aggregated relational data to feasibly identify network structure without network data.

Butts, Carter (2009), Using potential games to parameterize erg models. working paper.

Caimo, Alberto and Nial Friel (2011), 'Bayesian inference for exponential random graph models', *Social Networks* **33**(1), 41–55.

Calvo-Armengol, Antoni, Eleonora Patacchini and Yves Zenou (2009), 'Peer effects and social networks in education', *Review of Economic Studies* **76**, 1239–1267.

Chandrasekhar, Arun G. (2016), *Oxford handbook on the economics of networks.*, Oxford Univerisity Press, chapter Econometrics of network formation.

Chandrasekhar, Arun and Matthew Jackson (2014), Tractable and consistent exponential random graph models.

Chandrasekhar, Arun and Matthew Jackson (2016), A network formation model based on subgraphs. working paper.

Charbonneau, Karyne B. (2017), 'Multiple fixed effects in binary response panel data models', *Econometrics Journal* **20**(3), S1–S13.

Conley, Timothy and Christopher Udry (2010), 'Learning about a new technology: Pineapple in ghana', *American Economic Review* **100**(1), 35–69.

Currarini, Sergio, Matthew O. Jackson and Paolo Pin (2009), 'An economic model of friendship: Homophily, minorities, and segregation', *Econometrica* **77**(4), 1003–1045.

Currarini, Sergio, Matthew O. Jackson and Paolo Pin (2010), 'Identifying the roles of race-based choice and chance in high school friendship network formation', *the Proceedings of the National Academy of Sciences* **107**(11), 4857–4861.

De Giorgi, Giacomo, Michele Pellizzari and Silvia Redaelli (2010), 'Identification of social interactions through partially overlapping peer groups', *American Economic Journal: Applied Economics* **2**(2).

DePaula, Aureo (2017), Econometrics of network models, *in* B.Honore, A.Pakes, M.Piazzesi and L.Samuelson, eds, 'Advances in Economics and Econometrics: Eleventh World Congress', Cambridge University Press.

DePaula, Aureo, Seth Richards-Shubik and Elie Tamer (2018), 'Identifying preferences in networks with bounded degree', *Econometrica* **86**(1), 263–288.

Diaconis, Persi and Sourav Chatterjee (2013), 'Estimating and understanding exponential random graph models', *Annals of Statistics* **41**(5), 2428–2461.

Dzemski, Andreas (2017), An empirical model of dyadic link formation in a network with unobserved heterogeneity. Working Paper.

Echenique, Federico and Roland Fryer (2007), 'A measure of segregation based on social interactions', *Quarterly Journal of Economics* **122**(2), 441–485.

Fafchamps, Marcel and Flore Gubert (2007), 'Risk sharing and network formation', *American Economic Review Papers and Proceedings* **97**(2), 75–79.

Freeman, L. (1972), 'Segregation in social networks', *Sociological Methods and Research* **6**, 411–427.

Galeotti, Andrea (2006), 'One-way flow networks: the role of heterogeneity', *Economic Theory* **29**(1), 163–179.

Gelman, A., J. Carlin, H. Stern and D. Rubin (2003), *Bayesian Data Analysis, Second Edition*, Chapman & Hall/CRC.

Geyer, Charles and Elizabeth Thompson (1992), 'Constrained monte carlo maximum likelihood for depedendent data', *Journal of the Royal Statistical Society, Series B (Methodological)* **54**(3), 657–699.

Goldsmith-Pinkham, Paul and Guido W. Imbens (2013), 'Social networks and the identification of peer effects', *Journal of Business and Economic Statistics* **31**(3), 253–264.

Golub, Benjamin and Matthew Jackson (2011), 'Network structure and the speed of learning: Measuring homophily based on its consequences', *Annals of Economics and Statistics* .

Graham, Bryan (2016), Homophily and transitivity in dynamic network formation. working paper.

Graham, Bryan (2017), 'An empirical model of network formation: with degree heterogeneity', *Econometrica* **85**(4), 1033–1063.

Heckman, James J. (1978), 'Dummy endogenous variables in a simultaneous equation system', *Econometrica* **46**(4), 931–959.

Hsieh, Chih-Sheng and Lung-Fei Lee (2012), A structural modeling approach for network formation and social interactions with applications to students' friendship choices and selectivity on activities.

Ishwaran, Hemant and Lancelot F James (2001), 'Gibbs sampling methods for stick-breaking priors', *Journal of the American Statistical Association* **96**(453), 161–173.

Jackson, Matthew and Alison Watts (2001), 'The existence of pairwise stable networks', *Seoul Journal of Economics* **14**(3), 299–321.

Jackson, Matthew and Asher Wolinsky (1996), 'A strategic model of social and economic networks', *Journal of Economic Theory* **71**(1), 44–74.

Jackson, Matthew O. (2008), *Social and Economics Networks*, Princeton.

Jackson, Matthew O. and Brian W. Rogers (2007), 'Meeting strangers and friends of friends: How random are social networks?', *American Economic Review* **97**(3), 890–915.

Jochmans, Koen (2017), 'Two-way models for gravity', *Review of Economics and Statistics* .

Koskinen, Johan H. (2008), The linked importance sampler auxiliary variable metropolis hastings algorithm for distributions with intractable normalising constants. MelNet Social Networks Laboratory Technical Report 08-01, Department of Psychology, School of Behavioural Science, University of Melbourne, Australia.

Laschever, Ron (2009), The doughboys network: Social interactions and labor market outcomes of world war i veterans. working paper.

Lehman, E. L. (1983), *Theory of Point Estimation*, Wiley and Sons.

Leung, Michael (2014), A random-field approach to inference in large models of network formation. working paper.

Liang, Faming (2010), 'A double metropolis-hastings sampler for spatial models with intractable normalizing constants', *Journal of Statistical Computing and Simulation* **80**, 1007–1022.

Mayer, Adalbert and Steven L. Puller (2008), 'The old boy (and girl) network: Social network formation on university campuses.', *Journal of Public Economics* **92**(1-2), 329–347.

McCormick, T. H. and T. Zheng (2015), 'Latent surface models for networks using aggregated relational data', *Journal of the American Statistical Association* (110), 1684–1695.

McLachlan, G. J. and D. Peel (2000), *Finite mixture models*, Wiley Series in Probability and Statistics.

Mele, Angelo (2017), 'A structural model of dense network formation', *Econometrica* **85**, 825–850.

Mele, Angelo and Lingjiong Zhu (2020), Approximate variational estimation for a model of network formation.

Mele, Angelo, Lingxin Hao, Joshua Cape and Carey E. Priebe (2019), Spectral inference for large stochastic blockmodels with nodal covariates. working paper.

Menzel, Konrad (2015), Strategic network formation with many agents, Working papers, NYU.

Monderer, Dov and Lloyd Shapley (1996), 'Potential games', *Games and Economic Behavior* **14**, 124–143.

Moody, James (2001), 'Race, school integration, and friendship segregation in america', *American Journal of Sociology* **103**(7), 679–716.

Murray, Iain A., Zoubin Ghahramani and David J. C. MacKay (2006), 'Mcmc for doubly-intractable distributions', *Uncertainty in Artificial Intelligence* .

Nakajima, Ryo (2007), 'Measuring peer effects on youth smoking behavior', *Review of Economic Studies* **74**(3), 897–935.

Ridder, Geert and Shuyang Sheng (2015), Estimation of large network formation games, Working papers, UCLA.

Schweinberger, Michael (2020), 'Consistent structure estimation of exponential family random graph models with block structure', *Bernoulli* **26**, 1205–1233.

Schweinberger, Michael and Jonathan Stewart (2020), 'Concentration and consistency results for canonical and curved exponential-family models of random graphs', *Annals of Statistics* **48**(1), 374–396.

Schweinberger, Michael and Mark S Handcock (2015), 'Local dependence in random graph models: char- acterization, properties and statistical inference.', *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* (77), 1–30.

Schweinberger, Michael and Pamela Luna (forthcoming), 'Hergm: Hierarchical exponential-family random graph models', *Journal of Statistical Software* .

Sheng, Shuyang (2012), Identification and estimation of network formation games.

Snijders, Tom A.B (2002), 'Markov chain monte carlo estimation of exponential random graph models', *Journal of Social Structure* **3**(2).

Stephens, Matthew (2000), 'Dealing with label switching in mixture models', *Journal of the Royal Statistical Society B* .

Topa, Giorgio (2001), 'Social interactions, local spillovers and unemployment', *Review of Economic Studies* **68**(2), 261–295.

Appendix A. Additional theoretical results

A.1. **Proof of Proposition 1.** Consider the network $g = (g_{ij} = 1, g_{-ij})$, where g_{ij} is the entry at row i and column j of g; and g_{-ij} is the network g excluding entry g_{ij}. Let $g' = (g_{ij} = 0, g_{-ij})$ be a network in which link ij is deleted. It is straighforward to show that for any pair i and j the surplus from link g_{ij} is

$$\begin{aligned} U_i(g,x,z;\theta) + U_j(g,x,z;\theta) \quad &- \quad [U_i(g',x,z;\theta) + U_j(g',x,z;\theta)] \\ &= \quad u_{ij}(\theta_u) + u_{ji}(\theta_u) + 2\sum_{r\neq i,j}^{n} g_{jr}g_{ri}v_{ijr}(\theta_v) + 2\sum_{r\neq i,j}^{n} g_{ir}g_{rj}v_{jir}(\theta_v) \end{aligned}$$

and using the fact that that $g_{ij} = g_{ji}$ in an undirected network and by Assumption 2 $v_{ijr}(\theta_v) = v_{jir}(\theta_v)$, we obtain

$$\begin{aligned} U_i(g,x,z;\theta) + U_j(g,x,z;\theta) \quad &- \quad [U_i(g',x,z;\theta) + U_j(g',x,z;\theta)] \\ &= \quad u_{ij}(\theta_u) + u_{ji}(\theta_u) + 4\sum_{r\neq i,j}^{n} g_{jr}g_{ri}v_{ijr}(\theta_v) \end{aligned} \tag{30}$$

We need to show that the difference in potential $Q(g,x,z;\theta) - Q(g',x,z;\theta)$ is the same as the expression above. In network g, we have $g_{ij}=1$, so we write the potential as follows

$$\begin{aligned}
Q(g,x,z;\theta) &= g_{ij}u_{ij}(\theta_u) + g_{ji}u_{ji}(\theta_u) + \sum_{l\neq i}^{n}\sum_{s\neq j}^{n} g_{ls}u_{ls}(\theta_u) \\
&+ \frac{2}{3}\Big[\sum_{r\neq i,j}^{n} g_{ij}g_{jr}g_{ri}v_{ijr}(\theta_v) + \sum_{r\neq i,j}^{n} g_{ji}g_{ir}g_{rj}v_{jir}(\theta_v) + \sum_{r\neq i,j}^{n} g_{jr}g_{ri}g_{ij}v_{jri}(\theta_v) \\
&+ \sum_{r\neq i,j}^{n} g_{ir}g_{rj}g_{ji}v_{irj}(\theta_v) + \sum_{r\neq i,j}^{n} g_{ri}g_{ij}g_{jr}v_{rij}(\theta_v) + \sum_{r\neq i,j}^{n} g_{rj}g_{ji}g_{ir}v_{rji}(\theta_v)\Big] \\
&+ \frac{2}{3}\sum_{l\neq i}^{n}\sum_{s\neq j,i}\sum_{q\neq i,j,r} g_{ls}g_{sq}g_{gl}v_{lsq}(\theta_v) \\
&= u_{ij}(\theta_u) + u_{ji}(\theta_u) + \sum_{l\neq i}^{n}\sum_{s\neq j}^{n} g_{ls}u_{ls}(\theta_u) \\
&+ \frac{2}{3}\Big[\sum_{r\neq i,j}^{n} g_{jr}g_{ri}v_{ijr}(\theta_v) + \sum_{r\neq i,j}^{n} g_{ir}g_{rj}v_{jir}(\theta_v) + \sum_{r\neq i,j}^{n} g_{jr}g_{ri}v_{jri}(\theta_v) \\
&+ \sum_{r\neq i,j}^{n} g_{ir}g_{rj}v_{irj}(\theta_v) + \sum_{r\neq i,j}^{n} g_{ri}g_{jr}v_{rij}(\theta_v) + \sum_{r\neq i,j}^{n} g_{rj}g_{ji}g_{ir}v_{rji}(\theta_v)\Big] \\
&+ \frac{2}{3}\sum_{l\neq i}^{n}\sum_{s\neq j,i}\sum_{q\neq i,j,r} g_{ls}g_{sq}g_{gl}v_{lsq}(\theta_v) \\
&= u_{ij}(\theta_u) + u_{ji}(\theta_u) + \sum_{l\neq i}^{n}\sum_{s\neq j}^{n} g_{ls}u_{ls}(\theta_u) \\
&+ 4\sum_{r\neq i,j}^{n} g_{jr}g_{ri}v_{ijr}(\theta_v) + \frac{2}{3}\sum_{l\neq i}^{n}\sum_{s\neq j,i}\sum_{q\neq i,j,r} g_{ls}g_{sq}g_{ql}v_{lsq}(\theta_v)
\end{aligned}$$

where the second equality follows from the fact that $g_{ij} = g_{ji} = 1$, and the third equality uses the assumption $v_{\phi(ijr)} = v_{ijr}$ and that $g_{ls} = g_{sl}$ for any s, l.

The value of the potential when with g', where the link $g_{ij} = 0$, is

$$Q(g',x,z;\theta) = \sum_{l\neq i}^{n}\sum_{s\neq j}^{n} g_{ls}u_{ls}(\theta_u) + \frac{2}{3}\sum_{l\neq i}^{n}\sum_{s\neq j,i}\sum_{q\neq i,j,r} g_{ls}g_{sq}g_{ql}v_{lsq}(\theta_v)$$

Therefore we obtain that the difference

$$Q(g,x,z;\theta) - Q(g',x,z;\theta) \quad = \quad u_{ij}(\theta_u) + u_{ji}(\theta_u) + 4\sum_{r\neq i,j}^{n} g_{jr}g_{ri}v_{ijr}(\theta_v)$$

where we are using the fact that $g_{ij} = g_{ji}$ in an undirected network and by Assumption 2 $v_{ijr}(\theta_v) = v_{jir}(\theta_v)$. We can repeat the same reasoning for any pair ij to obtain the proposition.

A.2. **Perfect segregation meeting technology.** Let's consider a special case in which the meeting probability is $\rho(g_{-ij}, x_i, x_j, z_i, z_j) = \rho_w(g_{-ij}, x_i, x_j)$ if $z_i = z_j$ and $\rho_b(g_{-ij}, x_i, x_j, z_i, z_j) = 0$ otherwise, for any pair (i,j). This is a case in which the meeting process generetes perfect segregation. It is easy to show that when this is the case I can factorize the likelihood of observing the network as the product of K subnetwork likelihoods, one for each community.

PROPOSITION 3. ***Perfect segregation meeting technology.***
If $\rho_b(g_{=ij}, x_i, x_j) = 0$, then the likelihood of observing network g, conditional on the community structure z and covariates x is

$$\pi(g,x,z;\theta) = \prod_{k=1}^{K} \frac{\exp\left[Q_{k,k}(g_{k,k}, x^{(k)}, z;\theta)\right]}{c_{k,k}(\mathcal{G}_{k,k}, x^{(k)};\theta)} \tag{31}$$

where the potential $Q_{k,k}(g_{k,k}, x^{(k)}, z;\theta)$ is

$$Q_{k,k}(g_{k,k}, x^{(k)}, z;\theta) = \sum_{i\in\mathcal{C}_k}\sum_{j\in\mathcal{C}_k} g_{ij}u(x_i,x_j,z_i,z_j;\alpha,\beta) + \frac{2\gamma_k}{3}\sum_{i\in\mathcal{C}_k}\sum_{j\in\mathcal{C}_k}\sum_{r\in\mathcal{C}_k} g_{ij}g_{jr}g_{ri} \tag{32}$$

and the normalizing constant $c_{k,k}(\mathcal{G}_{k,k}, x^{(k)};\theta)$ is

$$c_{k,k}(\mathcal{G}_{k,k}, x^{(k)};\theta) = \sum_{\omega_{k,k}\in\mathcal{G}_{k,k}} \exp\left[Q_{k,k}(\omega_{k,k}, x^{(k)}, z;\theta)\right] \tag{33}$$

Proof. When $\rho_b(g_{-ij}, x_i, x_j) = 0$, there is no meeting of players across communities, therefore no links across communities will be formed. Each community is independent; within the

community, the network formation process is the same as Mele (2017) and Mele and Zhu (2020) (see Theorem 1 in Mele (2017) for a proof). Therefore each community subnetwork converges to an exponential random graph, with potential function $Q_{k,k}(g_{k,k}, x^{(k)}, z; \theta)$. □

The likelihood in (31) factorizes as a product of K independent exponential random graphs. The factorization is quite useful in estimation and for the identification of the parameters. In terms of estimation, one can parallelize the estimation routines, simulating a subnetwork on each processor. In addition, given the i.i.d. nature of the sample, identification can be obtained with standard regularity conditions for the exponential family. In particular, as the number of communities grows large, the parameters are identified. See Lehman (1983) for a general discussion. See also Badev (2013).

A.3. **Likelihood factorization with local transitivity.** When the model specification includes local transitivity, $v_{ijr} = \gamma_k$ if $z_{ik} = z_{jk} = z_{rk} = 1$ and $v_{ijr} = 0$ otherwise, the likelihood in (8) factorizes in a convenient way as shown in the next lemma.

