Disaster Response on September 11, 2001 Through the Lens of Statistical Network Analysis

Michael Schweinberger[*], **Miruna Petrescu-Prahova**[†], and **Duy Quang Vu**[‡]

[*]Department of Statistics, Rice University, Houston, TX, USA

[†]Department of Health Services, University of Washington, Seattle, WA, USA

[‡]Department of Mathematics and Statistics, University of Melbourne, Melbourne, Australia

Abstract

The rescue and relief operations triggered by the September 11, 2001 attacks on the World Trade Center in New York City demanded collaboration among hundreds of organisations. To shed light on the response to the September 11, 2001 attacks and help to plan and prepare the response to future disasters, we study the inter-organisational network that emerged in response to the attacks. Studying the inter-organisational network can help to shed light on (1) whether some organisations dominated the inter-organisational network and facilitated communication and coordination of the disaster response; (2) whether the dominating organisations were supposed to coordinate disaster response or emerged as coordinators in the wake of the disaster; and (3) the degree of network redundancy and sensitivity of the inter-organisational network to disturbances following the initial disaster. We introduce a Bayesian framework which can answer the substantive questions of interest while being as simple and parsimonious as possible. The framework allows organisations to have varying propensities to collaborate, while taking covariates into account, and allows to assess whether the inter-organisational network had network redundancy—in the form of transitivity—by using a test which may be regarded as a Bayesian score test. We discuss implications in terms of disaster management.

Keywords

discrete exponential families; hierarchical models; mixture models; model-based clustering; social networks; stochastic block models

1 Introduction

Large-scale disasters can strike without warning, leaving death and destruction behind. Examples are natural disasters, such as earthquakes and tsunamis, and man-made disasters, such as industrial accidents and terrorist attacks (e.g., Topper and Carley, 1999; Butts et al., 2007; Petrescu-Prahova and Butts, 2008). When disasters strike, governmental as well as private organisations converge at the scene of the disaster to conduct rescue and relief operations (Auf der Heide, 1989). The scale of operations may overtax the resources of

[*]Corresponding author; phone number: +1 713 348 2278; fax number: +1 713 348 5476; michael.schweinberger@rice.edu.

individual organisations and compel them to collaborate (Haas and Drabek, 1973). In addition, organisations may be forced to collaborate in order to coordinate operations. As a result, inter-organisational networks of collaboration tend to emerge in response to disasters. Studying the topological structure of inter-organisational networks is important with a view to understanding the weaknesses and strengths of the response to past disasters and planning and preparing the response to future disasters.

We consider here data on the massive rescue and relief operations triggered by the September 11, 2001 attacks on the World Trade Center in New York City (Tierney and Trainor, 2004; Bevc, 2010). The network of collaborations between the $n = 717$ organisations which responded to the attacks is interesting for at least three reasons. First, an important question is whether some organisations were more involved in the disaster response than others. The presence of such organisations is considered critical to the disaster response, because in the unstructured and changing environment in which rescue and relief operations are conducted such organisations facilitate communication and coordination of the disaster response (e.g., Auf der Heide, 1989). Second, if such organisations were present, then the question is whether those organisations were supposed to coordinate disaster response or emerged as coordinators in the wake of the disaster. The possible presence of emergent coordinators has implications in terms of resource allocation to responders (e.g., Petrescu-Prahova and Butts, 2008). Third, while the presence of established and emergent coordinators can improve the disaster response, it can make the disaster response more vulnerable to disturbances which follow the initial disaster and can disrupt the disaster response (e.g., Topper and Carley, 1999). Such disturbances are not uncommon, e.g., earthquakes may be followed by after-shocks, and terrorists may attack responders on purpose by first detonating explosive devices to attract by-standers and responders and then detonating more explosive devices to harm them (see, e.g., the Bali bombing in 2002 and the Madrid bombing in 2004, Koschade, 2006). In particular, the impact of the second airplane into the South Tower of the World Trade Center on September 11, 2001 disrupted the response to the September 11, 2001 attacks. Therefore, it is important to assess how much network redundancy inter-organisational networks possess and how sensitive inter-organisational networks are to disturbances.

These questions can be framed as questions about the organisations and the topological structure of the network of organisations, first and foremost the propensities of organisations to collaborate as well as the redundancy of the network in the form of transitivity, while taking available covariates into account.

A simple approach to answering the first question—whether some organisations were more involved in the disaster response than others—is based on clustering organisations by using deterministic clustering methods in social network analysis (e.g., Borgatti et al., 2002), physics (e.g., Girvan and Newman, 2002), and computer science (e.g., Leskovec et al., 2008). However, deterministic clustering methods are not model-based and ignore the uncertainty about the clustering; neglect covariates and all other features of the network (e.g., transitivity); and lack desirable statistical properties: e.g., Bickel and Chen (2009) showed that the Girvan-Newman modularity (Girvan and Newman, 2002)—which has attracted much attention—is not consistent.

An alternative approach to answering the first question is based on modeling the sequence of degrees, corresponding to the numbers of collaborations of organisations (e.g, Erdös and Rényi, 1959; Barabàsi and Albert, 1999; Newman et al., 2001; Jones and Handcock, 2004). However, such one- and two-parameter models are not flexible models of the sequence of degrees and may not be able to account for the heterogeneity of the degrees. In addition, such models ignore covariates and all other features of networks (e.g., transitivity).

A second, alternative approach to answering the first question is based on mixture models and other latent variable models, such as stochastic block models (e.g., Nowicki and Snijders, 2001; Tallberg, 2005; Airoldi et al., 2008), random effect models (e.g., van Duijn et al., 2004; Hoff, 2005), and latent space models (e.g., Hoff et al., 2002; Schweinberger and Snijders, 2003; Handcock et al., 2007). However, some of those models make strong model assumptions about the data: e.g., latent space models assume that organisations are embedded in a latent, metric space, which is a strong model assumption and is not needed to answer the substantive questions of interest. In addition, the computing time of most of them scales with n^2. Indeed, most of the statistical software which implements these statistical models and methods is either not publicly available (e.g., Tallberg, 2005) or cannot handle networks with $n \gg 200$: e.g., neither the statistical software BLOCKS (Snijders and Nowicki, 2007) implementing the stochastic block models of Nowicki and Snijders (2001) nor the R package lda (Chang, 2011) implementing the extended stochastic block models of Airoldi et al. (2008) can handle networks with $n \gg 200$ organisations.

A third, alternative approach to answering the first question is based on exponential-family models with so-called *k*-stars as sufficient statistics (Frank and Strauss, 1986) and curved exponential-family models with so-called geometrically weighted degree terms (Hunter and Handcock, 2006; Snijders et al., 2006; Hunter, 2007). However, these exponential-family models are, at present, not well-understood and sometimes possess undesirable properties, e.g., model degeneracy, and in accordance have attracted much critisism (Strauss, 1986; Jonasson, 1999; Snijders, 2002; Handcock, 2003; Bhamidi et al., 2008; Rinaldo et al., 2009; Butts, 2011; Schweinberger, 2011; Chatterjee and Diaconis, 2013).

Last, an approach to answering the first and the third question—whether there is more network redundancy than expected—is based on clustering coefficients (e.g., Kolaczyk, 2009, pp. 96–97), though clustering coefficients neither allow to cluster organisations based on the propensity to collaborate nor allow to test whether there is more network redudancy than expected.

To overcome the shortcomings of existing approaches, we introduce a Bayesian exponential-family framework which helps to answer the substantive questions of interest and is as simple and parsimonious as possible. The framework allows us to incorporate covariates as well as variation in the propensities of organisations to collaborate. In addition, the framework allows us to test whether there is network redundancy in the form of transitivity by using a test which can be considered to be a Bayesian analogue of the score test (Rao and Poti, 1946; Rao, 1948; Bera and Bilias, 2001a,b). The advantage of using a Bayesian score test is that models with transitivity—which are hard to estimate—need not be estimated, while the question of primary interest—whether there is an excess of transitivity—can be answered. An additional advantage of the framework is that the computing time scales with n rather than n^2 and thus the framework can be applied to the inter-organisational network with $n = 717$.