LEMMA 1. *Assume that the model specification is:* $v_{ijr} = \gamma_k$ *if* $z_{ik} = z_{jk} = z_{rk} = 1$ *and* $v_{ijr} = 0$ *otherwise. Then the likelihood factorizes in within- and between-communities components.*

$$P_\theta\left(G = g | Z = z, X = x\right) = \prod_{k=1}^{K} P\left(G_{k,k} = g_{k,k} | Z = z, X = x; \theta\right) \times \left[\prod_{l>k}^{K} P\left(G_{k,l} = g_{k,l} | Z = z, X = x; \theta\right)\right] \tag{34}$$

Proof. Let $g_{k,l}$ denote the subnetwork formed by individuals of communities $\mathcal{C}_k$ and $\mathcal{C}_l$. Let $x^{(k)}$ denote the covariates of players in community $\mathcal{C}_k$. The potential can be decomposed into the sum of sub-potentials for the sub-networks $g_{k,l}$'s. That is, we can decompose the potential $Q(g, x, z, \theta)$ as sum of subpotentials $Q_{k,l}(g_{k,l}, x^{(k)}, x^{(l)}, z)$, separating the within-community

and between-community contributions as follows:

$$
\begin{aligned}
(35)\quad Q(g,x,z;\theta) &= \sum_{k=1}^{K} Q_{k,k}(g_{k,l},x^{(k);\theta},z) + \sum_{k=1}^{K}\sum_{l>k}^{K} Q_{k,l}(g_{k,l},x^{(k)},x^{(l)},z;\theta) \\
(36) \qquad &= \sum_{k=1}^{K}\left[\sum_{i\in\mathcal{C}_k}\sum_{j\in\mathcal{C}_k} g_{ij}u(x_i,x_j,z_i,z_j;\alpha,\beta) + \frac{2\gamma_k}{3}\sum_{i\in\mathcal{C}_k}\sum_{j\in\mathcal{C}_k}\sum_{r\in\mathcal{C}_k} g_{ij}g_{jr}g_{ri}\right] \\
(37) \qquad &+ \sum_{k=1}^{K}\sum_{l>k}^{K}\left[\sum_{i\in\mathcal{C}_k}\sum_{j\in\mathcal{C}_l} g_{ij}u(x_i,x_j,z_i,z_j;\alpha,\beta)\right]
\end{aligned}
$$

This decomposition of the potential function allows us to rewrite the likelihood as follows

$$
\begin{aligned}
(38)\quad \pi(g,x,z;\theta) &= \prod_{k=1}^{K}\frac{\exp\left[Q_{k,k}(g_{k,k},x^{(k)},z;\theta)\right]}{c_{k,k}(\mathcal{G}_{k,k},x^{(k)};\theta)}\left[\prod_{l>k}^{K}\frac{\exp\left[Q_{k,l}(g_{k,l},x^{(k)},x^{(l)},z;\theta)\right]}{c_{k,l}(\mathcal{G}_{k,l},x^{(k)},x^{(l)};\theta)}\right] \\
(39) \qquad &= \prod_{k=1}^{K} P\left(g_{k,k}|z,x;\theta\right)\left[\prod_{l>k}^{K} P\left(g_{k,l}|z,x;\theta\right)\right]
\end{aligned}
$$

where the normalizing constants $c_{k,k}(\mathcal{G}_{k,k},x^{(k)};\theta)$ and $c_{k,l}(\mathcal{G}_{k,l},x^{(k),x^{(l)}};\theta)$ are

$$
\begin{aligned}
(40)\qquad c_{k,k}(\mathcal{G}_{k,k},x^{(k)};\theta) &= \sum_{\omega_{k,k}\in\mathcal{G}_{k,k}} \exp\left[Q_{k,k}(\omega_{k,k},x^{(k)},z;\theta)\right] \\
(41)\qquad c_{k,l}(\mathcal{G}_{k,l},x^{(k)},x^{(l)};\theta) &= \sum_{\omega_{k,l}\in\mathcal{G}_{k,l}} \exp\left[Q_{k,l}(g_{k,l},x^{(k)},x^{(l)},z;\theta)\right]
\end{aligned}
$$

Notice that the potential decomposition above, is consistent with the result in Lemma **??**; that is, the sufficient statistics of the model can be written as sum of within- and between-communities sufficient statistics. □

Appendix B. Estimation details

All the computations have been performed on a desktop Dell Precision T7620 with 2 Intel Xeon CPUs E5-2697 v2 with 12 Dual core processors at 2.7GHZ each and 64GB of RAM. I estimated all the models using the package `hergm` in R, developed by Schweinberger and

Handcock (2015) and Schweinberger and Luna (forthcoming). The code for estimation and replication is available from the author.

Each estimate is obtained with a 100,000 simulation run of the exchange algorithm. I collect 10,000 samples and discard the first 2,000 as burnin. I also experimented with a longer run of 200,000 steps, without changes in the results.

Add Health restricted-use data used in this paper can be obtained by applying at the website: http://www.cpc.unc.edu/projects/addhealth

B.1. **Block recovery in simulated data.** In this section, I estimate the model using simulated data for different specifications, and check the ability of the model and estimation algorithm to recover the community structure and the payoffs' parameters.
I generate 1000 networks with $n = 60$ nodes using the different spefications described in the text. In each simulation, there are $K = 3$ blocks and each block contains $n_k = 20$ nodes. For each simulated network I estimate the block structure and the parameters.

Note that in each simulation, the block memberships are estimated using the correct specification and assuming $K = 3$.

The parameters used to generate the simulated networks:

(1) **restricted, local transitivity**: $\alpha = (\alpha_1, \alpha_2, \alpha_3) = (-3, -2, -1)$; $\gamma = (\gamma_1, \gamma_2, \gamma_3) = (0.5, 0.5, 0.5)$; $\alpha_b = -6$;
(2) **restricted, global transitivity**: $\alpha = (\alpha_1, \alpha_2, \alpha_3) = (-3, -2, -1)$; $\gamma = 0.2$; $\alpha_b = -6$;
(3) **unrestricted, local transitivity**: $\alpha = (\alpha_1, \alpha_2, \alpha_3) = (-3, -2, -1)$; $\gamma = (\gamma_1, \gamma_2, \gamma_3) = (0.5, 0.5, 0.5)$;
(4) **unrestricted, global transitivity**: $\alpha = (\alpha_1, \alpha_2, \alpha_3) = (-3, -2, -1)$; $\gamma = 0.2$;
(5) **restricted, no transitivity (sbm)**: $\alpha = (\alpha_1, \alpha_2, \alpha_3) = (-3, -2, -1)$; $\alpha_b = -5$;
(6) **unrestricted, no transitivity (sbm)**: $\alpha = (\alpha_1, \alpha_2, \alpha_3) = (-3, -2, -1)$.

The performance in recovering the block structure of the simulated data is assessed using the Yule's ϕ-coefficient as in Babkin et al. (2020)

$$\phi(\widehat{z}, z) = \frac{n_{0,0}n_{1,1} - n_{0,1}n_{1,0}}{\sqrt{(n_{0,0}+n_{0,1})(n_{1,0}+n_{1,1})(n_{0,0}+n_{1,0})(n_{0,1}+n_{1,1})}} \tag{42}$$

where $n_{0,0}$ is

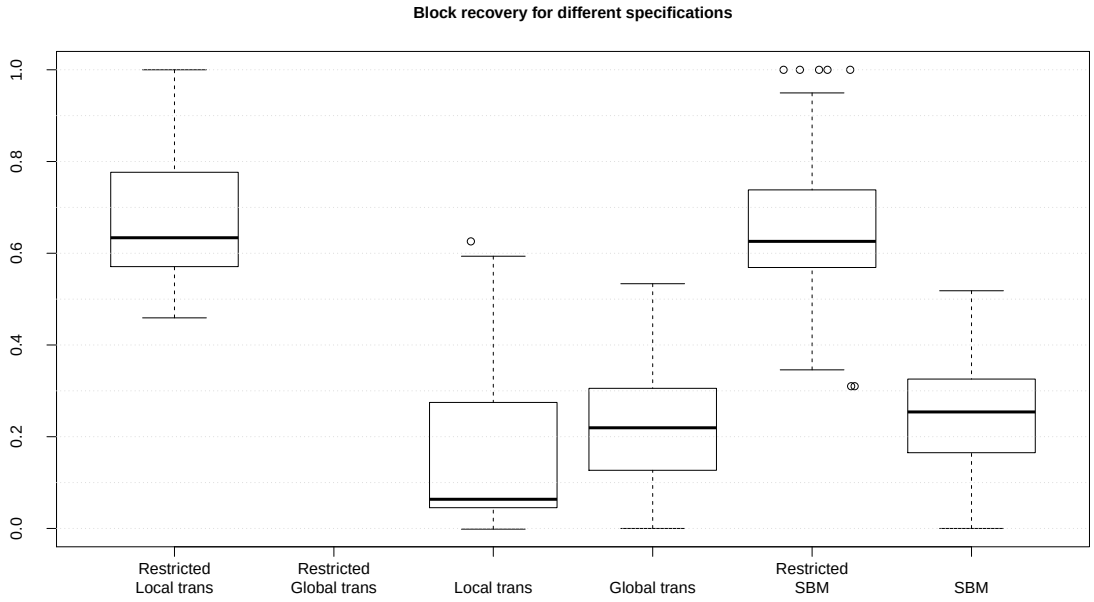
$$n_{0,0} := \sum_{i<j}^{n} \mathbb{I}\left(\mathbb{I}\left(\widehat{z}_i = \widehat{z}_j\right) = 0\right) \mathbb{I}\left(\mathbb{I}\left(z_i = z_j\right) = 0\right) \tag{43}$$

and $\mathbb{I}$ is an indicator variable. The terms $n_{0,1}$, $n_{1,0}$ and $n_{1,1}$ are defined analogously. The quantity $n_{0,0}$ is the number of nodes pairs assigned to distinct blocks under both the true z and the estimated $\widehat{z}$; on the other hand, $n_{1,1}$ is the number of nodes pairs assigned to same block under both the true z and the estimated $\widehat{z}$. The number of pairs where true z and the estimated $\widehat{z}$ block structure disagree are $n_{0,1} + n_{1,0}$. The coefficient (42) is equal to 1 when all the block assignments coincide under the true z and the estimated $\widehat{z}$, and it is invariant to the labeling of the blocks.

The boxplots of the simulation results are in Figure 4. First, note that the specification with restricted block dependence and global transitivity estimates only one block, thus the Yule's ϕ-coefficient is not defined. This specification gives the worst performance in these simulations. Second, the retstricted block dependence and local transitivity specification performs best. Fianlly, the model with local transitivity performs as well as the stochastic blockmodel.

B.2. **Choice of K with Add Health data.** There are no established methods to estimate the number of communities in this class of models. To choose the number of blocks for our model specification, I use an heuristic approach and compute Root MSE (RMSE) and Median Absolute Deviation (MAD) of the posterior predictions for edges, 2-stars and triangles. These results are shown in Tables 3, 4 and 5. The columns represents the different specifications, while the rows indicate the number of communities.

FIGURE 4. Monte Carlo experiments: block recovery performance of different specifications



Boxplots of Yule's coefficient (42) for several specifications of the model. Values closer to 1 indicate better recovery of the block structure. Each plot is generate by estimating 1000 simulated networks with $n = 60$ nodes, and $K = 3$ blocks of equal size. The model with restricted block dependence and global transitivity estimates only one block.

TABLE 3. Root Mean Squared Error and Median Absolute Deviation of posterior predictions, number of edges

	Restricted Local trans		Restricted Global trans		Unrestricted Local trans		Unrestricted Global trans		Restricted SBM		Unestricted SBM	
	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD
K=2	235.54	86.73	77.55	54.86	97.55	47.44	74.75	17.79	23.92	23.72	23.06	23.72
K=3	138.97	70.42	91.04	57.82	117.22	16.31	111.22	17.79	22.65	22.24	22.95	22.24
K=4	172.09	67.46	78.32	62.27	119.17	18.53	80.49	56.34	23.97	23.72	179.32	23.72
K=5	136.84	72.65	76.78	53.37	183.74	75.61	71.57	12.60	23.08	22.24	53.54	23.72
K=6	138.95	65.98	1605.83	69.68	77.70	63.75	73.09	13.34	22.81	22.24	23.91	22.24
K=10	137.54	75.61	80.78	51.89	234.25	85.25	69.86	62.27	25.77	22.24	168.81	23.72

The method suggested in Schweinberger and Handcock (2015) is to estimate the model with an increasing K, and choose the specification with the lowest RMSE for posterior prediction of the number of triangles. Their specification only includes the model with restricted block-dependent α and local transitivity. According to this method, one would choose a model with $K = 3$, based on the RMSE in the first column of Table 5.

TABLE 4. Root Mean Squared Error and Median Absolute Deviation of posterior predictions, number of 2-stars

	Restricted Local trans		Restricted Global trans		Unrestricted Local trans		Unrestricted Global trans		Restricted SBM		Unestricted SBM	
	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD
K=2	8433.64	1209.80	1473.34	675.32	6867.51	955.54	6681.56	652.34	247.08	243.15	267.66	246.11
K=3	3811.00	991.12	1938.73	968.14	10044.25	897.71	10342.54	778.37	218.02	215.72	272.03	236.47
K=4	5663.17	924.40	1442.05	832.48	9955.56	959.24	5832.33	1656.81	220.38	205.34	7683.64	253.52
K=5	3787.86	989.64	1512.53	762.06	7216.44	1177.18	6345.71	603.42	227.01	223.87	1657.37	240.92
K=6	3989.07	911.06	120081.28	1089.71	1364.88	781.33	7088.31	678.29	212.29	212.01	282.13	223.87
K=10	4002.89	976.29	1500.49	759.09	10207.10	1557.47	1642.81	1049.68	248.68	201.63	6947.38	255.01

TABLE 5. Root Mean Squared Error and Median Absolute Deviation of posterior predictions, number of triangles

	Restricted Local trans		Restricted Global trans		Unrestricted Local trans		Unrestricted Global trans		Restricted SBM		Unestricted SBM	
	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD	RMSE	MAD
K=2	999.66	106.75	145.07	69.68	199.60	97.85	20.39	20.02	79.84	16.31	95.67	10.38
K=3	399.38	97.11	137.30	78.58	252.04	44.48	37.43	34.10	70.39	19.27	94.87	11.86
K=4	618.44	94.89	141.91	85.99	217.98	63.75	35.08	28.91	64.93	20.76	1013.65	11.86
K=5	373.55	90.44	112.82	66.72	797.80	100.82	27.16	31.13	72.21	16.31	192.77	11.86
K=6	419.11	93.40	19794.28	93.40	141.62	87.47	18.31	17.79	64.25	20.76	95.47	11.86
K=10	435.15	94.15	92.88	62.27	1167.92	111.94	120.01	79.32	60.01	19.27	834.52	13.34

B.2.1. *Choosing K and specification using RMSE of posterior predictions.* For each specification, I choose the optimal number of K using the method of Schweinberger and Handcock (2015). So K is the minimum number of blocks that minimizes the RMSE of posterior predictions for the number of triangles. This results in a choice of $K = 3$ for restricted block-dependent and local transitivity, restricted block-dependent and global transitivity, and unrestricted block-dependent stochastic blockmodel; $K = 2$ for unrestricted block-dependent and local transitivity, and unrestricted block-dependent and global transitivity; $K = 4$ for unrestricted block-dependent stochastic blockmodel. Figure 5 shows the histograms of posterior predictions for edges, 2-stars and triangles for each of these models.

Note that the stochastic blockmodels are consistently under-estimating the number of triangles, and therefore the clustering level of the network (Panel (C)). If we consider the number of two-stars, it seems that the models with unrestricted block-dependence and transitivity (global and local) have a hard time fitting (Panel (B)). The model with restricted

block-dependence and transitivity (global and local) seems to be doing a better job in fitting these three network features.

The model with local transitivity puts most of the probability mass close to the observed network statistics. Therefore, I prefer this specification in the empirical application.

B.2.2. *Choosing K and specification using Median Absolute Deviation of posterior predictions.* Since in the simulations there are a few outliers, I also consider a more robust measure of distance from the observed sufficient statistics: the Mean Absolute Deviation (MAD) of posterior predictions for triangles. Indeed, in Table 5 I report the Median Absolute Deviation of posterior predictions from observed triangles. Using those columns to choose K, I show the histograms of the fit in Figure 6. Stochastic blockmodels underestimate the number of triangles (Panel (C)), while the models with restricted block-dependence seem to give a better fit. Models with unrestricted block-dependence consistently underestimate the number of stars (Panel (B)).

B.3. **Prior truncation.** In the empirical application, the priors are truncated, so that the number of communities is at most K. Therefore we have

$$\eta_1 = V_1 \tag{44}$$

$$\eta_k = V_k \prod_{j=1}^{K-1} (1 - V_j) \qquad k = 2, 3, 4, ..., K \tag{45}$$

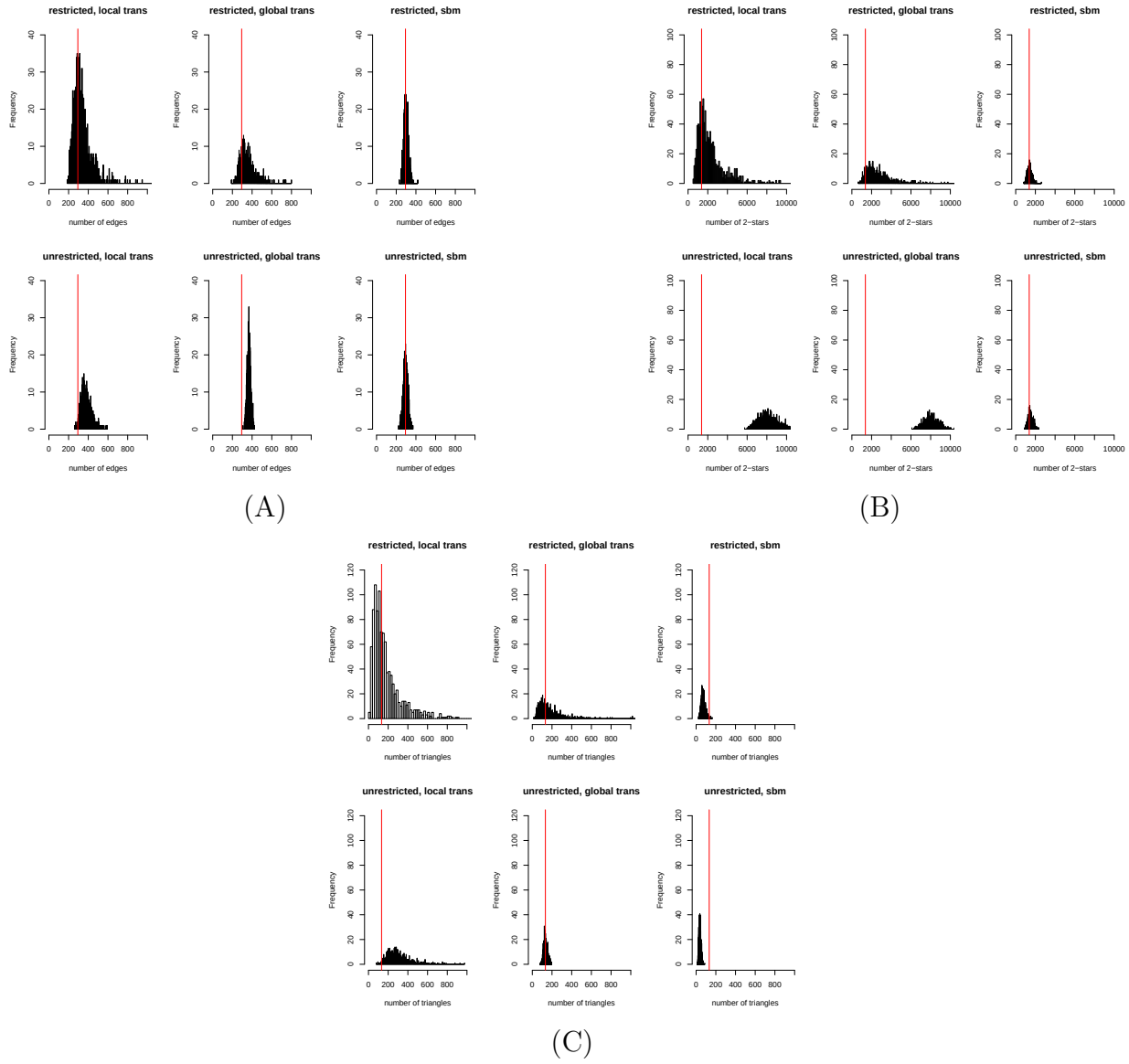
$$V_k|\phi \overset{iid}{\sim} Beta(1, \phi) \qquad k = 1, 2, 3, ..., K-1 \tag{46}$$

$$V_K = 1 \tag{47}$$

$$\phi > 0 \quad \text{and} \quad \sum_{k=1}^{K} \eta_k = 1 \ w.p.1 \tag{48}$$

This simpler formulation with truncation provides a more parsimonious model and improves the speed of computations and simulations from the posterior. Truncation is indeed needed

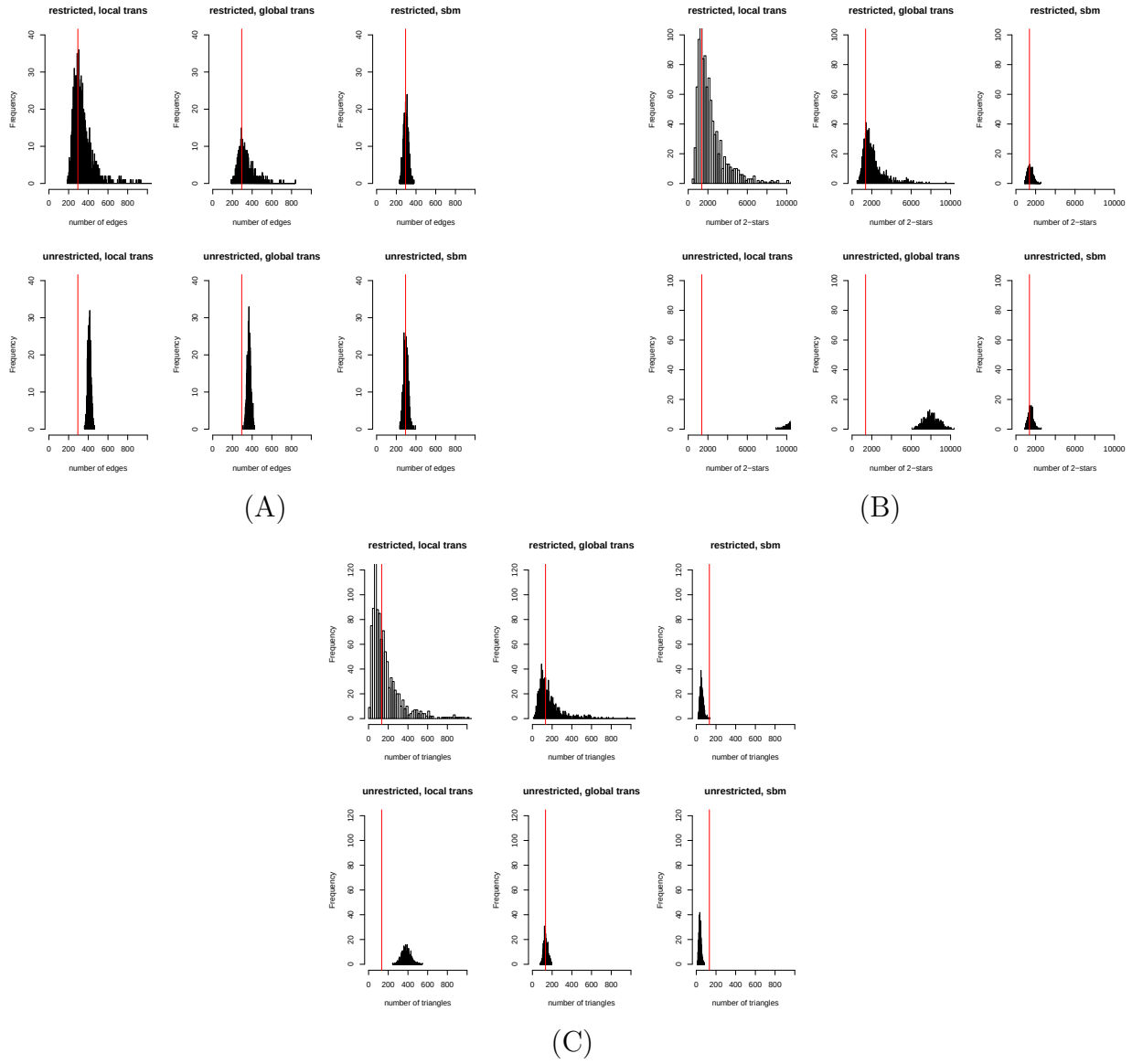
FIGURE 5. Fit of different specifications for edges, 2-stars and triangles, according to the choice of K using RMSE of posterior predictions of triangles, for each specification.



The graphs report the simulated posterior predictions for (A) edges, (B) 2-stars and (C) triangles. The number of blocks K for each specification is choses according to the heuristic method of Schweinberger and Handcock (2015), minimizing the RMSE for posterior predictions of triangles.

because the number of parameters to estimate depends on the number of communities and we have only one network observation.

FIGURE 6. Fit of different specifications for edges, 2-stars and triangles, according to the choice of K using Median Absolute Deviation (MAD) of posterior predictions of triangles, for each specification.



The graphs report the simulated posterior predictions for (A) edges, (B) 2-stars and (C) triangles. The number of blocks K for each specification is choses according to the heuristic method of Schweinberger and Handcock (2015), minimizing the MAD for posterior predictions of triangles.

B.4. **Hyper-priors used in estimation.** I use the default in the `hergm` package in R. The hyper prior on ϕ is

$$\phi \sim Gamma(1,1) \tag{49}$$

(50)

The hyper-priors for $\mu_w, \mu_b, \sigma_w, \sigma_b$ are

$$\mu_w \sim N(0,1) \tag{51}$$

$$\mu_b \sim N(0,1) \tag{52}$$

$$\sigma_w \sim Gamma(10,10) \tag{53}$$

$$\sigma_b \sim Gamma(10,10) \tag{54}$$

(55)

All the variables are independent.

B.5. **Data.** The network used in the estimation exercise is from the *National Longitudinal Study of Adolescent Health* (Add Health). This dataset contains information on a nationally representative sample of US schools. The survey started in 1994, when the 90118 participants were entering grades 7-12, and the project collected data in four successive waves.More details about the sampling design and the representativeness are contained in Moody (2001) and the Add Health website http://www.cpc.unc.edu/projects/addhealth/projects/addhealth. Each student responded to an *in-school* questionnaire, and a subsample of 20745 was given an *in-home* interview to collect more detailed information about behaviors, characteristics and health status. The survey asked each student a long set of demographic, health and socioeconomic questions. In addition, students were provided with the roster of their school and asked to identify up to 5 male and 5 female friends. One can think that this limit could bias the friendship data, but only 3% of the students nominated 10

friends (Moody, 2001). Moreover, the estimation routine could be easily extended to deal with missing links. I use this part of the survey to construct the (undirected) network of friendships.

The model estimated includes the following covariates: race, gender, grade and parental income. These are some of the variables that are considered good predictors of friendships during adolescence (Moody, 2001; Mayer and Puller, 2008; Boucher, 2015).

In this paper I use only data from school 28 and Wave I (1994), from the saturated sample. Each student in this sample completed both the in-school and in-home questionnaires, and the researchers made a significant effort to avoid any missing information on the students.

I use data on racial group, grade and gender of individuals. A student with a missing value in any of these variables is dropped from the sample. Each student that declares to be of Hispanic origin is considered Hispanic. The remaining non-Hispanic students are assigned to the racial group they declared. Therefore the racial categories are: White, Black, Asian, Hispanic and Other race. Other race contains Native Americans. I also control for homophily in income, using the family income reported in a question from the parent questionnaire. There are several cases in which the family income is missing. For those observations, I imputed values drawn from the unconditional income distribution of the community. An alternative but computationally very costly alternative is to introduce an additional step in the simulation, in which the imputation of missing incomes is done at each iteration.

The school contains 150 students, with 58.7% females. The school is very racially heterogeneous: 42% Whites/Caucasians, 45.3% African-Americans, 0.7% Asians, 10.7% Hispanics and 1.3% Other race. The racial fragmentation index is 0.606. The school offers all grades from 7 to 12, with a relatively balanced population among the different age groups, respectively 17.3%, 17.3%, 20%, 16.7%, 14%, and 14.7%.

This school exhibits a high level of segregation, measure using the index developed by Freeman (1972), that varies between a minimum of 0 (no segregation) and 1 (perfect segregation). The measured segregation level is 0.72 for Whites/Caucasians, 0.764 for African Americans, and 0.429 for Hispanics. The segregation by gender is 0.255.

B.6. **Label invariance and relabeling algorithm.** An additional challenge is that the likelihood of this model is invariant to permutations of the community labels. This problem is common in the literature on finite mixture models, where the likelihood is invariant to permutations of the labels of the mixture's components (Gelman et al. (2003), McLachlan and Peel (2000), Stephens (2000)). This complicates inference for the community-specific parameters, because the community labels may switch several times during the MCMC simulation.

Nonetheless, the use of nonparametric priors implies that the full posterior is not invariant to permutations of the community labels. This reduces the problem. Furthermore, after obtaining a MCMC sample from the posterior distribution, I use the algorithm of Schweinberger and Handcock (2015) to relabel the output of the posterior simulations. This approach is common in the Bayesian literature on finite mixture models.

Suppose to have a MCMC posterior simulation $\{\theta^s, z^s\}_{s=1}^n$ of length S. The relabeling algorithm minimizes the loss function

$$L(\xi, \nu(Z)) = \min_{\nu} L_0 \left[\xi, \nu(Z)\right] \tag{56}$$

where

$$L_0 \left[\xi, \nu(Z)\right] = -\log \prod_{i=1}^{n} \xi_{i,\mathcal{C}_i} \tag{57}$$

where ξ is an $n \times K$ matrix whose entry $\xi_{i,k}$ is the probability that individual i is reported to be in community/type k; and $\nu(Z)$ is a permutation of the community structure Z.

So the goal of the relabeling algorithm is to choose the matrix ξ that minimizes the posterior expectation of loss function $L[\xi, \nu(Z)]$. In practice the posterior expectation is approximated by the Monte Carlo sample

$$\frac{1}{S}\sum_{s=1}^{S}\min_{\nu_s}\left[L_0\left[\xi, \nu_s(z^s)\right]\right] = \min_{\nu_1,\ldots,\nu_S}\left[\frac{1}{S}\sum_{s=1}^{S}L_0\left[\xi, \nu_s(z^s)\right]\right] \tag{58}$$

and the algorithm starts from some initial permutation of the community labels $\nu_1, \nu_2, ..., \nu_S$ and iterates on the following two steps until convergence:

(1) choose $\hat{\xi}$ to minimize $\sum_{s=1}^{S}\left[L_0\left[\xi, \nu_s(z^s)\right]\right]$ subject to the constraint $\sum_{k=1}^{K}\xi_{i,k} = 1$ for $i = 1, ..., n$;

(2) for $s = 1, ..., S$ choose ν_s to minimize $L_0\left[\xi, \nu_s(z^s)\right]$

The second step is infeasible unless the number of communities K is very small. Schweinberger and Handcock (2015)'s implementation uses Simulated Annealing to perform the S minimizations in parallel. Similar algorithms for relabeling the output of the MCMC are discussed in Gelman et al. (2003) and McLachlan and Peel (2000). More details on the practical implementation are in Schweinberger and Handcock (2015) and Stephens (2000). This procedure is implemented in the R package `hergm`.

ONLINE APPENDIX:

NOT FOR PUBLICATION

Appendix C. Tables of estimates

Table 6. block restricted, local transitivity, $K = 1$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.025	0.688	3.629
Mean of parameters of hergm term 1: $\mu_{w,\alpha}$	-3.335	-1.745	-0.024
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-0.899	0.558	1.952
Precision of parameters of hergm term 1: $\sigma_{w,\alpha}$ 0.378	0.811	1.492	
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$	0.482	0.974	1.716
hergm term 1: parameter of block 1:	-4.737	-3.896	-2.887
hergm term 2: parameter of block 1:	0.628	1.161	1.465
ergm term 1 parameter:	0.022	0.657	1.171
ergm term 2 parameter:	0.073	0.643	1.145
ergm term 3 parameter:	-0.113	1.602	2.697
ergm term 4 parameter:	0.431	1.225	1.764
ergm term 5 parameter:	0.001	1.147	1.883
ergm term 6 parameter:	0.126	0.959	1.671
ergm term 7 parameter:	-0.272	0.883	1.755
ergm term 8 parameter:	-0.033	1.133	2.003
ergm term 9 parameter:	-0.639	0.808	1.467
ergm term 10 parameter:	-1.236	-0.365	0.294
ergm term 11 parameter:	-1.144	-0.383	0.118
ergm term 12 parameter:	-1.656	-0.869	-0.318

TABLE 7. block restricted, local transitivity, $K = 2$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.095	0.693	2.410
Mean of parameters of hergm term 1: $\mu_{w,\alpha}$	-3.830	-2.531	-1.073
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-0.330	0.874	2.059
Precision of parameters of hergm term 1: $\sigma_{w,\alpha}$	0.391	0.828	1.581
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$	0.507	0.992	1.727
hergm term 1 parameter of block 1:	-4.856	-4.155	-3.028
hergm term 1: parameter of block 2:	-4.915	-4.111	-3.144
hergm term 1: between-block parameter:	-6.766	-5.862	-4.895
hergm term 2: parameter of block 1:	0.666	0.991	1.272
hergm term 2: parameter of block 2:	0.736	1.618	2.758
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	0.348	0.933	1.317
ergm term 2 parameter:	0.394	0.942	1.345
ergm term 3 parameter:	0.805	2.156	2.937
ergm term 4 parameter:	0.689	1.402	1.938
ergm term 5 parameter:	0.422	1.287	1.911
ergm term 6 parameter:	0.434	1.297	1.839
ergm term 7 parameter:	-0.224	1.056	1.752
ergm term 8 parameter:	0.157	1.233	1.999
ergm term 9 parameter:	-0.088	1.030	1.629
ergm term 10 parameter:	-0.817	-0.016	0.539
ergm term 11 parameter:	-0.771	-0.129	0.308
ergm term 12 parameter:	-1.248	-0.549	-0.093

TABLE 8. block restricted, local transitivity, $K = 3$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.117	0.686	2.130
Mean of parameters of hergm term 1: $\mu_{w,\alpha}$	-3.729	-2.501	-1.114
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-0.252	0.874	2.035
Precision of parameters of hergm term 1: $\sigma_{w,\alpha}$	0.406	0.849	1.585
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$	0.512	1.021	1.767
hergm term 1: parameter of block 1:	-4.888	-4.086	-3.091
hergm term 1: parameter of block 2:	-4.883	-3.895	-2.589
hergm term 1: parameter of block 3:	-4.385	-2.609	-0.316
hergm term 1: between-block parameter:	-6.636	-5.763	-4.837
hergm term 2: parameter of block 1:	0.644	0.977	1.244
hergm term 2: parameter of block 2:	0.508	1.561	2.738
hergm term 2: parameter of block 3:	-0.889	0.969	2.920
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	0.500	1.017	1.420
ergm term 2 parameter:	0.424	0.938	1.364
ergm term 3 parameter:	0.789	1.920	3.128
ergm term 4 parameter:	0.685	1.409	1.831
ergm term 5 parameter:	0.627	1.327	1.892
ergm term 6 parameter:	0.568	1.172	1.883
ergm term 7 parameter:	0.207	1.127	1.929
ergm term 8 parameter:	0.249	1.291	1.973
ergm term 9 parameter:	0.435	1.033	1.562
ergm term 10 parameter:	-0.689	-0.029	0.450
ergm term 11 parameter:	-0.725	-0.135	0.294
ergm term 12 parameter:	-1.208	-0.568	-0.136

TABLE 9. block restricted, local transitivity, $K = 4$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.405	1.389	3.393
Mean of parameters of hergm term 1: $\mu_{w,\alpha}$	-3.787	-2.677	-1.388
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-0.303	0.812	1.893
Precision of parameters of hergm term 1: $\sigma_{w,\alpha}$	0.433	0.905	1.624
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$	0.532	1.030	1.774
hergm term 1: parameter of block 1:	-4.838	-4.155	-3.252
hergm term 1: parameter of block 2:	-4.390	-2.935	-1.060
hergm term 1: parameter of block 3:	-5.031	-3.019	-0.647
hergm term 1: parameter of block 4:	-4.618	-3.741	-2.614
hergm term 1: between-block parameter:	-6.568	-5.778	-5.043
hergm term 2: parameter of block 1:	0.657	0.961	1.228
hergm term 2: parameter of block 2:	-0.859	0.834	2.436
hergm term 2: parameter of block 3:	-1.426	0.896	3.136
hergm term 2: parameter of block 4:	0.497	1.380	2.339
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	0.491	1.050	1.486
ergm term 2 parameter:	0.491	0.924	1.330
ergm term 3 parameter:	0.972	2.053	3.068
ergm term 4 parameter:	0.877	1.533	1.901
ergm term 5 parameter:	0.635	1.364	1.884
ergm term 6 parameter:	0.599	1.242	1.768
ergm term 7 parameter:	0.387	1.136	1.962
ergm term 8 parameter:	0.475	1.297	1.848
ergm term 9 parameter:	0.416	1.014	1.622
ergm term 10 parameter:	-0.483	0.061	0.542
ergm term 11 parameter:	-0.533	-0.100	0.309
ergm term 12 parameter:	-1.013	-0.512	-0.105

TABLE 10. block restricted, local transitivity, $K = 5$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.145	0.793	2.360
Mean of parameters of hergm term 1: $\mu_{w,\alpha}$	-3.635	-2.381	-1.019
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-0.577	0.704	1.947
Precision of parameters of hergm term 1: $\sigma_{w,\alpha}$	0.434	0.890	1.611
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$	0.516	1.009	1.741
hergm term 1: parameter of block 1:	-4.898	-3.958	-2.951
hergm term 1: parameter of block 2:	-4.810	-2.839	-0.252
hergm term 1: parameter of block 3:	-4.784	-2.761	-0.606
hergm term 1: parameter of block 4:	-4.702	-2.550	-0.209
hergm term 1: parameter of block 5:	-4.757	-2.654	-0.105
hergm term 1: between-block parameter:	-6.485	-5.464	-4.521
hergm term 2: parameter of block 1:	0.635	1.016	1.311
hergm term 2: parameter of block 2:	-1.270	1.102	3.120
hergm term 2: parameter of block 3:	-1.516	0.753	2.822
hergm term 2: parameter of block 4:	-1.522	0.691	2.961
hergm term 2: parameter of block 5:	-1.522	0.777	2.926
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	0.302	0.867	1.447
ergm term 2 parameter:	0.213	0.831	1.394
ergm term 3 parameter:	-0.415	1.923	2.992
ergm term 4 parameter:	0.618	1.397	1.962
ergm term 5 parameter:	0.433	1.290	1.951
ergm term 6 parameter:	0.350	1.182	1.900
ergm term 7 parameter:	-0.247	1.035	1.838
ergm term 8 parameter:	0.181	1.307	2.082
ergm term 9 parameter:	0.107	1.016	1.742
ergm term 10 parameter:	-0.984	-0.143	0.530
ergm term 11 parameter:	-0.880	-0.326	0.268
ergm term 12 parameter:	-1.280	-0.685	-0.145

TABLE 11. block restricted, local transitivity, $K = 6$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.505	1.358	2.935
Mean of parameters of hergm term 1: $\mu_{w,\alpha}$	-3.505	-2.380	-1.170
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-0.241	0.891	1.986
Precision of parameters of hergm term 1: $\sigma_{w,\alpha}$	0.424	0.870	1.540
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$	0.547	1.034	1.773
hergm term 1: parameter of block 1:	-5.199	-4.187	-3.121
hergm term 1: parameter of block 2:	-4.781	-3.282	-1.360
hergm term 1: parameter of block 3:	-4.816	-2.749	-0.347
hergm term 1: parameter of block 4:	-4.680	-2.606	-0.335
hergm term 1: parameter of block 5:	-4.645	-2.469	-0.230
hergm term 1: parameter of block 6:	-3.680	-1.961	0.310
hergm term 1: between-block parameter:	-6.979	-5.887	-4.991
hergm term 2: parameter of block 1:	0.669	0.993	1.240
hergm term 2: parameter of block 2:	-0.057	1.358	2.665
hergm term 2: parameter of block 3:	-1.446	0.941	3.117
hergm term 2: parameter of block 4:	-1.293	0.920	3.022
hergm term 2: parameter of block 5:	-1.376	0.918	3.085
hergm term 2: parameter of block 6:	-0.500	1.138	2.876
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	0.456	1.054	1.626
ergm term 2 parameter:	0.349	0.969	1.529
ergm term 3 parameter:	1.002	2.253	3.010
ergm term 4 parameter:	0.767	1.482	2.046
ergm term 5 parameter:	0.436	1.390	1.980
ergm term 6 parameter:	0.409	1.368	1.977
ergm term 7 parameter:	0.285	1.247	2.037
ergm term 8 parameter:	0.515	1.377	2.065
ergm term 9 parameter:	0.198	0.983	1.645
ergm term 10 parameter:	-0.611	-0.008	0.437
ergm term 11 parameter:	-0.690	-0.164	0.149
ergm term 12 parameter:	-1.183	-0.534	-0.093