The paper is structured as follows. We describe the data and the substantive questions of interest in Section 2 and introduce Bayesian models in Section 3. We discuss Bayesian model estimation and model selection in Section 4. We shed light on the substantive questions of interest by using the Bayesian framework in Section 5 and discuss implications in terms of disaster management in Section 6.

2 Data

2.1 The inter-organisational network

On September 11, 2001, the United States experienced one of the deadliest and costliest events in its history: The attacks on the World Trade Center in New York City killed 2,749 individuals and resulted in immeasurable psychological, social, and economic costs (Tierney and Trainor, 2004; Bevc, 2010).

We consider here data on the massive rescue and relief operations triggered by the September 11, 2001 attacks. A detailed description of the data can be found in Tierney and Trainor (2004) and Bevc (2010). The data were collected during the first 12 days following September 11, 2001 and represent the rescue and relief phase of operations as well as the transition into the less communication-intensive recovery phase of operations. The resulting data set includes 717 organisations and two organisational attributes: scale of operations, which includes the categories of local, state, national, and international; and type of organisation, which includes the categories of collective, governmental, non-profit, and profit. Table 1 reports the number of organisations with these two attributes. Data on the $\binom{717}{2} = 256,686$ (possible) collaborations among the 717 organisations were collected by content analysis of field documents and newspaper articles. The collaborations are binary—i.e., either absent or present—and undirected. Figure 1 shows the collaborations among the 717 organisations.

2.2 Substantive questions of interest

A large body of data is hard to comprehend and Figure 1 shows as much as it hides. Despite the size and complexity of the data set, it is evident that the inter-organisational network is dominated by a small number of organisations which are involved in a large number of collaborations. Such organisations are considered to be critical to the disaster response, because they can (a) receive information and advice from many other organisations, directly and indirectly; (b) rapidly spread information and advice to many other organisations, directly and indirectly; (c) encourage organisations which do not collaborate with each other to start collaborating; and (d) coordinate rescue and relief activities of organisations. An important question is therefore whether some organisations were more involved in the disaster response than others.

A second, interesting question is whether these dominating organisations were supposed to coordinate disaster response or emerged as coordinators in the wake of the disaster. The possible presence of emergent coordinators implies that, e.g., organisations which are supposed to respond to disasters should be prepared to collaborate with emergent coordinators, and disaster management should be prepared to allocate resources to both established and emergent coordinators (Petrescu-Prahova and Butts, 2008).

Last, but not least, it is important to assess whether the inter-organisational network possessed network redundancy which could help to absorb the impact of disturbances which follow the initial disaster and have the potential to disrupt the disaster response. An important form of network redundancy is transitivity. Transitivity is illustrated by the triangle in Figure 2, where circles represent organisations and lines represent collaborations between organisations. If, for example, one edge of the triangle were removed, then every organisation could still reach every other organisation either directly or indirectly, through the third organisation. If one of the organisations were removed, then the remaining two organisations could still reach each other directly. In general, network redundancy can help to absorb disturbances by providing alternative channels of communication through which

information and advice could spread and which could be used to coordinate the disaster response. In particular, network redundancy can help to absorb disturbances that reduce the ability of established and emergent coordinators to coordinate the disaster response. It is therefore of interest to assess whether there is network redundancy in the form of transitivity.

In Section 3, we introduce a Bayesian framework which helps to shed light on the substantive questions of interest while being as simple and parsimonious as possible.

3 Bayesian model

We consider a set of n nodes (organisations), indexed by integers $1, \ldots, n$, and random variables $Y_{ij} \in \{0, 1\}$, where $Y_{ij} = 1$ indicates that there is an edge (collaboration) between nodes (organisations) i and j and $Y_{ij} = 0$ otherwise. The observed outcome of Y_{ij} is denoted by y_{ij}. Since collaborations are undirected and self-collaborations are meaningless, we assume that edges are undirected, $y_{ij} = y_{ji}$ ($1 \le i < j \le n$), and exclude self-edges, $y_{ii} = 0$ ($1 \le i \le n$). We denote by $\mathbf{Y}$ the set of random variables Y_{ij} ($1 \le i < j \le n$) and by $\mathscr{Y}$ the sample space of $\mathbf{Y}$.

The distribution of $\mathbf{Y}$ can be parameterised in exponential-family form (Besag, 1974; Frank and Strauss, 1986):

$$P(\boldsymbol{Y}=\mathbf{y}|\boldsymbol{\eta})=\exp[\boldsymbol{\eta}^{\top}s(\mathbf{y}) - \psi(\boldsymbol{\eta})], \mathbf{y} \in \mathscr{Y}, \quad (1)$$

where η is a d-vector of natural parameters, $s(\mathbf{y})$ is a d-vector of sufficient statistics, and $\psi(\eta)$ ensures that $P(\mathbf{Y} = \mathbf{y} \mid \eta)$ sums to 1.

While exponential families of distributions of the form (1) constitute a general modelling framework and admit the specification of countless models, from simple models with one parameter (e.g., Erdös and Rényi, 1959) to the saturated model with $\exp[\binom{n}{2}\log 2] - 1$ parameters (Handcock, 2003), it is desirable to narrow down the set of possible models to a subset of models which are deemed interesting on substantive grounds and acceptable on computational and statistical grounds.

Motivated by the substantive questions of interest on the one hand and computational and statistical considerations on the other hand, we introduce models in Sections 3.1—3.3 and discuss priors in Section 3.4.

3.1 Modelling degrees

To model the propensities of organisations to collaborate, we model the sequence of degrees $d_1(\mathbf{y}), \ldots, d_n(\mathbf{y})$, where $d_i(\mathbf{y})=\sum_{j \neq i}^{n} y_{ij}$ is the degree of organisation i, i.e., the number of collaborations in which organisation i is involved. A natural model of the sequence of degrees is given by the exponential family of distributions

$$P(\boldsymbol{Y}=\mathbf{y}|\boldsymbol{\delta})=\exp\left\{\sum_{i=1}^{n}\delta_i d_i(\mathbf{y}) - \psi(\boldsymbol{\delta})\right\}, \quad (2)$$

where the sufficient statistics $d_1(\mathbf{y}), \ldots, d_n(\mathbf{y})$ are the degrees of organisations, the natural parameters $\delta_1, \ldots, \delta_n$ can be interpreted as the propensities of organisations to collaborate—which are unconstrained and may thus vary from organisation to organisation—and $\psi(\boldsymbol{\delta})$ ensures that $P(\mathbf{Y} = \mathbf{y} \mid \boldsymbol{\delta})$ sums to 1. The exponential-family form of (2) can be motivated by

the principle of maximum entropy and the attractive properties of exponential families (e.g., Barndorff-Nielsen, 1978). It is worth noting that the exponential family (2), which is designed to model the sequence of degrees of undirected networks, is the natural relative of the exponential family proposed by Holland and Leinhardt (1981) to model the sequence of in-degrees and out-degrees of directed networks. Some large-n properties of the exponential family (2) were studied by Diaconis et al. (2011).

A convenient property of the exponential family (2) is that the probability mass function can be factorised according to

$$P(\boldsymbol{Y}=\mathbf{y}|\boldsymbol{\delta})=\exp\left\{\sum_{i=1}^{n}\delta_i d_i(\mathbf{y})-\psi(\boldsymbol{\delta})\right\}=\prod_{i<j}^{n}\exp\{\lambda_{ij}(\boldsymbol{\delta})y_{ij}-\psi_{ij}(\boldsymbol{\delta})\}, \quad (3)$$

where

$$\lambda_{ij}(\boldsymbol{\delta})=\delta_i+\delta_j \quad (4)$$

and

$$\psi_{ij}(\boldsymbol{\delta})=\log[1+\exp\{\lambda_{ij}(\boldsymbol{\delta})\}]. \quad (5)$$

Model (3) arises as a natural model of the sequence of degrees as well as the assumption that the Y_{ij} are independent Bernoulli(π_{ij}) random variables with probability π_{ij} and log odds

$$\log\frac{\pi_{ij}}{1-\pi_{ij}}=\delta_i+\delta_j. \quad (6)$$

The first derivation—as a model of the degree sequence—clarifies that the exponential family is designed to model the sequence of degrees, while the second derivation—as a logistic regression model—clarifies that the log odds of the probability of a collaboration between organisations i and j is additive in the propensities of i and j to collaborate.

A second convenient property of the exponential family is that the sufficient statistics are linear in y_{ij}. In contrast, conventional exponential-family models with Markov dependence (Frank and Strauss, 1986) include sufficient statistics, called k-star statistics, of the form

$s_k(\mathbf{y})=\sum_{i=1}^{n}\sum_{j_1<\cdots<j_k}^{n}y_{ij_1}\cdots y_{ij_k}=\sum_{i=1}^{n}\binom{d_i(\mathbf{y})}{k}$, which are functions of the sequence of degrees, but are non-linear in y_{ij} and tend to suffer from model degeneracy (Strauss, 1986; Handcock, 2003; Schweinberger, 2011). The model degeneracy in turn tends to impede simulation and statistical inference (Snijders, 2002; Handcock, 2003; Rinaldo et al., 2009; Butts, 2011; Schweinberger, 2011; Chatterjee and Diaconis, 2013).

3.2 Modelling degrees and covariates

In addition to modelling the propensities of organisations to collaborate, it is desirable to take available covariates into account.

A natural extension of the exponential family (2) to include covariates is given by

$$P(\boldsymbol{Y}=\mathbf{y}|\boldsymbol{\beta},\boldsymbol{\delta})=\exp\left\{\sum_{i=1}^{n}\delta_i d_i(\mathbf{y})+\boldsymbol{\beta}^{\top}\sum_{i<j}^{n}f(\mathbf{x}_i,\mathbf{x}_j)y_{ij}-\psi(\boldsymbol{\beta},\boldsymbol{\delta})\right\}, \quad (7)$$

where $f(\mathbf{x}_i, \mathbf{x}_j)$ is a vector-valued function of covariates $\mathbf{x}_i$ and $\mathbf{x}_j$—which may be vectors, $\boldsymbol{\beta}$ is a vector of parameters, and $\psi(\boldsymbol{\beta}, \boldsymbol{\delta})$ ensures that $P(\mathbf{Y} = \mathbf{y} \mid \boldsymbol{\beta}, \boldsymbol{\delta})$ sums to 1. An example of $f(\mathbf{x}_i, \mathbf{x}_j)$ is given by an indicator of whether the values of organisations i and j on some covariate match. Other examples are discussed in Handcock et al. (2010).

The extended model (7) shares with (2) the convenient property that the probability mass function factorises according to

$$P(\mathbf{Y}=\mathbf{y}|\boldsymbol{\beta},\boldsymbol{\delta})=\exp\left\{\sum_{i=1}^{n}\delta_i d_i(\mathbf{y})+\boldsymbol{\beta}^{\top}\sum_{i<j}^{n}f(\mathbf{x}_i,\mathbf{x}_j)y_{ij}-\psi(\boldsymbol{\beta},\boldsymbol{\delta})\right\}=\prod_{i<j}^{n}\exp\{\lambda_{ij}(\boldsymbol{\beta},\boldsymbol{\delta})y_{ij}-\psi_{ij}(\boldsymbol{\beta},\boldsymbol{\delta})\}, \quad (8)$$

where

$$\lambda_{ij}(\boldsymbol{\beta},\boldsymbol{\delta})=\delta_i+\delta_j+\boldsymbol{\beta}^{\top}f(\mathbf{x}_i,\mathbf{x}_j) \quad (9)$$

and

$$\psi_{ij}(\boldsymbol{\beta},\boldsymbol{\delta})=\log[1+\exp\{\lambda_{ij}(\boldsymbol{\beta},\boldsymbol{\delta})\}]. \quad (10)$$

3.3 Modelling degrees, covariates, and transitivity

To assess whether there is network redundancy in the form of transitivity, we extend the exponential family (7) to include transitivity:

$$P(\mathbf{Y}=\mathbf{y}|\boldsymbol{\beta},\gamma,\boldsymbol{\delta})=\exp\left\{\sum_{i=1}^{n}\delta_i d_i(\mathbf{y})+\boldsymbol{\beta}^{\top}\sum_{i<j}^{n}f(\mathbf{x}_i,\mathbf{x}_j)y_{ij}+\gamma g(\mathbf{y})-\psi(\boldsymbol{\beta},\gamma,\boldsymbol{\delta})\right\}, \quad (11)$$

where the sufficient statistic $g(\mathbf{y})=\sum_{i<j<k}^{n}y_{ij}y_{jk}y_{ik}$ is the number of triangles, the natural parameter γ is the weight of the number of triangles $g(\mathbf{y})$, and $\psi(\boldsymbol{\beta}, \gamma, \boldsymbol{\delta})$ ensures that $P(\mathbf{Y} = \mathbf{y} \mid \boldsymbol{\beta}, \gamma, \boldsymbol{\delta})$ sums to 1. A drawback of the exponential family (11) is that its probability mass function cannot be factorised. As a result, the likelihood function is intractable and the posterior distribution is doubly intractable, requiring time-consuming auxiliary-variable Markov chain Monte Carlo methods (e.g., Koskinen et al., 2010; Caimo and Friel, 2011).

We are here less interested in estimating model (11)—which would be time-consuming—than in assessing the evidence against the null hypothesis

$$H_0:\gamma=0. \quad (12)$$

In a frequentist framework, a test—which would not require to estimate model (11)—could be based on the vector-valued score function

$$\nabla_{\eta}\log P(\mathbf{Y}=\mathbf{y}|\eta)\Big|_{\hat{\eta}_0}=s(\mathbf{y})-E_{\hat{\eta}_0}[s(\mathbf{Y})], \quad (13)$$

where η is the vector of all parameters, $s(\mathbf{y})$ is the vector of all sufficient statistics, $P(\mathbf{Y} = \mathbf{y} \mid \eta)$ is the likelihood function corresponding to (11), and the gradient of the log-likelihood function $\nabla_{\eta} \log P(\mathbf{Y} = \mathbf{y} \mid \eta)$ is evaluated at the restricted maximum likelihood estimate of η under $H_0 : \gamma = 0$, which is denoted by $\hat{\eta}_0$. By definition of the restricted maximum likelihood estimate, the elements of the vector $s(\mathbf{y}) - E_{\hat{\eta}_0}[s(\mathbf{Y})]$ vanish, with the exception of

$$\left.\frac{\partial}{\partial_\gamma}\log P(\boldsymbol{Y}=\mathbf{y}|\boldsymbol{\eta})\right|_{\hat{\eta}_0}=g(\mathbf{y}) - E_{\hat{\eta}_0}[g(\boldsymbol{Y})]. \quad (14)$$

Therefore, a test of $H_0 : \gamma = 0$ could be based on test statistics of the form

$$t(\mathbf{y}) \propto g(\mathbf{y}) - E_{\hat{\eta}_0}[g(\boldsymbol{Y})]. \quad (15)$$

Test statistics based on the score function are natural, because estimators and tests based on the score function possess attractive properties (e.g., Bera and Bilias, 2001a,b; Lehmann and Romano, 2005). The goodness-of-fit test of Pearson (1900), the score tests of Rao and Poti (1946) and Rao (1948), and countless other tests are based on the score function (e.g., Bera and Bilias, 2001a,b). In canonical exponential families, the score function is given by the difference between observed and expected values of the sufficient statistics—see, e.g., (15) —and therefore score-based tests can be interpreted as goodness-of-fit tests. The classic example is the goodness-of-fit test of Pearson (1900).