TABLE 12. block restricted, local transitivity, $K = 10$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.526	1.441	3.090
Mean of parameters of hergm term 1: $\mu_{w,\alpha}$	-3.629	-2.524	-1.389
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-0.283	0.832	1.886
Precision of parameters of hergm term 1: $\sigma_{w,\alpha}$	0.491	0.953	1.661
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$	0.556	1.058	1.787
hergm term 1: parameter of block 1:	-4.722	-3.979	-3.177
hergm term 1: parameter of block 2:	-4.732	-2.839	-0.631
hergm term 1: parameter of block 3:	-4.816	-2.752	-0.546
hergm term 1: parameter of block 4:	-4.854	-2.699	-0.446
hergm term 1: parameter of block 5:	-3.910	-2.975	-1.899
hergm term 1: parameter of block 6:	-4.695	-2.620	-0.434
hergm term 1: parameter of block 7:	-4.553	-2.641	-0.246
hergm term 1: parameter of block 8:	-4.774	-2.557	-0.327
hergm term 1: parameter of block 9:	-4.792	-2.502	-0.168
hergm term 1: parameter of block 10:	-4.728	-2.581	-0.331
hergm term 1: between-block parameter:	-6.558	-5.735	-5.110
hergm term 2: parameter of block 1:	0.619	0.948	1.218
hergm term 2: parameter of block 2:	-1.355	0.994	3.010
hergm term 2: parameter of block 3:	-1.325	0.909	2.955
hergm term 2: parameter of block 4:	-1.410	0.825	3.023
hergm term 2: parameter of block 5:	-0.111	0.928	1.984
hergm term 2: parameter of block 6:	-1.284	0.902	2.949
hergm term 2: parameter of block 7:	-0.871	1.236	2.847
hergm term 2: parameter of block 8:	-1.372	0.817	2.999
hergm term 2: parameter of block 9:	-1.410	0.867	3.120
hergm term 2: parameter of block 10:	-1.403	0.801	3.035
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	0.602	1.038	1.390
ergm term 2 parameter:	0.425	0.887	1.242
ergm term 3 parameter:	0.628	1.956	3.016
ergm term 4 parameter:	0.716	1.321	1.837
ergm term 5 parameter:	0.517	1.227	1.824
ergm term 6 parameter:	0.443	1.289	1.941
ergm term 7 parameter:	0.395	1.202	2.122
ergm term 8 parameter:	0.552	1.298	2.103
ergm term 9 parameter:	0.105	0.905	1.520
ergm term 10 parameter:	-0.665	-0.004	0.503
ergm term 11 parameter:	-0.652	-0.201	0.206
ergm term 12 parameter:	-1.042	-0.548	-0.229

TABLE 13. block restricted, global transitivity, $K = 1$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.025	0.718	3.658
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-4.191	-2.304	-0.307
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.213	0.482	0.970
hergm term 1: parameter of block 1:	-7.681	-7.004	-6.224
ergm term 1 parameter:	0.762	1.127	1.396
ergm term 2 parameter:	1.661	2.330	3.067
ergm term 3 parameter:	1.691	2.296	3.115
ergm term 4 parameter:	1.526	3.136	4.183
ergm term 5 parameter:	1.625	2.247	2.927
ergm term 6 parameter:	0.690	1.539	2.292
ergm term 7 parameter:	0.746	1.748	2.491
ergm term 8 parameter:	0.120	1.290	2.182
ergm term 9 parameter:	-0.184	1.749	2.700
ergm term 10 parameter:	0.824	1.721	2.543
ergm term 11 parameter:	0.131	0.861	1.384
ergm term 12 parameter:	0.298	0.849	1.354
ergm term 13 parameter:	-0.145	0.277	0.680

TABLE 14. block restricted, global transitivity, $K = 2$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.061	0.452	1.520
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-4.935	-2.961	-0.814
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.124	0.332	0.861
hergm term 1: parameter of block 1:	-11.126	-8.847	-6.777
hergm term 1: parameter of block 2:	-8.287	-5.919	-3.954
hergm term 1: between-block parameter:	-5.227	-4.536	-3.878
ergm term 1 parameter:	0.742	1.016	1.271
ergm term 2 parameter:	2.788	4.588	6.757
ergm term 3 parameter:	2.779	4.689	6.722
ergm term 4 parameter:	0.759	1.917	2.990
ergm term 5 parameter:	1.302	1.825	2.359
ergm term 6 parameter:	0.853	1.581	2.296
ergm term 7 parameter:	0.737	1.674	2.276
ergm term 8 parameter:	-0.200	1.269	2.343
ergm term 9 parameter:	0.859	1.799	2.553
ergm term 10 parameter:	0.773	1.779	2.530
ergm term 11 parameter:	-0.037	0.619	1.152
ergm term 12 parameter:	0.134	0.614	1.199
ergm term 13 parameter:	-0.404	-0.028	0.344

TABLE 15. block restricted, global transitivity, $K = 3$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.058	0.469	1.834
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-4.818	-2.960	-0.913
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.174	0.510	1.148
hergm term 1: parameter of block 1:	-10.314	-7.070	-5.340
hergm term 1: parameter of block 2:	-8.224	-4.943	-0.602
hergm term 1: parameter of block 3:	-6.208	-3.371	-0.261
hergm term 1: between-block parameter:	-5.164	-4.330	-1.824
ergm term 1 parameter:	0.692	0.924	1.174
ergm term 2 parameter:	0.959	1.512	3.011
ergm term 3 parameter:	1.363	2.994	6.605
ergm term 4 parameter:	0.584	1.969	3.252
ergm term 5 parameter:	1.373	2.023	2.743
ergm term 6 parameter:	0.911	1.626	2.188
ergm term 7 parameter:	1.017	1.714	2.486
ergm term 8 parameter:	0.146	1.148	2.042
ergm term 9 parameter:	0.551	1.474	2.532
ergm term 10 parameter:	-0.329	1.426	2.136
ergm term 11 parameter:	-0.102	0.373	0.965
ergm term 12 parameter:	-0.004	0.410	0.903
ergm term 13 parameter:	-0.739	-0.162	0.340

TABLE 16. block restricted, global transitivity, $K = 4$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.043	0.188	0.808
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.863	-2.236	-0.516
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.323	0.674	1.269
hergm term 1: parameter of block 1:	-6.108	-5.572	-5.064
hergm term 1: parameter of block 2:	-4.861	-2.269	0.387
hergm term 1: parameter of block 3:	-4.786	-2.215	0.453
hergm term 1: parameter of block 4:	-4.749	-2.244	0.346
hergm term 1: between-block parameter:	-4.512	-2.662	0.053
ergm term 1 parameter:	0.938	1.157	1.392
ergm term 2 parameter:	0.969	1.450	1.911
ergm term 3 parameter:	1.010	1.460	1.895
ergm term 4 parameter:	1.248	2.293	3.077
ergm term 5 parameter:	1.307	1.800	2.234
ergm term 6 parameter:	0.838	1.490	2.005
ergm term 7 parameter:	0.761	1.422	1.890
ergm term 8 parameter:	0.251	1.270	1.940
ergm term 9 parameter:	0.811	1.532	2.142
ergm term 10 parameter:	0.580	1.306	1.854
ergm term 11 parameter:	-0.069	0.480	0.929
ergm term 12 parameter:	-0.055	0.359	0.715
ergm term 13 parameter:	-0.508	-0.098	0.216

TABLE 17. block restricted, global transitivity, $K = 5$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.127	0.750	2.462
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-4.916	-3.092	-0.983
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.091	0.386	1.063
hergm term 1: parameter of block 1:	-15.020	-8.588	-6.121
hergm term 1: parameter of block 2:	-7.427	-5.048	-1.639
hergm term 1: parameter of block 3:	-7.856	-3.899	-0.073
hergm term 1: parameter of block 4:	-6.774	-3.327	-0.146
hergm term 1: parameter of block 5:	-6.529	-3.286	-0.268
hergm term 1: between-block parameter:	-5.042	-4.415	-3.687
ergm term 1 parameter:	0.690	0.896	1.108
ergm term 2 parameter:	0.567	1.244	2.071
ergm term 3 parameter:	2.571	4.848	11.021
ergm term 4 parameter:	0.508	1.713	3.134
ergm term 5 parameter:	0.985	1.753	2.413
ergm term 6 parameter:	0.738	1.478	2.116
ergm term 7 parameter:	0.711	1.404	2.034
ergm term 8 parameter:	0.129	1.366	2.419
ergm term 9 parameter:	0.933	1.751	2.886
ergm term 10 parameter:	0.795	1.451	2.143
ergm term 11 parameter:	-0.144	0.365	0.961
ergm term 12 parameter:	-0.162	0.282	0.894
ergm term 13 parameter:	-0.806	-0.224	0.275

TABLE 18. block restricted, global transitivity, $K = 6$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.215	1.073	2.926
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-4.535	-2.340	-0.075
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.029	0.113	0.845
hergm term 1: parameter of block 1:	-7.876	-4.463	-0.601
hergm term 1: parameter of block 2:	-26.801	-15.546	-6.458
hergm term 1: parameter of block 3:	-8.851	-4.050	0.228
hergm term 1: parameter of block 4:	-8.236	-3.075	0.977
hergm term 1: parameter of block 5:	-7.762	-3.018	0.844
hergm term 1: parameter of block 6:	-8.143	-2.981	0.837
hergm term 1: between-block parameter:	-5.050	-4.323	-3.504
ergm term 1 parameter:	0.609	0.863	1.070
ergm term 2 parameter:	0.458	1.021	1.661
ergm term 3 parameter:	2.671	11.838	22.826
ergm term 4 parameter:	-0.135	1.680	2.979
ergm term 5 parameter:	0.875	1.645	2.248
ergm term 6 parameter:	0.331	1.357	2.009
ergm term 7 parameter:	0.434	1.352	1.902
ergm term 8 parameter:	-0.296	1.079	1.984
ergm term 9 parameter:	0.223	1.378	2.200
ergm term 10 parameter:	0.560	1.446	2.204
ergm term 11 parameter:	-0.544	0.292	0.886
ergm term 12 parameter:	-0.399	0.157	0.751
ergm term 13 parameter:	-0.884	-0.274	0.368

TABLE 19. block restricted, global transitivity, $K = 10$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.166	0.495	1.889
Mean of parameter of hergm term 2: $\mu_{w,\alpha}$	-4.654	-3.007	-1.167
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.338	0.700	1.296
hergm term 1: parameter of block 1:	-7.464	-6.020	-5.141
hergm term 1: parameter of block 2:	-7.375	-4.553	-0.528
hergm term 1: parameter of block 3:	-5.702	-3.044	-0.281
hergm term 1: parameter of block 4:	-5.669	-3.019	-0.373
hergm term 1: parameter of block 5:	-5.596	-2.964	-0.300
hergm term 1: parameter of block 6:	-5.616	-2.991	-0.257
hergm term 1: parameter of block 7:	-5.491	-3.004	-0.346
hergm term 1: parameter of block 8:	-5.563	-3.011	-0.308
hergm term 1: parameter of block 9:	-5.589	-3.008	-0.330
hergm term 1: parameter of block 10:	-5.571	-2.980	-0.284
hergm term 1: between-block parameter:	-4.706	-4.025	-1.415
ergm term 1 parameter:	0.707	0.925	1.209
ergm term 2 parameter:	0.877	1.421	1.841
ergm term 3 parameter:	1.307	2.280	3.705
ergm term 4 parameter:	0.054	2.031	2.788
ergm term 5 parameter:	1.322	1.928	3.020
ergm term 6 parameter:	0.815	1.481	2.000
ergm term 7 parameter:	0.844	1.545	2.227
ergm term 8 parameter:	0.658	1.483	2.584
ergm term 9 parameter:	0.031	1.445	2.193
ergm term 10 parameter:	0.903	1.446	2.118
ergm term 11 parameter:	-0.181	0.407	1.012
ergm term 12 parameter:	-0.187	0.352	0.742
ergm term 13 parameter:	-0.712	-0.232	0.289

TABLE 20. block unrestricted, local transitivity, $K = 1$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.025	0.716	3.704
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-2.240	-0.759	0.727
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-1.000	0.432	1.858
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.480	0.961	1.686
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$ 0.500	0.985	1.713	
hergm term 1: parameter of block 1:	-2.069	-1.585	-0.909
hergm term 2: parameter of block 1:	0.148	0.904	1.377
ergm term 1 parameter:	-0.396	0.423	1.073
ergm term 2 parameter:	-0.319	0.461	1.100
ergm term 3 parameter:	-1.107	1.313	2.856
ergm term 4 parameter:	0.023	1.140	1.949
ergm term 5 parameter:	-0.399	0.986	2.034
ergm term 6 parameter:	-0.466	0.840	1.743
ergm term 7 parameter:	-0.977	0.777	1.933
ergm term 8 parameter:	-0.878	1.002	2.042
ergm term 9 parameter:	-0.583	0.801	1.884
ergm term 10 parameter:	-1.915	-0.738	0.130
ergm term 11 parameter:	-1.659	-0.693	-0.016
ergm term 12 parameter:	-2.461	-1.378	-0.677

TABLE 21. block unrestricted, local transitivity, $K = 2$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.043	0.307	1.054
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-1.859	0.103	2.070
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-0.728	0.807	2.367
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.000	0.000	0.054
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$ 0.452	0.926	1.675	
hergm term 1: parameter of block 1:	-3.241	-2.851	-2.474
hergm term 1: parameter of block 2:	20.273	398.882	843.026
hergm term 1: between-block parameter:	0.000	0.000	0.000
hergm term 2: parameter of block 1:	1.318	1.840	2.299
hergm term 2: parameter of block 2:	-2.013	0.584	3.639
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	0.065	2.410	3.230
ergm term 2 parameter:	1.256	2.243	3.051
ergm term 3 parameter:	2.197	3.869	5.215
ergm term 4 parameter:	1.688	2.412	3.048
ergm term 5 parameter:	1.516	2.371	3.175
ergm term 6 parameter:	0.498	1.615	2.314
ergm term 7 parameter:	-1.153	0.866	2.549
ergm term 8 parameter:	0.730	2.318	3.290
ergm term 9 parameter:	0.782	2.562	3.591
ergm term 10 parameter:	-10.131	-7.462	0.121
ergm term 11 parameter:	-0.086	0.376	0.814
ergm term 12 parameter:	-4.759	-3.539	-0.737

TABLE 22. block unrestricted, local transitivity, $K = 3$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.430	1.854	5.009
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-1.925	-0.010	1.996
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-1.854	0.119	2.005
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.000	0.000	0.000
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$ 0.000	0.000	0.001	
hergm term 1: parameter of block 1:	2113.920	5747.676	9386.677
hergm term 1: parameter of block 2:	-2625.832	-1443.274	-346.442
hergm term 1: parameter of block 3:	-1.389	-0.692	0.047
hergm term 1: between-block parameter:	0.000	0.000	0.000
hergm term 2: parameter of block 1:	89.809	243.393	365.219
hergm term 2: parameter of block 2:	158.017	169.733	180.386
hergm term 2: parameter of block 3:	-0.292	1.429	3.735
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	17.196	20.430	23.722
ergm term 2 parameter:	-8.654	-6.799	-5.104
ergm term 3 parameter:	8.342	10.663	12.818
ergm term 4 parameter:	22.925	26.738	30.011
ergm term 5 parameter:	9.173	10.976	12.725
ergm term 6 parameter:	-9.946	-8.151	-5.512
ergm term 7 parameter:	8.853	11.207	13.225
ergm term 8 parameter:	1.520	3.844	5.787
ergm term 9 parameter:	14.061	16.690	18.527
ergm term 10 parameter:	-41.698	-39.273	-36.643
ergm term 11 parameter:	9.259	11.323	13.307
ergm term 12 parameter:	-119.403	-116.877	-114.490

TABLE 23. block unrestricted, local transitivity, $K = 4$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.434	1.458	3.565
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-1.957	0.020	1.989
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-1.923	0.064	1.980
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.000	0.000	0.000
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$ 0.000	0.000	0.000	
hergm term 1: parameter of block 1:	-21.885	-17.364	-10.467
hergm term 1: parameter of block 2:	6.853	10.857	49.437
hergm term 1: parameter of block 3:	1468.200	5088.213	8699.978
hergm term 1: parameter of block 4:	-2951.686	-1307.000	-110.136
hergm term 1: between-block parameter:	0.000	0.000	0.000
hergm term 2: parameter of block 1:	9.874	87.145	92.848
hergm term 2: parameter of block 2:	47.134	69.337	87.363
hergm term 2: parameter of block 3:	-51.028	-26.690	-11.123
hergm term 2: parameter of block 4:	280.866	615.584	633.113
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	22.502	26.784	39.693
ergm term 2 parameter:	-24.196	-20.002	-15.894
ergm term 3 parameter:	7.630	10.366	13.588
ergm term 4 parameter:	20.153	24.625	31.499
ergm term 5 parameter:	8.995	12.804	17.870
ergm term 6 parameter:	-13.000	-7.835	-4.130
ergm term 7 parameter:	7.482	11.065	13.998
ergm term 8 parameter:	-1.342	2.309	5.416
ergm term 9 parameter:	8.500	12.515	18.801
ergm term 10 parameter:	-53.667	-49.739	-46.095
ergm term 11 parameter:	13.597	17.373	37.038
ergm term 12 parameter:	-229.861	-226.243	-218.558

TABLE 24. block unrestricted, local transitivity, $K = 5$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.073	0.531	3.229
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-2.539	-1.067	0.945
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-0.756	0.714	2.037
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.311	0.869	1.610
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$ 0.494	0.975	1.714	
hergm term 1: parameter of block 1:	-2.981	-2.594	-2.102
hergm term 1: parameter of block 2:	-3.463	-0.895	3.485
hergm term 1: parameter of block 3:	-3.539	-0.944	3.762
hergm term 1: parameter of block 4:	-3.567	-0.852	3.837
hergm term 1: parameter of block 5:	-4.113	-1.601	2.177
hergm term 1: between-block parameter:	0.000	0.000	0.000
hergm term 2: parameter of block 1:	0.990	1.271	1.939
hergm term 2: parameter of block 2:	-1.871	0.671	3.121
hergm term 2: parameter of block 3:	-1.812	0.733	3.119
hergm term 2: parameter of block 4:	-1.825	0.663	3.118
hergm term 2: parameter of block 5:	-1.815	0.982	2.828
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	0.139	1.153	1.661
ergm term 2 parameter:	0.688	1.231	1.899
ergm term 3 parameter:	0.845	2.232	3.124
ergm term 4 parameter:	1.200	1.731	2.748
ergm term 5 parameter:	0.869	1.536	2.287
ergm term 6 parameter:	0.746	1.403	2.197
ergm term 7 parameter:	0.393	1.241	2.091
ergm term 8 parameter:	0.522	1.537	2.225
ergm term 9 parameter:	0.048	1.177	1.914
ergm term 10 parameter:	-0.683	0.227	0.762
ergm term 11 parameter:	-0.711	0.126	0.575
ergm term 12 parameter:	-2.368	-0.356	0.065