In a Bayesian framework, the evidence against $H_0 : \gamma = 0$ can be assessed by the posterior predictive p-value (Meng, 1994) given by

$$p_t=\Pr(t(\boldsymbol{Y}^\star) \geq t(\mathbf{y})|\boldsymbol{Y}=\mathbf{y}, H_0), \quad (16)$$

where $\boldsymbol{Y}\star$ is a posterior prediction of the network and $t(\boldsymbol{Y}\star)$ is a test statistic. A Bayesian analogue of (15) is given by the score-based test statistic

$$t(\mathbf{y}^\star)=g(\mathbf{y}^\star) - E[g(\boldsymbol{Y}^\star)|\boldsymbol{Y}=\mathbf{y}, H_0], \quad (17)$$

where

$$E[g(\boldsymbol{Y}^\star)|\boldsymbol{Y}=\mathbf{y}, H_0]=\sum_{\mathbf{y}^\star}g(\mathbf{y}^\star)P(\boldsymbol{Y}^\star=\mathbf{y}^\star|\boldsymbol{Y}=\mathbf{y}, H_0) \quad (18)$$

is the posterior predictive expectation of $g(\boldsymbol{Y}\star)$ under $H_0 : \gamma = 0$ and

$$P(\boldsymbol{Y}^\star=\mathbf{y}^\star|\boldsymbol{Y}=\mathbf{y}, H_0)=\int_{\eta_0} P(\boldsymbol{Y}^\star=\mathbf{y}^\star|\boldsymbol{\eta}_0, H_0)p(\boldsymbol{\eta}_0|\boldsymbol{Y}=\mathbf{y}, H_0)\mathrm{d}\boldsymbol{\eta}_0 \quad (19)$$

is the posterior predictive distribution of $\boldsymbol{Y}\star$ under $H_0 : \gamma = 0$. In contrast to the frequentist approach, the Bayesian approach averages over the unspecified parameters $\boldsymbol{\beta}$ and $\delta_1, \ldots, \delta_n$ and thus takes the uncertainty about the unspecified parameters into account. In addition, the Bayesian approach does not require approximations based on large-n asymptotics. The most important advantage of using posterior predictive p-values based on the score-based test statistic (17) is that the evidence against $H_0 : \gamma = 0$ can be assessed by generating posterior draws from model (11) under $H_0 : \gamma = 0$, which possesses the convenient factorisation property (8)—in contrast to model (11) with $\gamma \neq 0$. In other words, the advantage is that model (11)—which is hard to estimate—needs not be estimated, while the question of primary interest—whether there is more transitivity than expected—can be answered.

To explore the finite-sample behavior of the Bayesian score test, we conducted a simulation study with 100 nodes and 1,000 simulated networks. The networks were simulated under $\gamma = 0$ by using model M4 described in Section 5, using the posterior medians of the degree parameters reported in Section 5, and under $\gamma = 1$ by using model M4 with transitivity within blocks but without transitivity between blocks. The simulation results with $\gamma = 0$ and 1 show that the Bayesian score test with nominal significance level .05 rejects $H_0 : \gamma = 0$ in

0% and 98.8% of the cases, respectively. It is thus evident that the Bayesian score test is conservative, but can detect local transitivity.

Last, we note that an alternative Bayesian score test was explored by Kleibergen et al. (2000), though the score test calculates the expectation of the score statistic under the posterior distribution rather than the posterior predictive distribution and is tailored to continuous data without network structure.

3.4 Prior

We assume that the natural parameters $\delta_1, \ldots, \delta_n$ of the exponential families (2), (7), and (11) are governed by truncated Dirichlet process priors, which approximate Dirichlet process priors and have computational advantages relative to Dirichlet process priors (Ishwaran and James, 2001).

We assume that the set of organisations is partitioned into subsets, called blocks. The maximum number of blocks K is assumed to be known and can be selected by the Bayesian model selection method described in Section 4.2. We assume that organisations which are members of the same block have the same propensity to collaborate:

$$\delta_i = \boldsymbol{\theta}^\top \boldsymbol{Z}_i, i=1, \ldots, n, \quad (20)$$

where the elements θ_k of the vector $\boldsymbol{\theta}$ are block-dependent propensities to collaborate,

$$\theta_k | \mu, \sigma^2 \overset{\text{iid}}{\sim} N(\mu, \sigma^2), k=1, \ldots, K, \quad (21)$$

and the vectors $\boldsymbol{Z}_i$ are organisation-dependent indicators of block membership,

$$\boldsymbol{Z}_i | \omega_1, \ldots, \omega_K \overset{\text{iid}}{\sim} \text{Multinomial}(1; \omega_1, \ldots, \omega_K), i=1, \ldots, n. \quad (22)$$

The multinomial parameters $\omega_1, \ldots, \omega_K$ are constructed by truncated stick-breaking (Ishwaran and James, 2001):

$$\omega_1 = V_1 \ \omega_k = V_k \prod_{j=1}^{k-1}(1 - V_j), k=2, \ldots, K-1 \ \omega_K = 1 - \sum_{k=1}^{K-1} \omega_k, \quad (23)$$

where

$$V_k | \alpha \overset{\text{iid}}{\sim} \text{Beta}(1, \alpha), k=1, \ldots, K-1. \quad (24)$$

Truncated Dirichlet process priors approximate Dirichlet process priors and imply that the multinomial parameters $\omega_1, \ldots, \omega_K$ are governed by a generalised Dirichlet distribution (Ishwaran and James, 2001).

Since the posterior distribution may be sensitive to the choice of the parameters α, μ, and σ^2 of the truncated Dirichlet process prior and the parameters cannot be specified with confidence, it makes sense to express the uncertainty about the parameters by specifying hyper-priors as follows:

$$\alpha | A_1, B_1 \sim \text{Gamma}(A_1, B_1) \ \mu | M, S^2 \sim N(M, S^2) \ \sigma^{-2} | A_2, B_2 \sim \text{Gamma}(A_2, B_2), \quad (25)$$

where $A_1, B_1, M, S^2, A_2, B_2$ are specified constants.

The covariate parameters $\boldsymbol{\beta}$ can be assigned Gaussian priors, while the transitivity parameter γ is assigned a degenerate prior placing all of its probability mass at $\gamma = 0$.

4 Bayesian inference

We discuss the estimation of the Bayesian model given K blocks in Section 4.1 and the selection of the number of blocks K in Section 4.2.

4.1 Bayesian model estimation

We approximate the posterior distribution of the Bayesian model given K blocks by combining the following Markov chain Monte Carlo steps by means of cycling or mixing (Tierney, 1994; Liu, 2008). We assume throughout that the null hypothesis $H_0 : \gamma = 0$ holds. Bayesian estimation of the resulting dyad-independence model is straightforward, in contrast to Bayesian estimation of dyad-dependence models, which would necessitate time-consuming auxiliary-variable methods (Caimo and Friel, 2011).

Scaling parameter α—If the hyper-prior distribution of scaling parameter α is $\mathrm{Gamma}(A_1, B_1)$, we can sample α from its full conditional distribution:

$$\alpha | A_1, B_1, \omega_1, \ldots, \omega_K \sim \mathrm{Gamma}(A_1 + K - 1, B_1 - \log \omega_K). \quad (26)$$

Mean parameter μ—If the hyper-prior distribution of mean parameter μ is $N(M, S^2)$, we can sample μ from its full conditional distribution:

$$\mu | M, S^2, \sigma^2, \theta_1, \ldots, \theta_K \sim N \left(\frac{S^{-2}M + \sigma^{-2} \sum_{k=1}^{K} \theta_k}{S^{-2} + K\sigma^{-2}}, \frac{1}{S^{-2} + K\sigma^{-2}} \right). \quad (27)$$

Precision parameter σ⁻²—If the hyper-prior distribution of precision parameter σ^{-2} is given by $\mathrm{Gamma}(A_2, B_2)$, we can sample σ^{-2} from its full conditional distribution:

$$\sigma^{-2} | A_2, B_2, \mu, \theta_1, \ldots, \theta_K \sim \mathrm{Gamma} \left(A_2 + K/2, B_2 + \sum_{k=1}^{K} (\theta_k - \mu)^2 / 2 \right). \quad (28)$$

Multinomial parameters ω₁, …, ω_K—We sample $\omega_1, \ldots, \omega_K$ from the full conditional distribution by sampling

$$V_k^\star | \alpha, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n \overset{\mathrm{ind}}{\sim} \mathrm{Beta} \left(1 + n_k, \alpha + \sum_{j=k+1}^{K} n_j \right), k = 1, \ldots, K - 1 \quad (29)$$

and setting

$$\omega_1 = V_1^\star \quad \omega_k = V_k^\star \prod_{j=1}^{k-1} (1 - V_j^\star), k = 2, \ldots, K - 1 \quad \omega_K = 1 - \sum_{k=1}^{K-1} \omega_k, \quad (30)$$

where n_k denotes the number of organisations in block k.