TABLE 25. block unrestricted, local transitivity, $K = 6$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.077	0.242	0.932
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-2.754	-1.334	0.093
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-0.836	0.583	1.979
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.480	0.922	1.627
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$ 0.510	0.982	1.700	
hergm term 1: parameter of block 1:	-2.968	-2.738	-2.531
hergm term 1: parameter of block 2:	-3.848	-1.332	1.020
hergm term 1: parameter of block 3:	-3.761	-1.351	1.037
hergm term 1: parameter of block 4:	-3.722	-1.358	1.161
hergm term 1: parameter of block 5:	-3.801	-1.347	1.082
hergm term 1: parameter of block 6:	-3.689	-1.297	1.124
hergm term 1: between-block parameter:	0.000	0.000	0.000
hergm term 2: parameter of block 1:	1.022	1.212	1.416
hergm term 2: parameter of block 2:	-1.836	0.556	2.963
hergm term 2: parameter of block 3:	-1.804	0.605	2.930
hergm term 2: parameter of block 4:	-1.845	0.574	2.991
hergm term 2: parameter of block 5:	-1.801	0.605	2.919
hergm term 2: parameter of block 6:	-1.817	0.596	3.049
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	1.031	1.376	1.771
ergm term 2 parameter:	0.997	1.330	1.729
ergm term 3 parameter:	1.357	2.190	2.902
ergm term 4 parameter:	1.352	1.739	2.088
ergm term 5 parameter:	1.076	1.509	1.921
ergm term 6 parameter:	0.929	1.428	1.840
ergm term 7 parameter:	0.577	1.242	1.781
ergm term 8 parameter:	0.928	1.564	2.039
ergm term 9 parameter:	0.668	1.270	1.757
ergm term 10 parameter:	-0.021	0.440	0.827
ergm term 11 parameter:	0.004	0.329	0.642
ergm term 12 parameter:	-0.423	-0.127	0.144

TABLE 26. block unrestricted, local transitivity, $K = 10$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.183	0.727	2.298
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-2.445	-1.092	0.422
Mean of parameter of hergm term 2: $\mu_{w,\gamma}$	-0.790	0.650	2.106
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.456	0.930	1.620
Precision of parameter of hergm term 2: $\sigma_{w,\gamma}$ 0.518	0.985	1.703	
hergm term 1: parameter of block 1:	-2.815	-2.511	-2.091
hergm term 1: parameter of block 2:	-3.549	-1.031	1.544
hergm term 1: parameter of block 3:	-3.447	-1.007	1.407
hergm term 1: parameter of block 4:	-3.359	-1.048	1.475
hergm term 1: parameter of block 5:	-3.441	-1.016	1.517
hergm term 1: parameter of block 6:	-3.457	-1.066	1.579
hergm term 1: parameter of block 7:	-3.395	-1.119	1.448
hergm term 1: parameter of block 8:	-3.551	-1.086	1.405
hergm term 1: parameter of block 9:	-3.493	-1.042	1.575
hergm term 1: parameter of block 10:	-3.456	-1.081	1.547
hergm term 1: between-block parameter:	0.000	0.000	0.000
hergm term 2: parameter of block 1:	0.997	1.270	1.852
hergm term 2: parameter of block 2:	-1.834	0.675	3.047
hergm term 2: parameter of block 3:	-1.789	0.728	3.150
hergm term 2: parameter of block 4:	-1.774	0.673	3.086
hergm term 2: parameter of block 5:	-1.829	0.652	3.080
hergm term 2: parameter of block 6:	-1.872	0.632	3.043
hergm term 2: parameter of block 7:	-1.838	0.626	3.119
hergm term 2: parameter of block 8:	-1.790	0.658	3.165
hergm term 2: parameter of block 9:	-1.787	0.664	3.121
hergm term 2: parameter of block 10:	-1.785	0.655	3.150
hergm term 2: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	0.578	1.140	1.573
ergm term 2 parameter:	0.605	1.119	1.555
ergm term 3 parameter:	0.952	2.112	2.815
ergm term 4 parameter:	1.102	1.618	2.044
ergm term 5 parameter:	0.747	1.450	1.955
ergm term 6 parameter:	0.641	1.317	1.797
ergm term 7 parameter:	0.374	1.154	1.788
ergm term 8 parameter:	0.601	1.444	2.019
ergm term 9 parameter:	0.446	1.145	1.738
ergm term 10 parameter:	-0.555	0.144	0.654
ergm term 11 parameter:	-0.625	0.107	0.499
ergm term 12 parameter:	-1.403	-0.365	0.014

TABLE 27. block unrestricted, global transitivity, $K = 1$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.027	0.685	3.631
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-2.790	-1.338	0.213
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.442	0.901	1.617
hergm term 1: parameter of block 1:	-3.113	-2.832	-2.570
ergm term 1 parameter:	1.022	1.208	1.428
ergm term 2 parameter:	1.106	1.506	1.899
ergm term 3 parameter:	1.088	1.460	1.845
ergm term 4 parameter:	1.576	2.412	3.159
ergm term 5 parameter:	1.317	1.793	2.197
ergm term 6 parameter:	1.014	1.533	1.992
ergm term 7 parameter:	0.864	1.430	1.881
ergm term 8 parameter:	0.641	1.319	1.904
ergm term 9 parameter:	0.936	1.614	2.194
ergm term 10 parameter:	0.743	1.324	1.862
ergm term 11 parameter:	0.032	0.478	0.909
ergm term 12 parameter:	0.018	0.372	0.734
ergm term 13 parameter:	-0.385	-0.070	0.222

TABLE 28. block unrestricted, global transitivity, $K = 2$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.044	0.314	1.066
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-1.915	0.006	1.965
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.000	0.000	0.000
hergm term 1: parameter of block 1:	-2.507	-2.278	-2.053
hergm term 1: parameter of block 2:	1320.302	3008.395	4693.625
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	-1.217	-0.879	-0.562
ergm term 2 parameter:	0.223	0.801	1.404
ergm term 3 parameter:	2.316	2.766	3.256
ergm term 4 parameter:	2.244	3.201	4.123
ergm term 5 parameter:	2.882	3.505	4.109
ergm term 6 parameter:	0.234	1.128	1.917
ergm term 7 parameter:	1.978	2.605	3.197
ergm term 8 parameter:	1.944	2.750	3.471
ergm term 9 parameter:	1.750	2.520	3.281
ergm term 10 parameter:	0.512	1.334	2.126
ergm term 11 parameter:	0.181	0.594	1.005
ergm term 12 parameter:	-1.972	-1.427	-0.954
ergm term 13 parameter:	-2.400	-1.909	-1.468

TABLE 29. block unrestricted, global transitivity, $K = 3$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.126	0.553	1.501
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-1.965	0.023	1.959
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.000	0.000	0.000
hergm term 1: parameter of block 1:	2.679	3.891	5.048
hergm term 1: parameter of block 2:	-71.812	-60.524	-55.409
hergm term 1: parameter of block 3:	2088.061	5820.823	9586.061
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	-27.393	-25.246	-23.401
ergm term 2 parameter:	-14.035	-11.872	-9.451
ergm term 3 parameter:	32.461	35.222	37.320
ergm term 4 parameter:	5.146	8.007	10.283
ergm term 5 parameter:	22.365	24.307	26.698
ergm term 6 parameter:	3.227	6.199	8.607
ergm term 7 parameter:	4.435	6.730	9.287
ergm term 8 parameter:	11.235	13.814	16.271
ergm term 9 parameter:	9.665	12.318	14.839
ergm term 10 parameter:	0.490	2.685	4.904
ergm term 11 parameter:	14.089	16.307	18.274
ergm term 12 parameter:	-35.623	-33.627	-31.390
ergm term 13 parameter:	-28.009	-25.636	-24.068

TABLE 30. block unrestricted, global transitivity, $K = 4$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.494	1.729	4.198
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-1.969	-0.018	1.919
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.000	0.000	0.048
hergm term 1: parameter of block 1:	-23.914	-13.875	0.815
hergm term 1: parameter of block 2:	18.227	3335.955	7037.310
hergm term 1: parameter of block 3:	-3097.586	-1677.432	-13.495
hergm term 1: parameter of block 4:	-0.646	35.172	81.310
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	-92.845	-40.919	0.024
ergm term 2 parameter:	2.102	44.462	59.335
ergm term 3 parameter:	-47.134	-20.831	2.411
ergm term 4 parameter:	2.839	10.480	13.883
ergm term 5 parameter:	3.878	26.301	31.995
ergm term 6 parameter:	2.271	16.922	22.156
ergm term 7 parameter:	-7.549	2.972	8.403
ergm term 8 parameter:	1.898	12.199	16.697
ergm term 9 parameter:	-9.671	-5.972	0.245
ergm term 10 parameter:	2.193	16.454	20.217
ergm term 11 parameter:	-80.791	-47.851	-1.202
ergm term 12 parameter:	1.390	34.479	52.723
ergm term 13 parameter:	-283.134	-179.109	-4.896

TABLE 31. block unrestricted, global transitivity, $K = 5$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.369	1.210	3.004
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-1.929	0.041	2.025
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.000	0.000	0.015
hergm term 1: parameter of block 1:	-0.912	2.554	22.882
hergm term 1: parameter of block 2:	-16.621	-9.769	-5.148
hergm term 1: parameter of block 3:	4.487	691.667	1252.156
hergm term 1: parameter of block 4:	-28.512	-14.148	-7.007
hergm term 1: parameter of block 5:	-1276.172	-757.162	240.313
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	-3.547	-1.894	0.000
ergm term 2 parameter:	2.577	8.002	9.379
ergm term 3 parameter:	2.004	9.641	11.155
ergm term 4 parameter:	-2.882	-0.889	7.699
ergm term 5 parameter:	2.146	3.321	14.303
ergm term 6 parameter:	0.357	4.584	6.011
ergm term 7 parameter:	0.778	5.240	8.411
ergm term 8 parameter:	-0.357	5.864	7.969
ergm term 9 parameter:	-0.943	1.400	7.213
ergm term 10 parameter:	2.127	5.150	9.844
ergm term 11 parameter:	5.127	13.664	34.700
ergm term 12 parameter:	-39.906	-15.360	-5.126
ergm term 13 parameter:	-43.090	-6.855	-1.078

TABLE 32. block unrestricted, global transitivity, $K = 6$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.722	1.883	3.980
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-1.939	0.046	1.988
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.000	0.000	0.000
hergm term 1: parameter of block 1:	-23.168	-21.656	-20.056
hergm term 1: parameter of block 2:	-5.265	-4.568	-3.882
hergm term 1: parameter of block 3:	-36.131	-31.056	-28.236
hergm term 1: parameter of block 4:	6.328	7.969	9.525
hergm term 1: parameter of block 5:	371.186	1072.496	1774.379
hergm term 1: parameter of block 6:	17.383	18.795	19.969
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	-4.456	-3.837	-3.256
ergm term 2 parameter:	-13.452	-12.074	-10.656
ergm term 3 parameter:	27.510	29.410	31.126
ergm term 4 parameter:	4.416	6.268	7.683
ergm term 5 parameter:	9.350	10.947	12.846
ergm term 6 parameter:	-1.150	0.598	2.645
ergm term 7 parameter:	6.284	8.048	10.098
ergm term 8 parameter:	4.141	5.933	7.526
ergm term 9 parameter:	3.468	4.998	6.946
ergm term 10 parameter:	1.738	3.462	5.330
ergm term 11 parameter:	18.115	19.707	21.071
ergm term 12 parameter:	-24.249	-22.809	-21.190
ergm term 13 parameter:	-10.522	-9.671	-8.764

TABLE 33. block unrestricted, global transitivity, $K = 10$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.296	1.175	4.120
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-2.218	0.131	2.211
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.078	0.202	0.499
hergm term 1: parameter of block 1:	-8.776	2.001	6.445
hergm term 1: parameter of block 2:	-2.836	1.174	8.278
hergm term 1: parameter of block 3:	-2.629	2.298	7.830
hergm term 1: parameter of block 4:	-10.051	-8.266	-6.064
hergm term 1: parameter of block 5:	-2.256	1.330	6.495
hergm term 1: parameter of block 6:	-2.857	0.772	8.993
hergm term 1: parameter of block 7:	-2.818	0.810	7.219
hergm term 1: parameter of block 8:	-2.875	0.419	7.422
hergm term 1: parameter of block 9:	-2.903	0.402	7.711
hergm term 1: parameter of block 10:	-2.929	0.552	7.057
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	0.009	1.003	1.311
ergm term 2 parameter:	1.187	1.613	2.090
ergm term 3 parameter:	0.923	1.416	1.875
ergm term 4 parameter:	1.166	2.147	2.953
ergm term 5 parameter:	1.313	1.831	2.455
ergm term 6 parameter:	1.058	1.714	2.398
ergm term 7 parameter:	0.946	1.593	2.181
ergm term 8 parameter:	0.509	1.384	2.091
ergm term 9 parameter:	0.348	1.453	2.176
ergm term 10 parameter:	0.697	1.390	2.175
ergm term 11 parameter:	-14.687	-10.771	-6.407
ergm term 12 parameter:	7.227	11.375	15.304
ergm term 13 parameter:	-1.941	-0.346	0.159

TABLE 34. block restricted, stochastic blockmodel, $K = 2$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.104	0.734	2.477
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.944	-2.609	-0.975
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.334	0.719	1.352
hergm term 1: parameter of block 1:	-5.941	-5.510	-5.127
hergm term 1: parameter of block 2:	-3.725	-3.253	-2.857
hergm term 1: between-block parameter:	-6.699	-6.181	-5.729
ergm term 1 parameter:	1.327	1.621	1.950
ergm term 2 parameter:	1.549	1.868	2.195
ergm term 3 parameter:	1.772	2.486	3.199
ergm term 4 parameter:	2.173	2.584	3.000
ergm term 5 parameter:	1.268	1.768	2.238
ergm term 6 parameter:	1.133	1.641	2.069
ergm term 7 parameter:	0.737	1.365	1.944
ergm term 8 parameter:	0.270	0.999	1.663
ergm term 9 parameter:	0.491	1.117	1.770
ergm term 10 parameter:	0.039	0.387	0.730
ergm term 11 parameter:	0.031	0.309	0.615
ergm term 12 parameter:	-0.420	-0.173	0.052

TABLE 35. block restricted, stochastic blockmodel, $K = 3$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.231	1.014	2.784
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.745	-2.582	-1.209
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.380	0.776	1.402
hergm term 1: parameter of block 1:	-5.743	-5.316	-4.907
hergm term 1: parameter of block 2:	-3.568	-2.910	-2.224
hergm term 1: parameter of block 3:	-3.487	-2.899	-2.306
hergm term 1: between-block parameter:	-6.592	-6.094	-5.627
ergm term 1 parameter:	0.971	1.373	1.756
ergm term 2 parameter:	1.293	1.675	2.075
ergm term 3 parameter:	1.519	2.397	3.087
ergm term 4 parameter:	2.090	2.572	3.015
ergm term 5 parameter:	1.255	1.778	2.264
ergm term 6 parameter:	1.162	1.676	2.132
ergm term 7 parameter:	0.610	1.272	1.855
ergm term 8 parameter:	0.669	1.296	1.904
ergm term 9 parameter:	0.625	1.340	1.951
ergm term 10 parameter:	0.065	0.430	0.748
ergm term 11 parameter:	0.012	0.330	0.615
ergm term 12 parameter:	-0.500	-0.206	0.045

TABLE 36. block restricted, stochastic blockmodel, $K = 4$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.345	1.282	3.732
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.789	-2.711	-1.454
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.414	0.846	1.500
hergm term 1: parameter of block 1:	-5.662	-5.059	-2.573
hergm term 1: parameter of block 2:	-3.626	-2.930	-2.156
hergm term 1: parameter of block 3:	-3.505	-2.890	-2.258
hergm term 1: parameter of block 4:	-5.379	-3.573	-1.007
hergm term 1: between-block parameter:	-6.761	-6.239	-5.757
ergm term 1 parameter:	0.993	1.385	1.808
ergm term 2 parameter:	1.272	1.700	2.054
ergm term 3 parameter:	1.522	2.444	3.117
ergm term 4 parameter:	2.030	2.513	2.976
ergm term 5 parameter:	1.203	1.759	2.310
ergm term 6 parameter:	1.131	1.636	2.131
ergm term 7 parameter:	0.508	1.254	1.804
ergm term 8 parameter:	0.521	1.264	1.944
ergm term 9 parameter:	0.711	1.329	1.911
ergm term 10 parameter:	0.077	0.435	0.790
ergm term 11 parameter:	0.008	0.306	0.635
ergm term 12 parameter:	-0.484	-0.233	0.008

TABLE 37. block restricted, stochastic blockmodel, $K = 5$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.464	1.495	3.997
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.957	-2.860	-1.498
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.367	0.811	1.510
hergm term 1: parameter of block 1:	-5.944	-4.716	-2.086
hergm term 1: parameter of block 2:	-5.619	-4.394	-2.342
hergm term 1: parameter of block 3:	-3.768	-3.316	-2.876
hergm term 1: parameter of block 4:	-5.047	-2.806	0.664
hergm term 1: parameter of block 5:	-5.123	-3.203	-0.791
hergm term 1: between-block parameter:	-7.046	-6.572	-6.067
ergm term 1 parameter:	1.360	1.702	2.065
ergm term 2 parameter:	1.538	1.912	2.245
ergm term 3 parameter:	1.740	2.571	3.341
ergm term 4 parameter:	1.718	2.473	3.020
ergm term 5 parameter:	1.146	1.756	2.309
ergm term 6 parameter:	1.017	1.578	2.102
ergm term 7 parameter:	0.639	1.262	1.858
ergm term 8 parameter:	0.295	0.987	1.626
ergm term 9 parameter:	0.485	1.045	1.688
ergm term 10 parameter:	0.018	0.379	0.758
ergm term 11 parameter:	-0.018	0.270	0.590
ergm term 12 parameter:	-0.457	-0.188	0.055

TABLE 38. block restricted, stochastic blockmodel, $K = 6$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.713	2.197	5.456
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.750	-2.756	-1.403
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.401	0.869	1.597
hergm term 1: parameter of block 1:	-3.624	-2.963	-2.256
hergm term 1: parameter of block 2:	-5.561	-3.527	-1.190
hergm term 1: parameter of block 3:	-3.510	-2.930	-2.330
hergm term 1: parameter of block 4:	-5.833	-3.240	0.320
hergm term 1: parameter of block 5:	-5.008	-3.140	0.314
hergm term 1: parameter of block 6:	-5.471	-4.342	-0.430
hergm term 1: between-block parameter:	-6.799	-6.347	-5.905
ergm term 1 parameter:	1.010	1.446	1.858
ergm term 2 parameter:	1.312	1.721	2.136
ergm term 3 parameter:	1.581	2.508	3.193
ergm term 4 parameter:	1.974	2.592	3.073
ergm term 5 parameter:	1.156	1.747	2.294
ergm term 6 parameter:	1.118	1.639	2.157
ergm term 7 parameter:	0.465	1.164	1.759
ergm term 8 parameter:	0.653	1.354	1.999
ergm term 9 parameter:	0.565	1.290	1.929
ergm term 10 parameter:	0.042	0.430	0.781
ergm term 11 parameter:	-0.039	0.294	0.597
ergm term 12 parameter:	-0.477	-0.216	0.021

TABLE 39. block restricted, stochastic blockmodel, $K = 10$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	1.018	2.539	5.467
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.422	-2.572	-1.536
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.400	0.832	1.540
hergm term 1: parameter of block 1:	-5.696	-4.012	-1.227
hergm term 1: parameter of block 2:	-3.692	-2.982	-2.292
hergm term 1: parameter of block 3:	-3.504	-2.901	-2.259
hergm term 1: parameter of block 4:	-5.571	-3.395	-0.454
hergm term 1: parameter of block 5:	-5.365	-3.240	-1.245
hergm term 1: parameter of block 6:	-3.837	-2.226	-0.252
hergm term 1: parameter of block 7:	-5.231	-3.158	-0.586
hergm term 1: parameter of block 8:	-4.892	-2.553	0.337
hergm term 1: parameter of block 9:	-4.220	-1.520	0.826
hergm term 1: parameter of block 10:	-5.521	-3.228	-0.700
hergm term 1: between-block parameter:	-6.813	-6.331	-5.885
ergm term 1 parameter:	0.987	1.477	1.900
ergm term 2 parameter:	1.365	1.788	2.221
ergm term 3 parameter:	1.625	2.601	3.455
ergm term 4 parameter:	1.699	2.586	3.175
ergm term 5 parameter:	1.275	1.839	2.350
ergm term 6 parameter:	0.981	1.589	2.140
ergm term 7 parameter:	0.358	1.111	1.827
ergm term 8 parameter:	0.576	1.261	2.008
ergm term 9 parameter:	0.577	1.285	1.929
ergm term 10 parameter:	-0.001	0.437	0.797
ergm term 11 parameter:	-0.069	0.261	0.574
ergm term 12 parameter:	-0.512	-0.242	-0.004

TABLE 40. block unrestricted, stochastic blockmodel, $K = 2$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.161	1.102	3.679
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-2.958	-1.769	-0.412
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.456	0.919	1.627
hergm term 1: parameter of block 1:	-2.371	-2.118	-1.875
hergm term 1: parameter of block 2:	-3.653	-3.331	-3.067
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	1.455	1.838	2.214
ergm term 2 parameter:	1.664	2.051	2.457
ergm term 3 parameter:	1.358	2.156	2.904
ergm term 4 parameter:	2.123	2.650	3.216
ergm term 5 parameter:	1.170	1.877	2.605
ergm term 6 parameter:	1.016	1.513	2.067
ergm term 7 parameter:	0.595	1.302	1.901
ergm term 8 parameter:	0.915	1.547	2.064
ergm term 9 parameter:	1.098	1.625	2.198
ergm term 10 parameter:	0.050	0.433	0.835
ergm term 11 parameter:	0.079	0.429	0.756
ergm term 12 parameter:	-0.552	-0.292	-0.014

TABLE 41. block unrestricted, stochastic blockmodel, $K = 3$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.123	1.121	3.795
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.082	-1.992	-0.760
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.465	0.942	1.685
hergm term 1: parameter of block 1:	-4.074	-3.239	-1.792
hergm term 1: parameter of block 2:	-3.960	-2.270	-1.488
hergm term 1: parameter of block 3:	-3.875	-2.311	-0.898
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	1.462	1.871	2.262
ergm term 2 parameter:	1.685	2.073	2.467
ergm term 3 parameter:	1.291	2.101	2.905
ergm term 4 parameter:	2.158	2.689	3.220
ergm term 5 parameter:	1.152	1.830	2.467
ergm term 6 parameter:	1.025	1.566	2.108
ergm term 7 parameter:	0.667	1.274	1.866
ergm term 8 parameter:	0.914	1.567	2.200
ergm term 9 parameter:	1.087	1.665	2.224
ergm term 10 parameter:	0.016	0.431	0.808
ergm term 11 parameter:	0.062	0.410	0.752
ergm term 12 parameter:	-0.538	-0.268	-0.010

TABLE 42. block unrestricted, stochastic blockmodel, $K = 4$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.150	1.212	4.055
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.199	-2.142	-0.942
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.479	0.962	1.692
hergm term 1: parameter of block 1:	-4.174	-2.313	-1.774
hergm term 1: parameter of block 2:	-4.942	-3.234	-1.649
hergm term 1: parameter of block 3:	-4.384	-2.352	-0.602
hergm term 1: parameter of block 4:	-4.104	-2.543	-0.590
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	1.463	1.848	2.239
ergm term 2 parameter:	1.681	2.062	2.464
ergm term 3 parameter:	1.227	2.071	2.860
ergm term 4 parameter:	2.138	2.706	3.236
ergm term 5 parameter:	1.170	1.828	2.478
ergm term 6 parameter:	0.989	1.536	2.087
ergm term 7 parameter:	0.685	1.320	1.905
ergm term 8 parameter:	0.902	1.544	2.163
ergm term 9 parameter:	1.056	1.645	2.227
ergm term 10 parameter:	0.023	0.420	0.822
ergm term 11 parameter:	0.070	0.401	0.739
ergm term 12 parameter:	-0.548	-0.283	-0.020

TABLE 43. block unrestricted, stochastic blockmodel, $K = 5$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.233	1.539	4.530
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.213	-2.281	-1.159
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.504	0.983	1.706
hergm term 1: parameter of block 1:	-3.900	-2.619	-1.737
hergm term 1: parameter of block 2:	-4.374	-2.986	-1.674
hergm term 1: parameter of block 3:	-4.889	-2.854	-1.389
hergm term 1: parameter of block 4:	-4.382	-2.614	-0.673
hergm term 1: parameter of block 5:	-4.052	-2.718	-0.837
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	1.451	1.868	2.262
ergm term 2 parameter:	1.682	2.060	2.473
ergm term 3 parameter:	1.188	2.080	2.859
ergm term 4 parameter:	2.150	2.696	3.265
ergm term 5 parameter:	1.133	1.818	2.459
ergm term 6 parameter:	1.011	1.539	2.073
ergm term 7 parameter:	0.642	1.319	1.973
ergm term 8 parameter:	0.862	1.537	2.172
ergm term 9 parameter:	1.041	1.666	2.245
ergm term 10 parameter:	0.020	0.421	0.822
ergm term 11 parameter:	0.060	0.410	0.740
ergm term 12 parameter:	-0.538	-0.271	-0.001

TABLE 44. block unrestricted, stochastic blockmodel, $K = 6$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.155	1.067	4.334
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.204	-2.224	-1.014
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.509	0.994	1.720
hergm term 1: parameter of block 1:	-4.055	-3.025	-1.819
hergm term 1: parameter of block 2:	-4.169	-2.814	-1.771
hergm term 1: parameter of block 3:	-4.049	-2.392	-0.678
hergm term 1: parameter of block 4:	-4.236	-2.388	-0.216
hergm term 1: parameter of block 5:	-4.821	-2.440	-0.231
hergm term 1: parameter of block 6:	-4.415	-2.297	-0.186
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	1.462	1.847	2.234
ergm term 2 parameter:	1.650	2.040	2.468
ergm term 3 parameter:	1.264	2.092	2.916
ergm term 4 parameter:	2.153	2.688	3.266
ergm term 5 parameter:	1.064	1.764	2.446
ergm term 6 parameter:	0.959	1.545	2.087
ergm term 7 parameter:	0.643	1.301	1.940
ergm term 8 parameter:	0.954	1.571	2.148
ergm term 9 parameter:	1.045	1.628	2.230
ergm term 10 parameter:	0.026	0.435	0.822
ergm term 11 parameter:	0.041	0.385	0.734
ergm term 12 parameter:	-0.562	-0.287	-0.022

TABLE 45. block unrestricted, stochastic blockmodel, $K = 10$

Posterior quantiles	2.5%	50%	97.5%
Concentration parameter ϕ:	0.219	0.961	3.633
Mean of parameter of hergm term 1: $\mu_{w,\alpha}$	-3.253	-2.298	-1.080
Precision of parameter of hergm term 1: $\sigma_{w,\alpha}$	0.532	0.999	1.743
hergm term 1: parameter of block 1:	-3.796	-2.750	-1.871
hergm term 1: parameter of block 2:	-4.086	-3.137	-1.768
hergm term 1: parameter of block 3:	-4.853	-2.408	-0.840
hergm term 1: parameter of block 4:	-4.491	-2.518	-0.360
hergm term 1: parameter of block 5:	-4.688	-2.455	-0.149
hergm term 1: parameter of block 6:	-4.523	-2.373	-0.112
hergm term 1: parameter of block 7:	-4.541	-2.424	-0.068
hergm term 1: parameter of block 8:	-4.598	-2.352	-0.036
hergm term 1: parameter of block 9:	-4.402	-2.279	-0.018
hergm term 1: parameter of block 10:	-4.557	-2.351	-0.035
hergm term 1: between-block parameter:	0.000	0.000	0.000
ergm term 1 parameter:	1.445	1.845	2.231
ergm term 2 parameter:	1.653	2.044	2.474
ergm term 3 parameter:	1.233	2.046	2.788
ergm term 4 parameter:	2.111	2.682	3.232
ergm term 5 parameter:	1.170	1.794	2.477
ergm term 6 parameter:	1.003	1.536	2.089
ergm term 7 parameter:	0.657	1.279	1.885
ergm term 8 parameter:	0.894	1.545	2.156
ergm term 9 parameter:	1.054	1.650	2.279
ergm term 10 parameter:	0.018	0.422	0.805
ergm term 11 parameter:	0.039	0.398	0.741
ergm term 12 parameter:	-0.556	-0.276	-0.011