Indicators Z₁, …, Z_n—We sample indicator $\boldsymbol{Z}_i$ from its full conditional distribution:

$$\boldsymbol{Z}_i|\{\boldsymbol{Z}_j\}_{j\neq i}^n, \omega_1,\dots,\omega_K, \boldsymbol{\beta}, \theta_1,\dots,\theta_K, \mathbf{y} \sim \text{Multinomial}(1;\omega_{i,1},\dots,\omega_{i,K}), \quad (31)$$

where

$$\omega_{i,k}=\frac{\omega_k\prod\limits_{\substack{i\neq j}}^{n}P(Y_{ij}=y_{ij}|\boldsymbol{\beta},\delta_i=\theta_k,\{\delta_h\}_{h\neq i}^n)}{\sum\limits_{l=1}^{K}\left\{\omega_l\prod\limits_{\substack{i\neq j}}^{n}P(Y_{ij}=y_{ij}|\boldsymbol{\beta},\delta_i=\theta_l,\{\delta_h\}_{h\neq i}^n)\right\}}. \quad (32)$$

Parameters $\theta_1, \dots, \theta_K$ and β—The full conditional distribution of $\theta_1, \dots, \theta_K$ and β is intractable. We update $\theta_1, \dots, \theta_K$ and β by Metropolis-Hastings steps, where proposals are generated from random-walk, independence, or autoregressive proposal distributions (Tierney, 1994).

The so-called label-switching problem, arising from the invariance of the likelihood function to the labelling of the blocks and resulting in symmetric and multi-modal posterior distributions and Markov chain Monte Carlo samples with multiple labellings, can be solved by using the Bayesian decision-theoretic approach of Stephens (2000).

4.2 Bayesian model selection

We are interested in selecting the number of blocks K of the Bayesian model. Markov chain Monte Carlo-based approaches to Bayesian model selection (e.g., Richardson and Green, 1997; Chib and Jeliazkov, 2001) are time-consuming. Computing time is an important concern, because the number of units of analysis—i.e., $\binom{717}{2}=256,686$ possible collaborations—is large and estimating models can take between 1 and 3 weeks. We propose a variational approach to Bayesian model selection (Blei and Jordan, 2005), which tends to be much faster than Markov chain Monte Carlo-based approaches (e.g., McGrory and Titterington, 2006). A variational approach to Bayesian model estimation of latent space models was explored by Salter-Townshend and Murphy (2013), though Bayesian model selection was based on BIC approximations of the marginal likelihood rather than variational approximations of the marginal likelihood.

Variational methods can be motivated by observing that the marginal likelihood, the most important obstacle to Bayesian model selection, can be represented by

$$P(\boldsymbol{Y}=\mathbf{y})=\int \frac{P(\boldsymbol{Y}=\mathbf{y}|\boldsymbol{\eta})p(\boldsymbol{\eta})}{q(\boldsymbol{\eta}|\boldsymbol{\vartheta})}q(\boldsymbol{\eta}|\boldsymbol{\vartheta})\mathrm{d}\boldsymbol{\eta}, \quad (33)$$

where η denotes the vector of all parameters, $P(\mathbf{Y} = \mathbf{y} \mid \eta)$ denotes the likelihood function, $p(\eta)$ denotes the prior distribution, and $q(\eta \mid \vartheta)$ is a member of a family of auxiliary distributions (to be specified). The idea of variational methods is to translate the intractable integration problem (33) into a tractable optimisation problem by lower bounding $\log P(\mathbf{Y} = \mathbf{y})$ and making the lower bound on $\log P(\mathbf{Y} = \mathbf{y})$ as tight as possible. A lower bound on $\log P(\mathbf{Y} = \mathbf{y})$ can be obtained by Jensen's inequality:

$$\log P(\boldsymbol{Y}=\mathbf{y}) \geq \int \left\{\log \frac{P(\boldsymbol{Y}=\mathbf{y}|\boldsymbol{\eta})p(\boldsymbol{\eta})}{q(\boldsymbol{\eta}|\boldsymbol{\vartheta})}\right\}q(\boldsymbol{\eta}|\boldsymbol{\vartheta})\mathrm{d}\boldsymbol{\eta}. \quad (34)$$

The best lower bound on log $P(\mathbf{Y} = \mathbf{y})$ is obtained by maximising the lower bound on log $P(\mathbf{Y} = \mathbf{y})$ with respect to $\boldsymbol{\vartheta}$:

$$\hat{\boldsymbol{\vartheta}} = \arg \max_{\boldsymbol{\vartheta}} \left[\int \left\{ \log \frac{P(\mathbf{Y}=\mathbf{y}|\boldsymbol{\eta})p(\boldsymbol{\eta})}{q(\boldsymbol{\eta}|\boldsymbol{\vartheta})} \right\} q(\boldsymbol{\eta}|\boldsymbol{\vartheta}) \mathrm{d}\boldsymbol{\eta} \right]. \quad (35)$$

The geometry of the maximisation problem is simple: The difference between the left-hand side of (34) and the right-hand side of (34) is equal to the Kullback-Leibler divergence from the auxiliary distribution $q(\boldsymbol{\eta} \mid \boldsymbol{\vartheta})$ indexed by $\boldsymbol{\vartheta}$ to the posterior distribution $p(\boldsymbol{\eta} \mid \mathbf{Y} = \mathbf{y})$:

$$KL\{q(\boldsymbol{\eta}|\boldsymbol{\vartheta}); p(\boldsymbol{\eta}|\mathbf{Y}=\mathbf{y})\} = \int \left\{ \log \frac{q(\boldsymbol{\eta}|\boldsymbol{\vartheta})}{p(\boldsymbol{\eta}|\mathbf{Y}=\mathbf{y})} \right\} q(\boldsymbol{\eta}|\boldsymbol{\vartheta}) \mathrm{d}\boldsymbol{\eta}. \quad (36)$$

Therefore, maximising the lower bound on log $P(\mathbf{Y} = \mathbf{y})$ with respect to $\boldsymbol{\vartheta}$ amounts to minimising the Kullback-Leibler divergence from the auxiliary distribution $q(\boldsymbol{\eta} \mid \boldsymbol{\vartheta})$ indexed by $\boldsymbol{\vartheta}$ to the posterior distribution $p(\boldsymbol{\eta} \mid \mathbf{Y} = \mathbf{y})$. It is worth noting that variational approximations of log $P(\mathbf{Y} = \mathbf{y})$ penalise model complexity, because log $P(\mathbf{Y} = \mathbf{y})$ penalises model complexity (e.g., MacKay, 2003, Chapter 28).

In practice, it is important that the optimisation problem (35) is tractable, which depends first and foremost on the choice of the family of auxiliary distributions. To make the optimisation problem (35) tractable, we choose a family of fully factorised auxiliary distributions of the form

$$q(\boldsymbol{\eta}|\boldsymbol{\vartheta}) = q(\alpha|\vartheta_{\alpha,1}, \vartheta_{\alpha,2}) q(\mu|\vartheta_{\mu,1}, \vartheta_{\mu,2}) q(\sigma^{-2}|\vartheta_{\sigma^{-2},1}, \vartheta_{\sigma^{-2},2})$$
$$\times \left\{ \prod_{k=1}^{K-1} q(V_k|\vartheta_{V_k,1}, \vartheta_{V_k,2}) \right\} \left\{ \prod_{i=1}^{n} q(\mathbf{Z}_i|\boldsymbol{\vartheta}_{\mathbf{Z}_i}) \right\} \left\{ \prod_{k=1}^{K} q(\theta_k|\vartheta_{\theta_k,1}, \vartheta_{\theta_k,2}) \right\}, \quad (37)$$

where the marginal auxiliary distributions $q(\alpha|\vartheta_{\alpha,1}, \vartheta_{\alpha,2})$ and $q(\sigma^{-2}|\vartheta_{\sigma^{-2},1}, \vartheta_{\sigma^{-2},2})$ are Gamma distributions, $q(\mu|\vartheta_{\mu,1}, \vartheta_{\mu,2})$ and $q(\theta_k|\vartheta_{\theta_k,1}, \vartheta_{\theta_k,2})$ are Gaussian distributions, $q(V_k|\vartheta_{V_k,1}, \vartheta_{V_k,2})$ are Beta distributions, and $q(\mathbf{Z}_i|\boldsymbol{\vartheta}_{\mathbf{Z}_i})$ are multinomial distributions. The implementation of the variational approach is discussed in more detail in the supplement.

If the posterior correlations are weak, the fully factorised auxiliary distribution $q(\boldsymbol{\eta} \mid \boldsymbol{\vartheta})$ under the maximiser $\hat{\boldsymbol{\vartheta}}$ can be expected to be close to the posterior distribution $p(\boldsymbol{\eta} \mid \mathbf{Y} = \mathbf{y})$ in terms of Kullback-Leibler divergence and the lower bound on log $P(\mathbf{Y} = \mathbf{y})$ can be expected to be tight. In practice, as long as the posterior correlations are not too strong, one would expect that large changes of log $P(\mathbf{Y} = \mathbf{y})$ due to increases in the number of blocks K are reflected by large changes of the variational approximations of log $P(\mathbf{Y} = \mathbf{y})$. In terms of computing time, in the application in Section 5, the variational approximations of log $P(\mathbf{Y} = \mathbf{y})$ consumed .5 hours ($K = 2$ blocks) to 15 hours ($K = 10$ blocks), based on 1,000 runs to address the issue of local maxima, which is much less than Markov chain Monte Carlo-based approaches would consume (i.e., weeks).

5 Application

We exploit the Bayesian framework introduced in Sections 3—5 to answer the substantive questions of interest described in Section 2.2. We start by assessing whether some of the 717 organisations that responded to the September 11, 2001 attacks were more involved in collaboration than others. We compare competing models capturing the heterogeneity of organisations with respect to collaboration in Section 5.2. We shed light on which organisations were primary coordinators and whether the primary coordinators were

established or emergent coordinators in Section 5.3. We assess whether the inter-organisational network possessed network redundancy in the form of transitivity in Section 5.4. Last, we study the vulnerability of the inter-organisational network to disturbances in Section 5.5.

5.1 Assessing heterogeneity

To assess whether some of the 717 organisations that responded to the September 11, 2001 attacks were more involved in collaboration than others, we compare model (2) with one block to model (2) with more than one block. We use the truncated Dirichlet process prior with hyper-prior $\alpha \sim$ Gamma(1, 1), $\mu \sim N$(0, 25), and $\sigma^{-2} \sim$ Gamma(5, 5). We compare the model with K blocks in terms of the marginal likelihood $P(\mathbf{Y} = \mathbf{y})$. Figure 3 shows the approximations of log $P(\mathbf{Y} = \mathbf{y})$ under the model with 1, …, 10 blocks. The difference in log $P(\mathbf{Y} = \mathbf{y})$ between the model with K and $K + 1$ blocks can be interpreted as log Bayes factor (Kass and Raftery, 1995). There are dramatic increases in log $P(\mathbf{Y} = \mathbf{y})$ by going from 1 to 4 blocks, small changes by going from 4 to 9 blocks, and a large decrease by going from 9 to 10 blocks. Therefore, according to Kass and Raftery (1995), there is decisive evidence against the model with 1 block. In other words, some organisations were more involved in collaboration than others.

To select the number of blocks, we observe that, according to Figure 3, model (2) with 4 or more blocks is preferable to model (2) with 1, 2, or 3 blocks. In view of model parsimony, we select the model with 4 blocks.

5.2 Accounting for heterogeneity

There are multiple models which can capture the heterogeneity of organisations with respect to collaboration, take the available covariates into account, and can be estimated in reasonable time. We compare the Bayesian framework introduced above to three competing exponential-family models of the form (1) in terms of goodness-of-fit. The three competing exponential-family models include sufficient statistics which are functions of the sequence of degrees $d_1(\mathbf{y})$, …, $d_n(\mathbf{y})$:

M1 (Bevc, 2010): The exponential family with the number of edges $\sum_{i<j}^{n} y_{ij} = \sum_{i=1}^{n} d_i(\mathbf{y})/2$ as sufficient statistic, which is equivalent to assuming that the edges are independent Bernoulli random variables.

M2 : The exponential family with the number of edges $\sum_{i<j}^{n} y_{ij}$ and the number of 2-stars $\sum_{i=1}^{n} \sum_{j_1 < j_2}^{n} y_{ij_1} y_{ij_2} = \sum_{i=1}^{n} \binom{d_i(\mathbf{y})}{2}$ as sufficient statistics, which is a special case of the exponential family with Markov dependence of Frank and Strauss (1986).

M3 : The curved exponential family with the number of edges $\sum_{i<j}^{n} y_{ij}$ as sufficient statistic and geometrically weighted degree terms (Hunter and Handcock, 2006; Hunter, 2007), corresponding to the $n - 1$ sufficient statistics $d_k(\mathbf{y}) = \sum_{i=1}^{n} 1\{d_i(\mathbf{y}) = k\}$, the number of organisations with degree k, and the $n-1$ natural parameters $\delta_k(\boldsymbol{\theta}) = \theta_1 \exp(\theta_2) [1 - \{1 - \exp(-\theta_2)\}^k]$ ($k = 1, …, n - 1$), which are non-linear functions of two parameters θ_1 and θ_2.

We compare these three competing exponential-family models with two competing latent variable models:

M4 : Model (2) with 4 blocks.

M5 (Krivitsky et al., 2009): The organisations have unobserved positions in a 2-dimensional Euclidean space, where the positions of organisations are drawn from a mixture of 4 normal distributions, and the propensities of organisations to collaborate vary across organisations by using the sociality random effects as described by Krivitsky et al. (2009).

To take the covariates scale of operations and type of organisation into account, we follow Bevc (2010) and add the two covariates in the form of node-mix covariate terms to each of the models, with the exception of model M5, because the R package latentnet (Krivitsky and Handcock, 2012) implementing model M5 does not support node-mix covariate terms. That is, for each of the two covariates and each pair of categories, we include sufficient statistics of the form $\sum_{i<j}^{n} 1(x_i=k, x_j=l) y_{ij}$, where $1(x_i = k, x_j = l)$ is an indicator function, taking value 1 if $x_i = k$ and $x_j = l$ and 0 otherwise. To ensure that the model is identifiable, we restrict the natural parameters corresponding to the combinations local—local (scale of operations) and governmental—governmental (type of organisation) to 0. Model M1 can be considered to be a degenerate block model with one block, degenerate in the sense that all organisations are members of the same block with probability 1. We note that model M1 is equivalent to the logistic regression model of Bevc (2010), though we transformed the covariate scale of operations by aggregating some small categories (cf. Bevc, 2010, and Table 1).

Models M1—M3 were estimated by Monte Carlo maximum likelihood algorithms (Handcock, 2003; Hunter and Handcock, 2006) implemented in the R package ergm (Handcock et al., 2010). While it would be preferable to estimate models M1—M3 by Bayesian methods, models M2 and M3 lack conditional independence of edges, implying that special Markov chain Monte Carlo methods along the lines of Koskinen et al. (2010); Caimo and Friel (2011) would be needed, which tend to consume excessive amounts of computing time. The posterior distribution of model M4 was approximated by Markov chain Monte Carlo methods, using 40,000 burn-in iterations, 200,000 post-burn iterations, and keeping track of every 200-th post-burn-in iteration. To solve the so-called label-switching problem, the Markov chain Monte Carlo sample was relabelled along the lines of Stephens (2000). Model M5 was estimated by the maximum likelihood algorithm implemented in the R package latentnet (Krivitsky and Handcock, 2012), because Bayesian Markov chain Monte Carlo algorithms failed to converge in reasonable time.