Appendix D. Goodness of fit simulations

Figure 7. Restricted, local transitivity, $K = 2$, Goodness of fit

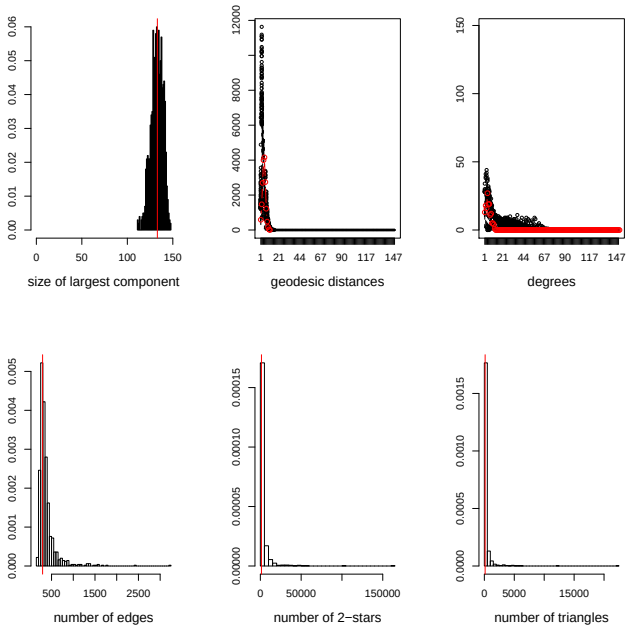


Figure 8. Restricted, local transitivity, $K = 3$, Goodness of fit

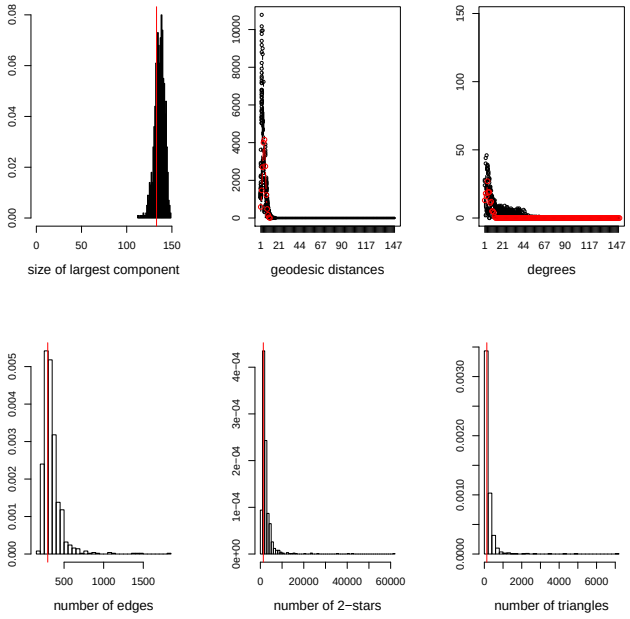


FIGURE 9. Restricted, local transitivity, $K = 4$, Goodness of fit

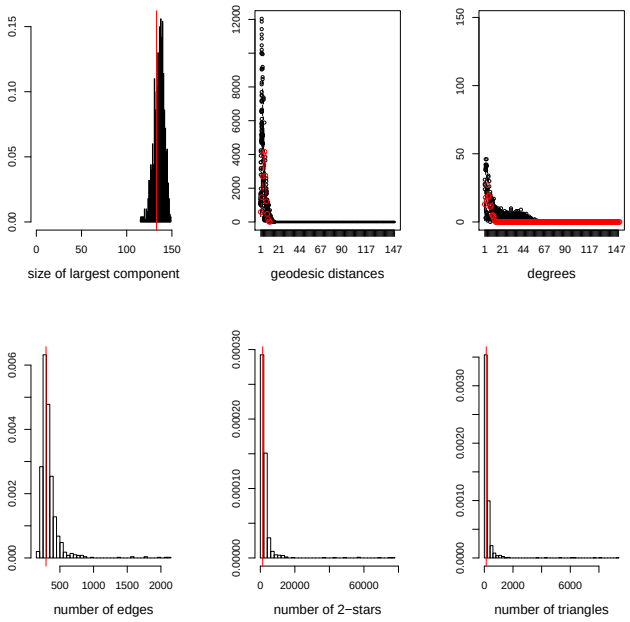


FIGURE 10. Restricted, local transitivity, $K = 5$, Goodness of fit

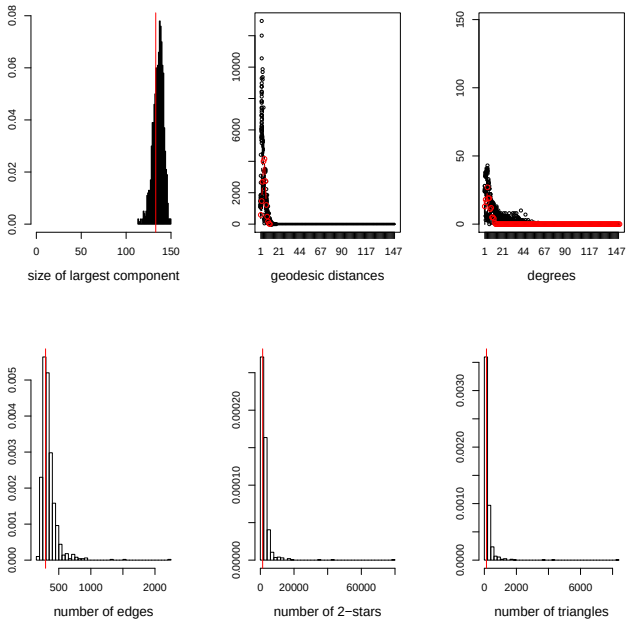


FIGURE 11. Restricted, local transitivity, $K = 6$, Goodness of fit

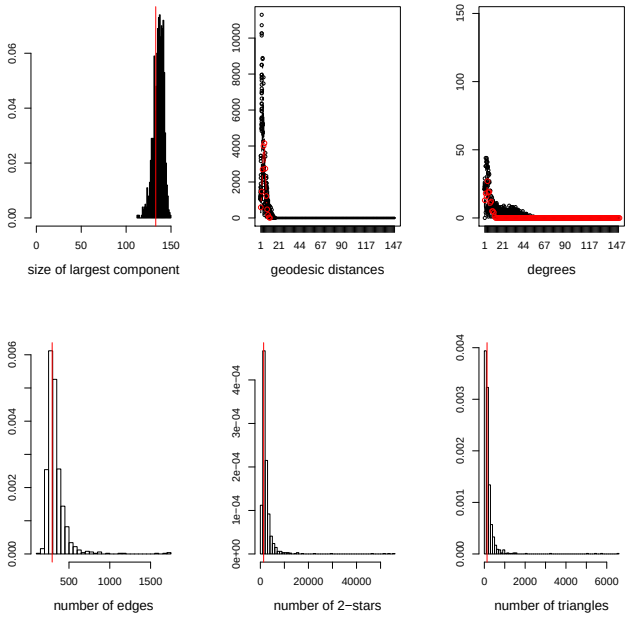


FIGURE 12. Restricted, local transitivity, $K = 10$, Goodness of fit

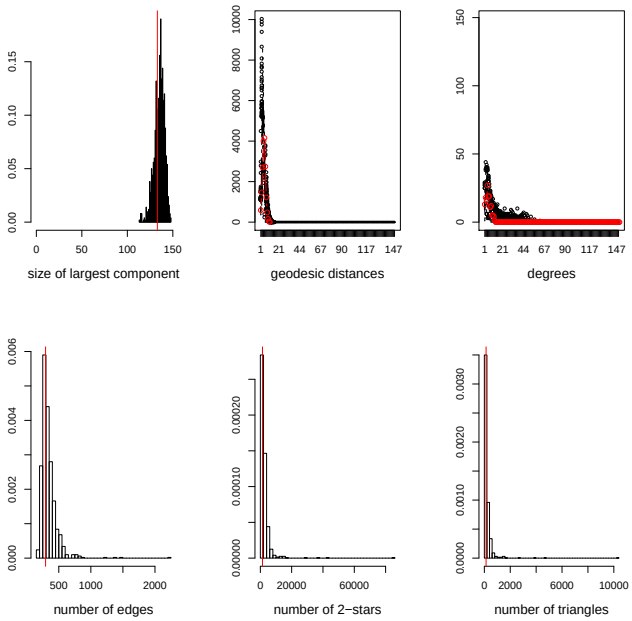


FIGURE 13. Restricted, global transitivity, $K = 2$, Goodness of fit

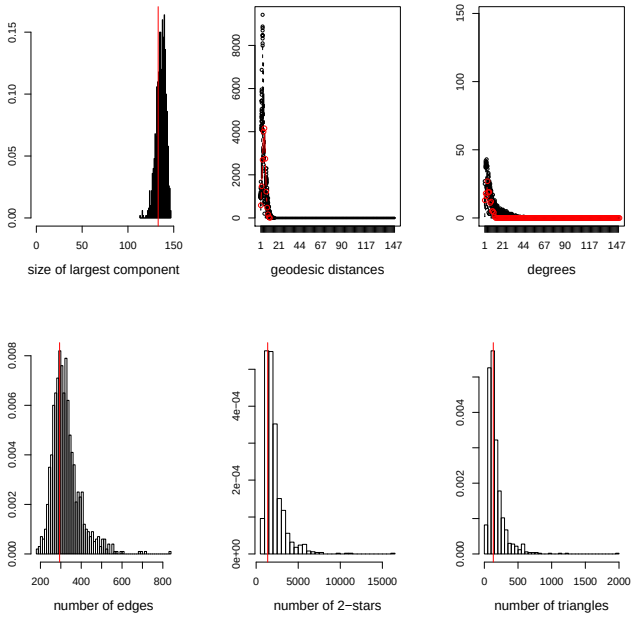


FIGURE 14. Restricted, global transitivity, $K = 3$, Goodness of fit

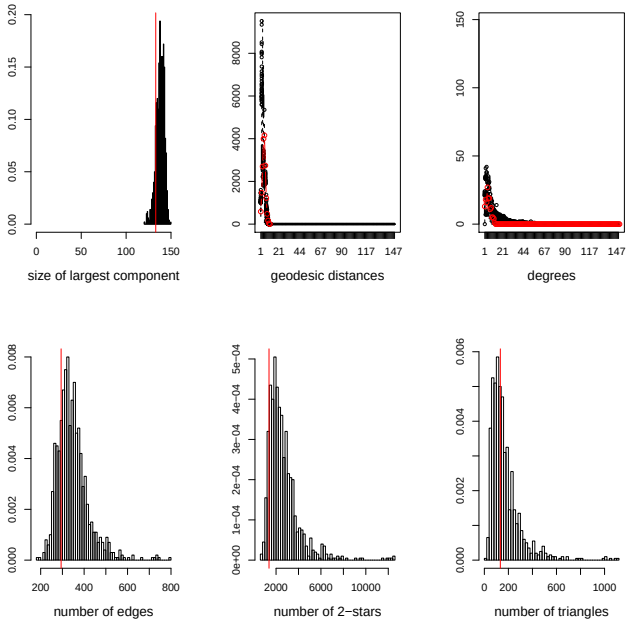


FIGURE 15. Restricted, global transitivity, $K = 4$, Goodness of fit

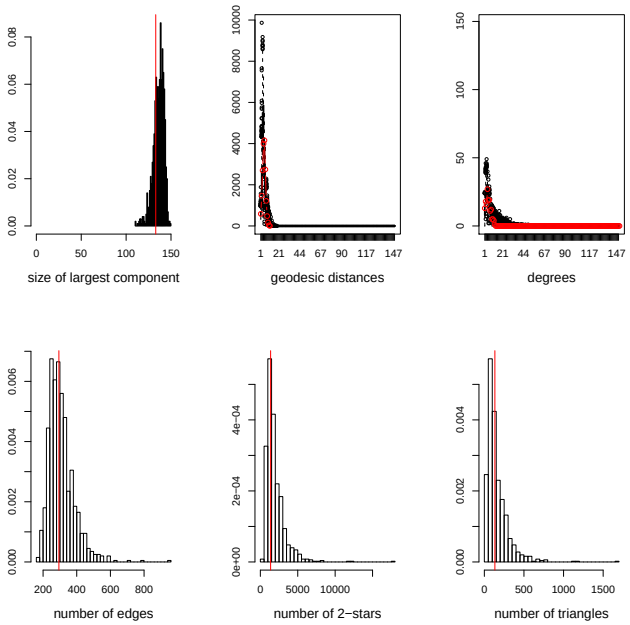


FIGURE 16. Restricted, global transitivity, $K = 5$, Goodness of fit

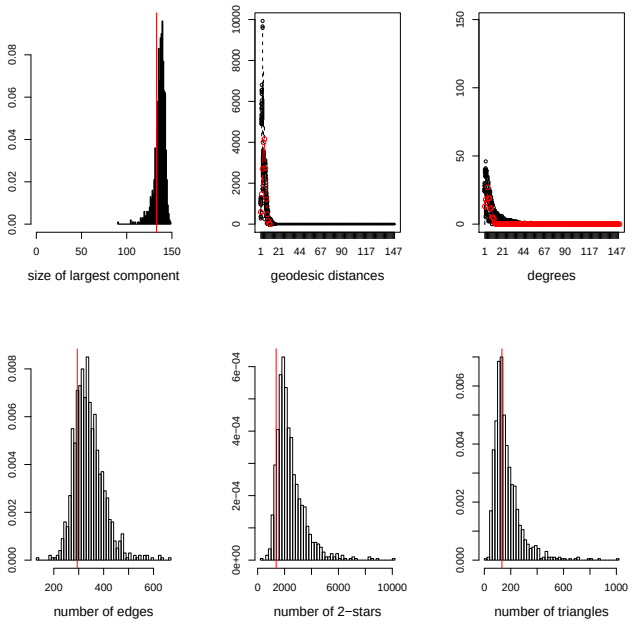


FIGURE 17. Restricted, global transitivity, $K = 6$, Goodness of fit

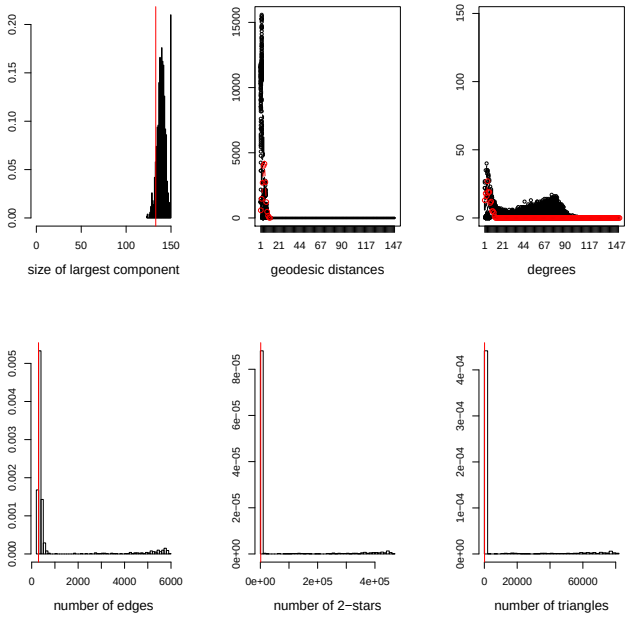


FIGURE 18. Restricted, global transitivity, $K = 10$, Goodness of fit

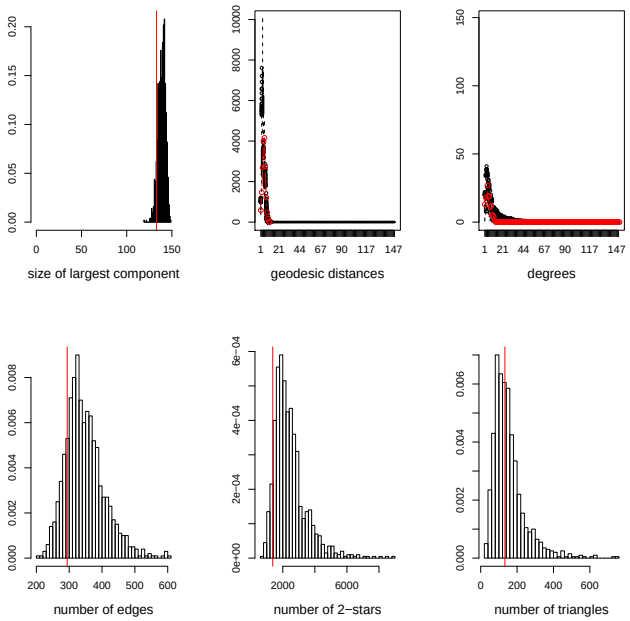


FIGURE 19. Unrestricted, local transitivity, $K = 2$, Goodness of fit

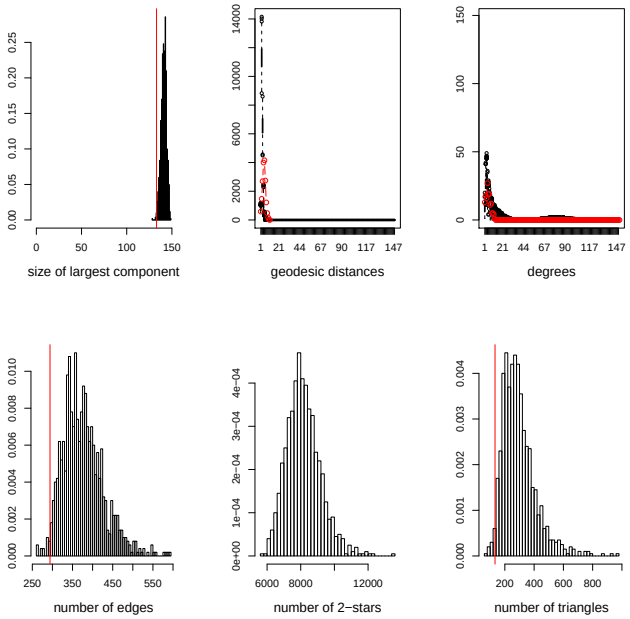


FIGURE 20. Unrestricted, local transitivity, $K = 3$, Goodness of fit

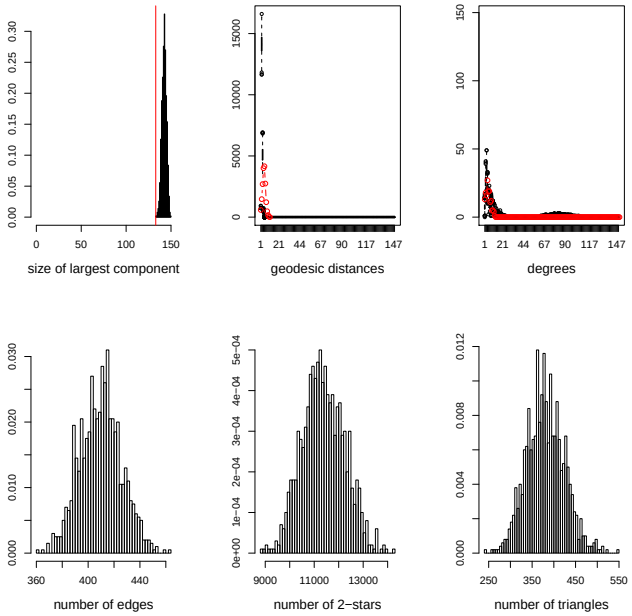


FIGURE 21. Unrestricted, local transitivity, $K = 4$, Goodness of fit

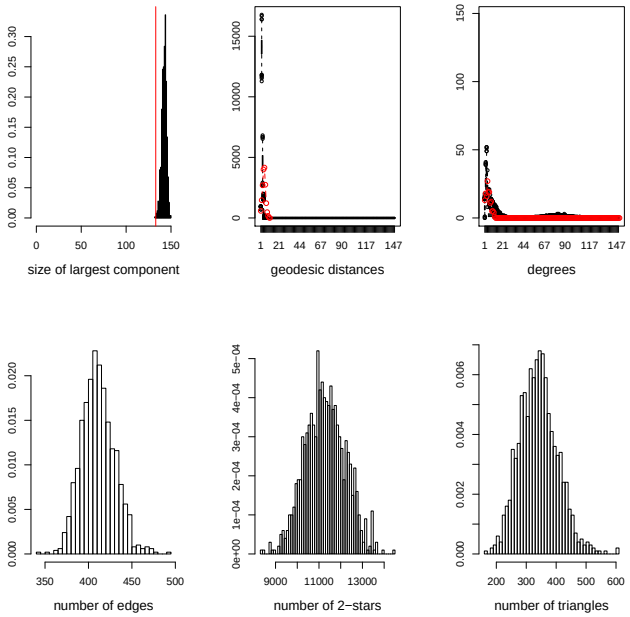


FIGURE 22. Unrestricted, local transitivity, $K = 5$, Goodness of fit

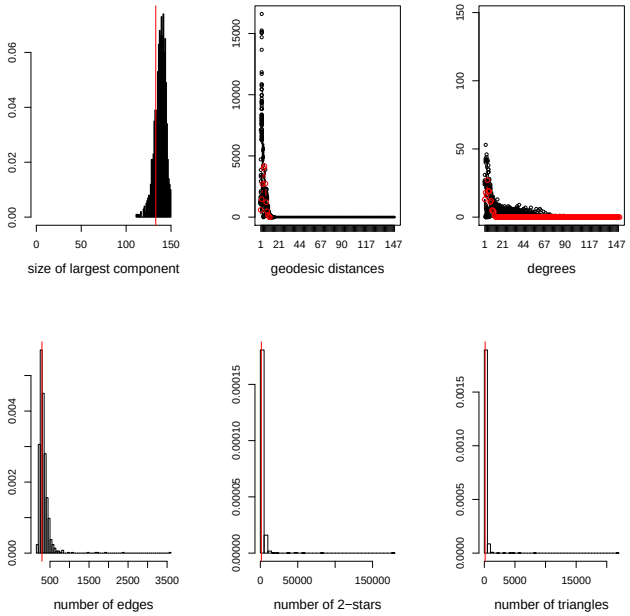


FIGURE 23. Unrestricted, local transitivity, $K = 6$, Goodness of fit

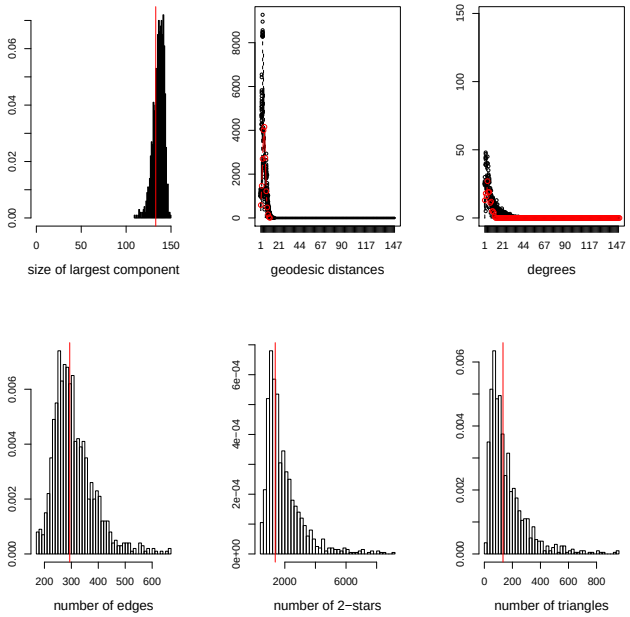


FIGURE 24. Unrestricted, local transitivity, $K = 10$, Goodness of fit

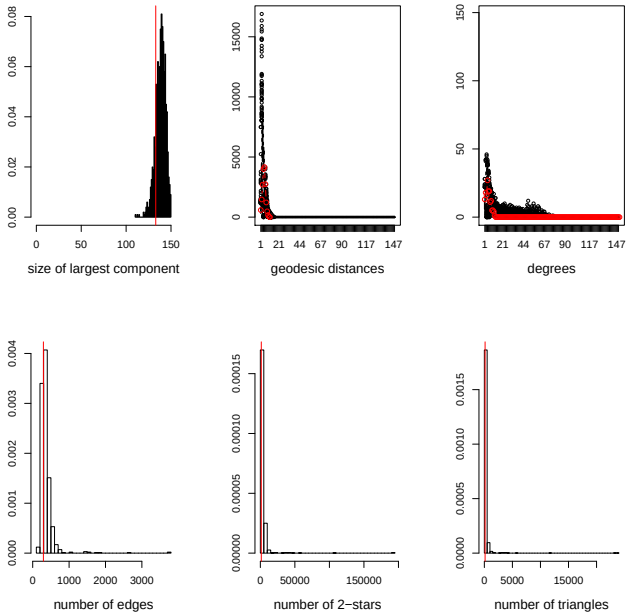