We compare models M1—M5 by comparing the goodness-of-fit of the models in terms of the following simple, but fundamental functions of the data (Wasserman and Faust, 1994; Hunter et al., 2008): (S1) the number of components (maximal subsets of organisations such that every organisation is connected by a finite path to every other organisation); (S2) the size of the largest component; (S3) the degrees $d_1(\mathbf{y})$, …, $d_n(\mathbf{y})$; and (S4) the number of 2-stars $\sum_{i=1}^{n} \sum_{j_1<j_2}^{n} y_{ij_1} y_{ij_2}$. The observed values of statistics S1 and S2 are given by 237 and 463, respectively. The distribution of the observed degrees S3 is shown in Figure 4 and the observed value of statistic S4 is given by 34,247. We note that goodness-of-fit tests of transitive closure based on the Bayesian score test are conducted in Section 5.4.

Figures 5 and 6 show the goodness-of-fit of models M1—M5 in terms of statistics S1—S4. The lack of fit of models M1—M3 is striking. Model M1, which assumes that edges are independent Bernoulli random variables conditional on covariates, implies too short-tailed degree distributions and is therefore too restrictive to model long-tailed degree distributions and other fundamental functions of the data. Model M2, on the other hand, tends to suffer from model degeneracy (Strauss, 1986; Handcock, 2003; Schweinberger, 2011). The model degeneracy of model M2 impedes Monte Carlo maximum likelihood algorithms, because

such algorithms generate samples of networks and sampled networks from degenerate models do not tend to cluster around the observed network in terms of the sufficient statistics (Snijders, 2002; Handcock, 2003; Schweinberger, 2011). As a result, Monte Carlo maximum likelihood estimators may not exist (Handcock, 2003; Rinaldo et al., 2009), resulting in computational failure and lack of fit (Handcock, 2003; Hunter et al., 2008; Rinaldo et al., 2009). In light of model degeneracy, the lack of fit of model M2 is therefore not surprising. In the case of model M3, the geometric sequence on which the natural parameters $\delta_k(\theta)$ are based may be too restrictive to model the degree distribution and in addition the estimation of curved exponential-family models tends to be challenging. In contrast, model M4 clearly outperforms models M1—M3 in terms of goodness-of-fit. The most notable fact is that the goodness-of-fit of model M4 with respect to the number of 2-stars is excellent—in sharp contrast to model M2 with the number of 2-stars as sufficient statistic. In other words, even when *k*-stars are of primary interest, it seems to be preferable to model degrees rather than *k*-stars. Model M5 outperforms models M1—M3 in terms of goodness-of-fit and compares favorably to model M4 with respect to the degrees and number of 2-stars, but M4 outperforms M5 with respect to the number of components and the size of the largest component. In addition, model M5—the 2-dimensional Euclidean space 4-component mixture model with random effects—is much more complex than model M4 in terms of the number of unknowns (i.e., latent variables and parameters). Thus, by Occam's razor, model M4 is preferable to model M5.

To examine how much the propensities to collaborate vary across organisations, we present in Table 2 the posterior medians and 95% credibility intervals of the sizes n_k and the parameters θ_k of blocks 1, 2, 3, 4. The 95% posterior credibility intervals do not overlap and indicate that the organisations which responded to the September 11, 2001 attacks had varying propensities to collaborate. In addition, the number of organisations which tends to collaborate is small: e.g., fewer than 2% of the organisations is member of block 1 and fewer than 10% of the organisations is member of blocks 1 and 2, which have the highest propensities to collaborate.

5.3 Formal and emergent coordinators

To detect organisations which were critical to coordinating the response to the September 11, 2001 attacks and to answer the question of whether primary coordinators were established or emergent coordinators, we predict the block membership of organisations by assigning every organisation to the block of which it is member with maximum posterior probability.

Based on these predictions, blocks 1, 2, 3, and 4 include 13, 39, 234, and 431 organisations, respectively, and the mean observed degree of organisations in block 1, 2, 3, and 4 is given by 56.54, 18.41, 5.07, and .62, respectively. The core of the network is made up by the collaborations within and between the organisations in block 1, 2, and 3, which we interpret as primary, secondary, and tertiary coordinators, respectively, while the organisations in block 4 are peripheral to the network. To gain more insight into the collaborations within and between blocks, we show the proportion of observed collaborations within and between blocks in Figure 7.

It is evident that primary coordinators dominate the network: Collaboration is most intense among primary coordinators; secondary coordinators tend to collaborate with primary coordinators rather than other secondary coordinators; and tertiary coordinators tend to collaborate most with primary coordinators, less with secondary coordinators, and least with other tertiary coordinators. The presence of primary coordinators had the potential to improve the disaster response, but at the same time may have made disaster response vulnerable to disturbances.

Considering the importance of primary coordinators, it is worthwhile to shed more light on the subset of primary coordinators. The organisations which are predicted to be primary coordinators are shown in Table 3 along with the uncertainty about the prediction in terms of the entropy of the posterior distribution (re-scaled to [0, 1]) and the classification as formal or emergent coordinator. The classification as formal coordinators is based on the rules laid out by the National Response Framework (see http://www.fema.gov/emergency/nrf/). Organisations which are not classified as formal coordinators are classified as emergent coordinators. The Port Authority of New York and New Jersey is an exception and is neither classified as formal nor emergent coordinator, beause it was the owner of the World Trade Center complex. As the owner, it had the floor plans of the entire World Trade Center complex and it is thus not surprising that is was a primary coordinator. Last, the entropy reflects the uncertainty about the prediction of the block memberships, which is ignored by deterministic clustering methods (e.g., Borgatti et al., 2002; Girvan and Newman, 2002; Leskovec et al., 2008). It is worth noting that model M5 could be used to cluster organisations as well. Model M5 seems to cluster close-knit groups of organisations, but the clustering (not shown) is hard to summarise and inconclusive, and organisations without collaborations can be found in all components. We suspect that the optimal number of components (in terms of the marginal likelihood) exceeds 4 by an order of magnitude, but model M5 with the optimal number of components would be even less parsimonious than with 4 components.

An interesting fact is that all of the emergent coordinators turn out to have operated offices close to the World Trade Center. An example is given by Verizon, which was head-quartered in the so-called Verizon building. The Verizon building is located on the edge of the World Trade Center and withstood the September 11, 2001 attacks. Verizon was therefore as close as it could be to the site of the disaster and, owing to the nature of its business, was able to serve as a communication hub in the days following September 11, 2001 (The Washington Post, September 13, 2001). A second example is given by the Health Care Financing Administration, which operated a regional office in Manhattan close to the World Trade Center. The Health Care Financing Administration assisted in the shifting of non-critical patients from hospitals to nursing homes and, in so doing, freed up beds which could be used to serve the wounded of the September 11, 2001 attacks (Comfort et al., 2004, p. 300).

It is worth noting that there are formal coordinators in other blocks. The block of secondary hubs includes many formal coordinators—including the American Red Cross, while the block of tertiary hubs includes some. Most of them are small, e.g., subdivisions of FEMA and the New York City administration. We do not know why these formal coordinators are not primary hubs, though we speculate that the distance to the site of the disaster, lack of resources, lack of communication equipment and procedures, and unwillingness to collaborate with other organisations are among the most important reasons.

5.4 Network redundancy

We argued in Section 2.2 that network redundancy in the form of transitivity can help to absorb disturbances by providing alternative channels of communication. An extension of model M4, which takes an excess of transitivity into account, is given by model (11). To assess whether there is an excess of transitivity, we follow the approach advocated in Section 3.3 and assess the evidence against the null hypothesis

$$H_0 : \gamma = 0, \quad (38)$$

where γ is the transitivity parameter of model (11). We can assess the evidence against H_0 : $\gamma = 0$ by the posterior predictive p-value (Meng, 1994) given by

$$p_t = \Pr(t(\mathbf{Y}^\star) \geq t(\mathbf{y}) | \mathbf{Y} = \mathbf{y}, H_0), \quad (39)$$

where $t(\mathbf{Y}\star)$ is given by the score-based test statistic (17). A Markov chain Monte Carlo approximation of the posterior predictive distribution of the score-based test statistic (17) under $H_0 : \gamma = 0$ is presented in Figure 8 and the corresponding posterior predictive *p*-value is given by $\hat{p_t} = .0006$. In other words, there is more transitivity than expected, which suggests that the inter-organisational network possessed network redundancy.