FIGURE 25. Unrestricted, global transitivity, $K = 2$, Goodness of fit

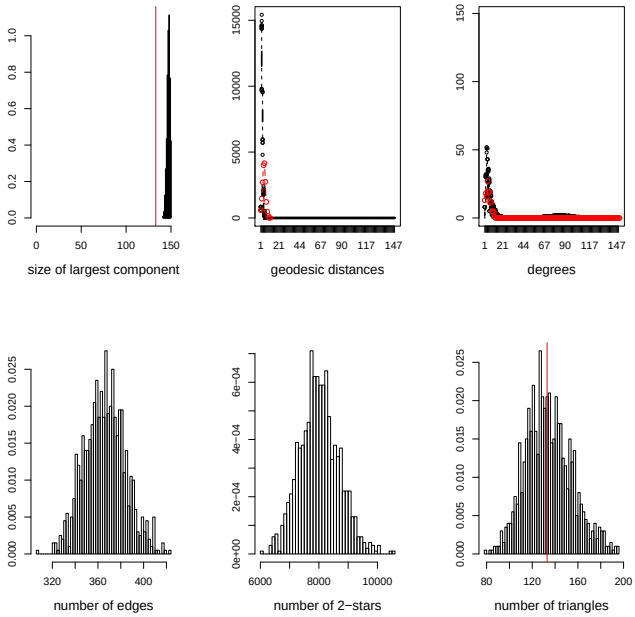


FIGURE 26. Unrestricted, global transitivity, $K = 3$, Goodness of fit

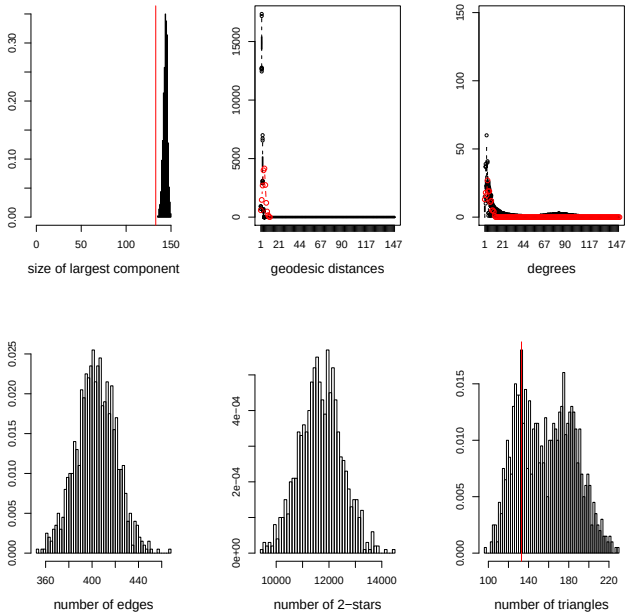


FIGURE 27. Unrestricted, global transitivity, $K = 4$, Goodness of fit

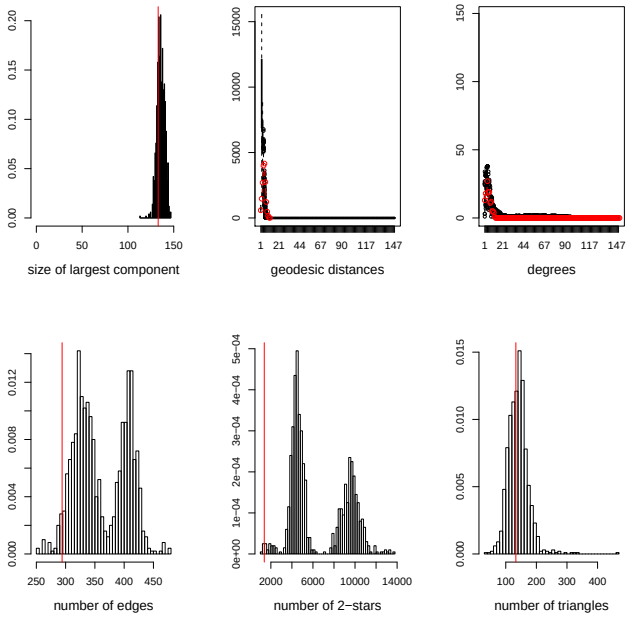


FIGURE 28. Unrestricted, global transitivity, $K = 5$, Goodness of fit

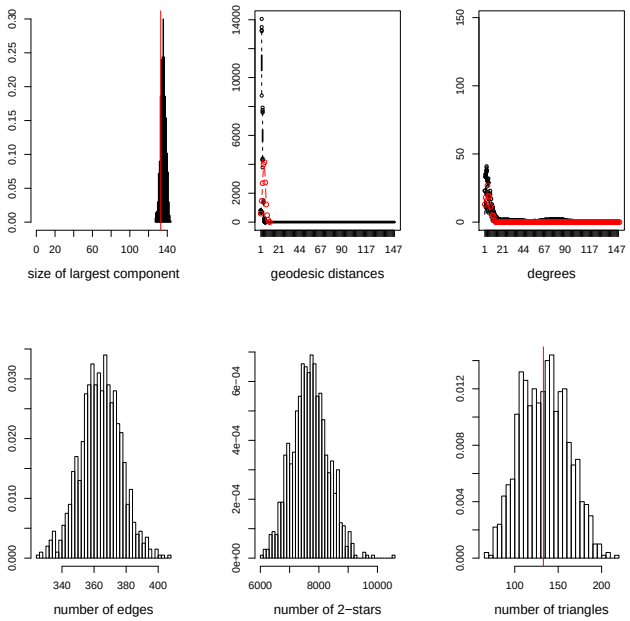


FIGURE 29. Unrestricted, global transitivity, $K = 6$, Goodness of fit

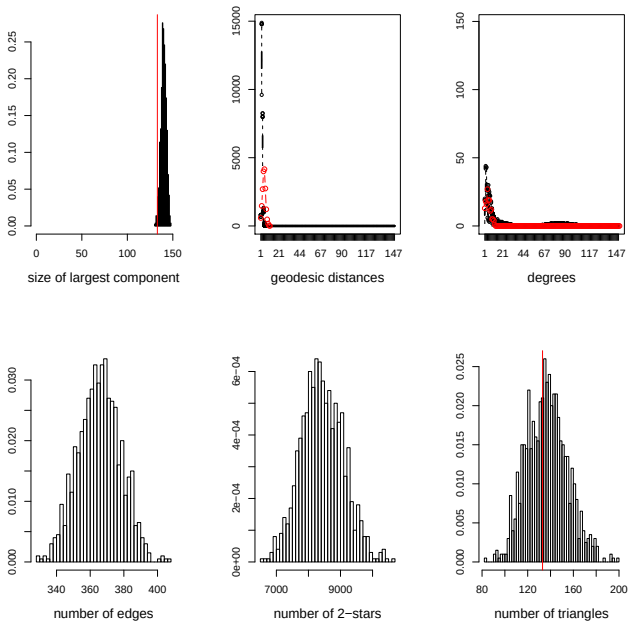


FIGURE 30. Unrestricted, global transitivity, $K = 10$, Goodness of fit

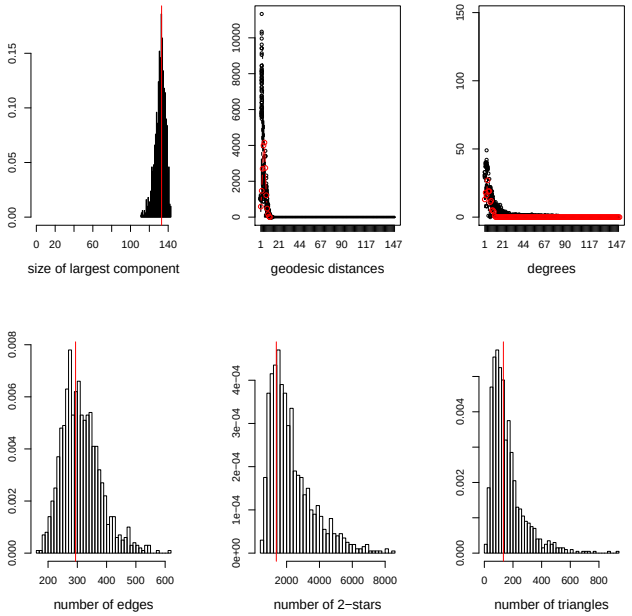


FIGURE 31. Restricted, no transitivity, $K = 2$, Goodness of fit

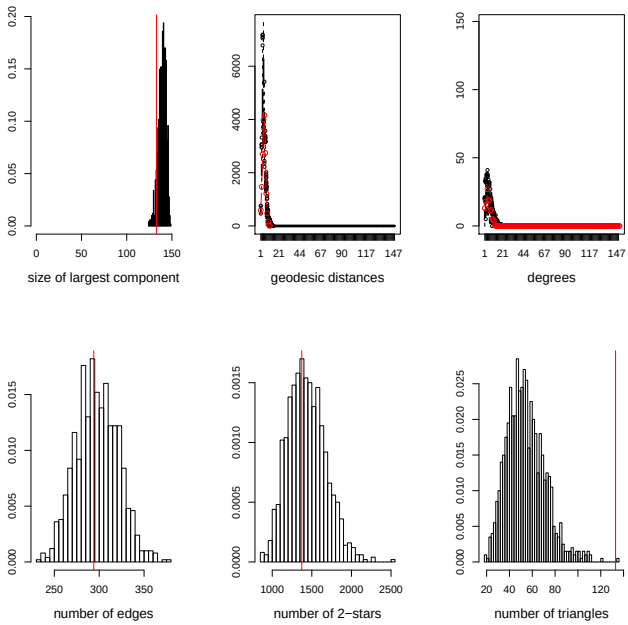


FIGURE 32. Restricted, no transitivity, $K = 3$, Goodness of fit

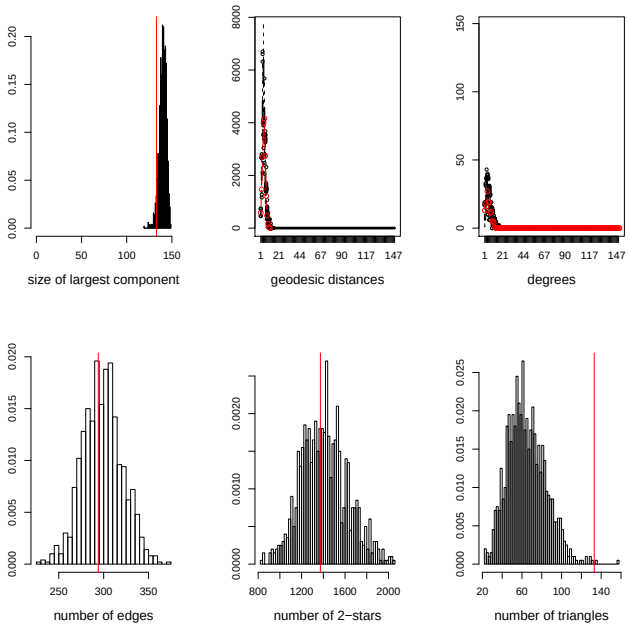


FIGURE 33. Restricted, no transitivity, $K = 4$, Goodness of fit

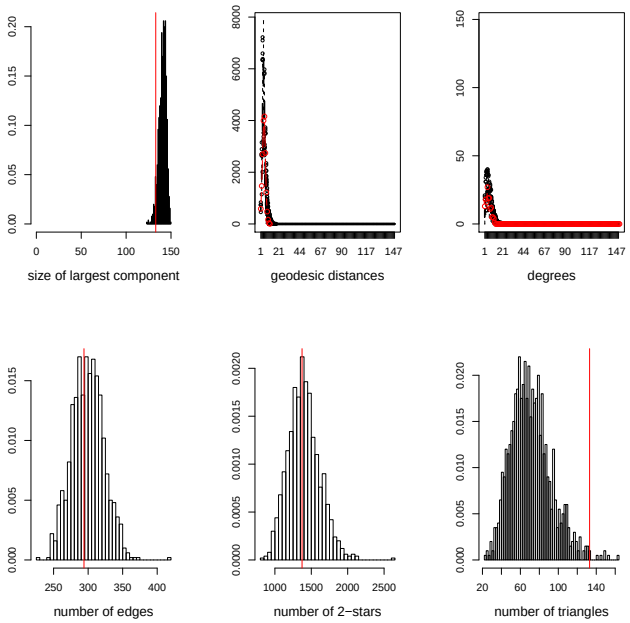


FIGURE 34. Restricted, no transitivity, $K = 5$, Goodness of fit

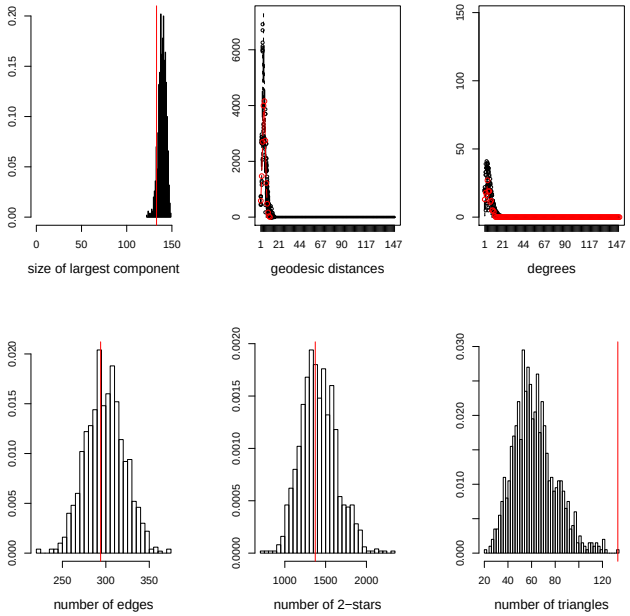


FIGURE 35. Restricted, no transitivity, $K = 6$, Goodness of fit

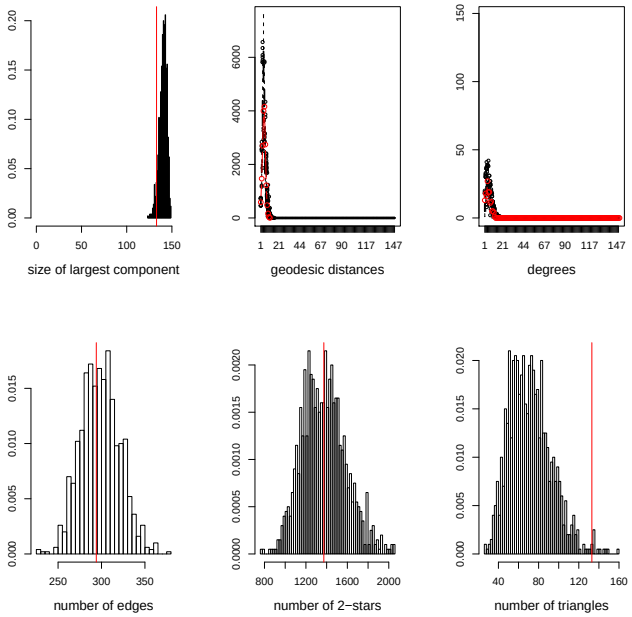


FIGURE 36. Restricted, no transitivity, $K = 10$, Goodness of fit

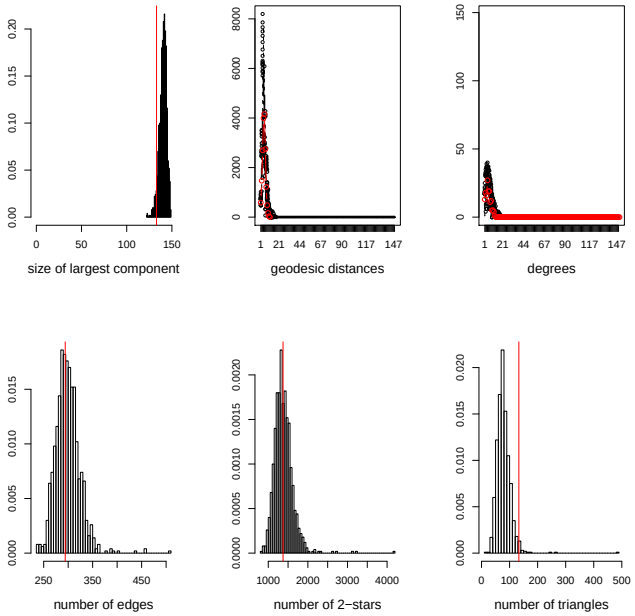


FIGURE 37. Unrestricted, no transitivity, $K = 2$, Goodness of fit

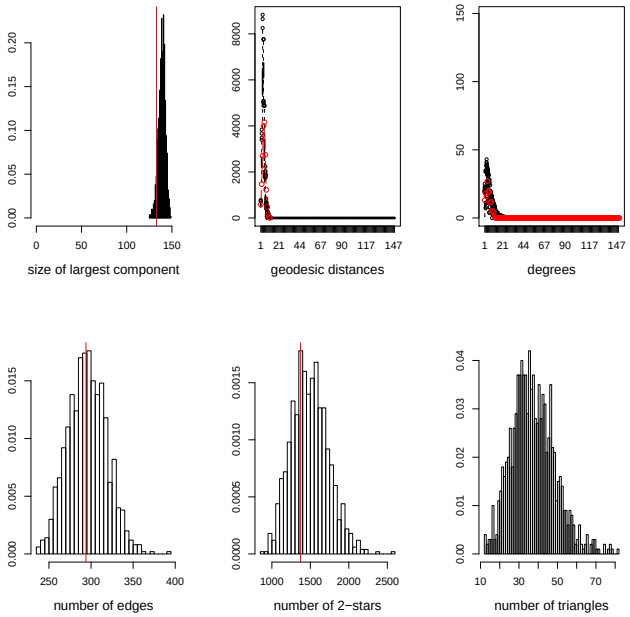


FIGURE 38. Unrestricted, no transitivity, $K = 3$, Goodness of fit

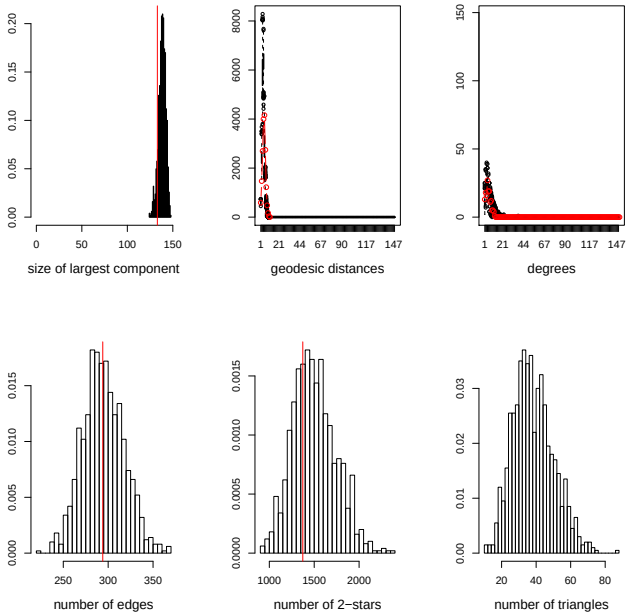


FIGURE 39. Unrestricted, no transitivity, $K = 4$, Goodness of fit

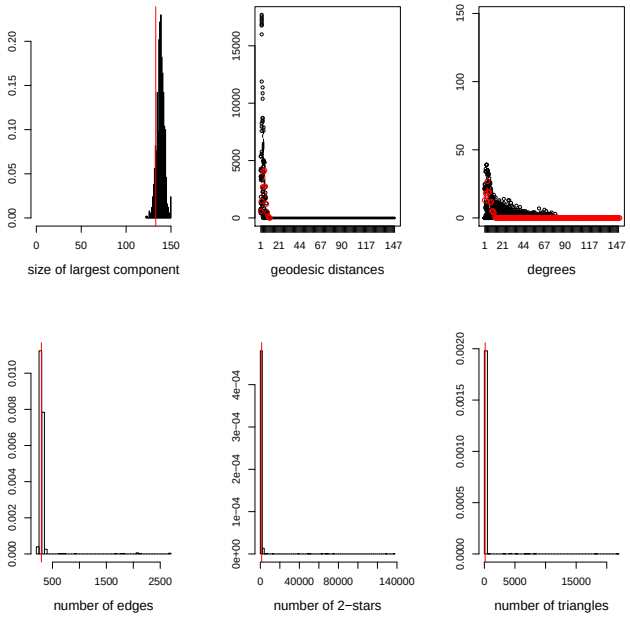


FIGURE 40. Unrestricted, no transitivity, $K = 5$, Goodness of fit

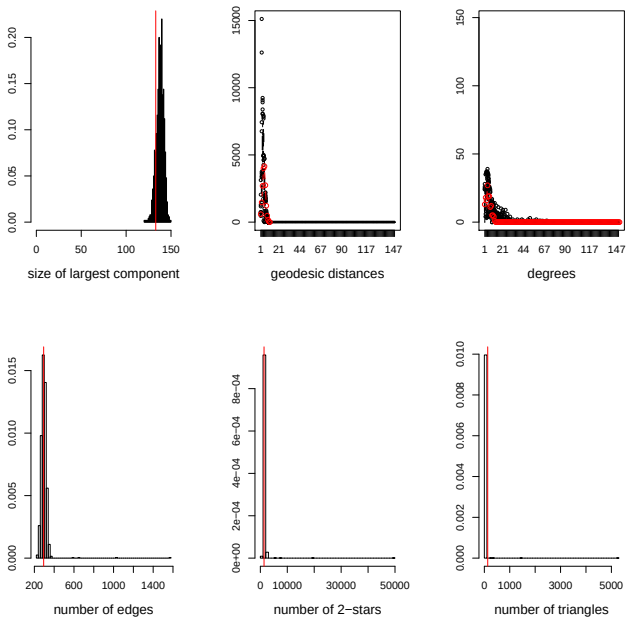


FIGURE 41. Unrestricted, no transitivity, $K = 6$, Goodness of fit

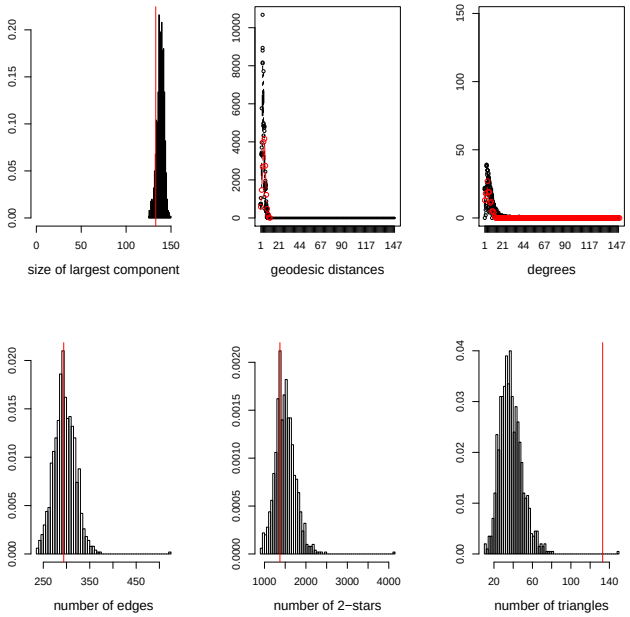
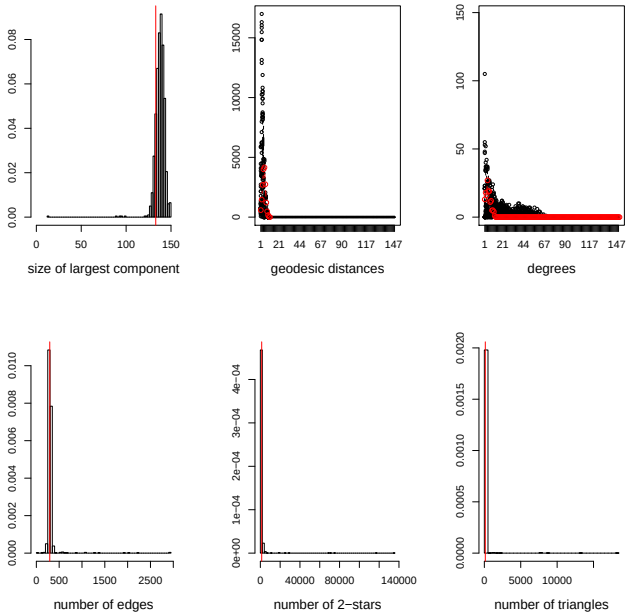


FIGURE 42. Unrestricted, no transitivity, $K = 10$, Goodness of fit



Johns Hopkins University, Carey Business School, 100 International Dr, Baltimore, MD 21202