To see where in the network the transitivity "resides"—i.e., whether the transitivity is concentrated in the core or the periphery of the network or both, we compute the number of triangles in which organisations in blocks 1, 2, 3, and 4 are involved. The number of triangles in which organisations in blocks 1, 2, 3, and 4 on average are involved turn out to be 112.33, 53.45, 4.62, and .08, respectively. Therefore, the transitivity resides first and foremost in the "neighbourhood" of primary and secondary coordinators, i.e., in the core of the inter-organisational network. These results are in line with Butts et al. (2012), who argued that transitivity more often than not is concentrated in the core of the network.

5.5 Network vulnerability

To gain more insight into how vulnerable the inter-organisational network was to disturbances following the initial disaster, we simulate disturbances. We note that, while the impact of the second airplane into the South Tower of the World Trade Center can be thought of as a disturbance, the cross-sectional data at the disposal do not allow us to assess the impact of the second airplane on the disaster response.

We generate 1,000 disturbances as follows. For 1,000 sample points from the posterior distribution of model M4, we sample $i = 1, \ldots, 10$ organisations from block 1, 2, and 3, excluding block 4 which is peripheral and thus ignored; note that, by using the posterior distribution of blocks, we take the uncertainty about the block memberships of organisations into account. We consider the worst-case scenario in which each sampled organisation is rendered inoperational by the disturbances and all of its edges are deleted. We assess the sensitivity of the inter-organisational network to the disturbances in terms of three statistics: (S1) the number of components and (S2) the size of the largest component, as defined in Section 5.2; and (S3) Krackhardt's connectedness (Krackhardt, 1994) defined by $1 - I /$ $\max(I)$, where I is the number of pairs of organisations which are not connected by a finite path and $\max(I) = \begin{pmatrix} n \\ 2 \end{pmatrix}$ is the maximum of I. The results are in Figure 9, which plots the 5%, 50%, and 95% quantiles of statistics S1—S3 against the number of deleted organisations per block.

The results indicate that the inter-organisational network was very vulnerable to deleting a small number of organisations. Deleting more than 5 primary, secondary, and tertiary coordinators can, in the worst case, break the inter-organisational network into many more disconnected components. In particular, by deleting 30 out of 717 organisations—i.e., no more than 5% of the organisations—Krackhardt's connectedness can drop by 50%, which represents a dramatic reduction in connectedness.

A possible explanation is that network redundancy is concentrated in the core of the inter-organisational network, and thus reducing the core of the network reduces network redundancy. Therefore, while small reductions of the core can be compensated, additional reductions of the core cannot.

These findings underscore that the primary, secondary, and tertiary coordinators, while small in number, were vital to the inter-organisational network and the disaster response.

6 Discussion

> *"A systems perspective in disaster preparedness requires inter-organizational planning. Some of the most critical difficulties in disaster response are due to the lack of inter-organizational coordination. Yet, many organizations plan for disaster as if they were to function in isolation. Their disaster plans are conceived with a focus on trees rather than forests" (Auf der Heide, 1989, p. 28).*

The results suggest that a small number of organisations dominated the inter-organisational network that emerged in response to the September 11, 2001 attacks. Some of the dominating organisations were formal coordinators and supposed to respond to disasters, such as FEMA, while others were emergent coordinators, first and foremost Verizon and the Health Care Financing Administration. The emergent coordinators were not supposed to respond to disasters, but may have become involved due to their proximity to the World Trade Center, the nature of their business, and the desire of their management to help. The number of formal and emergent coordinators was small and had both advantages and disadvantages. On the one hand, it had the potential to improve the spread of information and advice and increase the coordination of rescue and relief operations, thus improving the disaster response. On the other hand, it made the disaster response more vulnerable to disturbances. To some extent, the increased vulnerability of the disaster response to disturbances was counterbalanced by network redundancy, which provided alternative channels of communication which could be used in the event of disturbances. A potential problem with network redundancy, however, is that the network redundancy was concentrated in the core of the inter-organisational network, which was composed of key coordinators and associates. Therefore, the inter-organisational network was vulnerable to disturbances harming key coordinators, because reducing the operational capabilities of key coordinators tends to reduce network redundancy. The vulnerability of the inter-organisational network to disturbances harming key coordinators underscores the critical importance of key coordinators.

With a view to planning and preparing the response to future disasters, these results suggest that authorities are well-advised to support formal coordinators. At the same time, formal coordinators cannot be expected to carry the entire burden of the disaster response and must be prepared to cooperate with other organisations, first and foremost emergent coordinators. Formal coordinators should plan, prepare, and train employees in inter-organisational communication, cooperation, and collaboration. During disasters, disaster management should take into account the possible presence of emergent coordinators and allocate resources to formal coordinators as well as emergent coordinators. After disasters, the role of formal coordinators in the disaster response, which may or may not have been in line with the expectations of those who fund them, should be evaluated and emergent coordinators should be acknowledged. With regard to possible disturbances, network redundancy can help to alleviate the impact of disturbances on the disaster response, but the results suggest that key coordinators should be protected against anything that could reduce their ability to coordinate the disaster response.

Last, the Bayesian framework introduced here seems to be a simple and useful starting point in situations where degrees or functions of degrees, such as k-stars, are of primary interest. We have demonstrated that, despite its simple and parsimonious design, the Bayesian framework outperforms competing exponential-family models in terms of goodness-of-fit. Indeed, the Bayesian framework was not only superior in terms of goodness-of-fit, but

preferable in terms of simplicity and computational burden as well. An excess of transitivity and other network characteristics can be tested by using a Bayesian score test as proposed here. An additional advantage of the Bayesian framework is that the computing time is

$O(nK^2)$ without covariates, $O(nK^2 \prod_{i=1}^{I} C_i)$ with categorical covariates, where I is the number of covariates and C_i is the number of categories of the i-th covariate, and $O(n^2 K^2)$ with real-valued covariates. In contrast, the computing time of other, related stochastic block models is more excessive: e.g., the computing time of the stochastic block models of Nowicki and Snijders (2001) and Airoldi et al. (2008) without covariates is $O(n^2 K^2)$. Therefore, the Bayesian framework introduced here—with or without categorical covariates —can be applied to networks with $n \gg 200$, in contrast to competing approaches.

An R package implementing the Bayesian framework will be made publicly available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Airoldi E, Blei D, Fienberg S, Xing E. Mixed membership stochastic blockmodels. Journal of Machine Learning Research. 2008; 9:1981–2014. [PubMed: 21701698]

Auf der Heide, E. Disaster Response Principles of Preparation and Coordination. C. V. Mosby; 1989.

Barabàsi AL, Albert R. Emergence of scaling in random networks. Science. 1999; 286:509–512. [PubMed: 10521342]

Barndorff-Nielsen, OE. Information and Exponential Families in Statistical Theory. New York: Wiley; 1978.

Bera AK, Bilias Y. On some optimality properties of Fisher-Rao score function in testing and estimation. Communications in Statistics—Theory and Methods. 2001a; 30:1533–1559.

Bera AK, Bilias Y. Rao's score, Neyman's $C(\alpha)$ and Silvey's LM tests: an essay on historical developments and some new results. Journal of Statistical Planning and Inference. 2001b; 97:9–44.

Besag J. Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society, Series B. 1974; 36:192–225.

Bevc, C. Ph.D. thesis. Boulder: Department of Sociology, University of Colorado; 2010. Working on the Edge: A Study of Multiorganizational Networks in the Spatiotemporal Context of the World Trade Center Attacks on September 11 2001.

Bhamidi S, Bresler G, Sly A. Mixing time of exponential random graphs. 2008 IEEE 49th Annual IEEE Symposium on Foundations of Computer Science. 2008:803–812.

Bickel P, Chen A. A Nonparametric View of Network Models and Newman-Girvan and Other Modularities. Proceedings of the National Academy of Sciences. 2009; Vol. 106:21068–21073.

Blei D, Jordan M. Variational inference for Dirichlet process mixtures. Journal of Bayesian Analysis. 2005; 1:121–144.

Borgatti, SP.; Everett, M.; Freeman, L. UCINET. Harvard, MA: 2002.

Butts CT. Bernoulli Graph Bounds for General Random Graph Models. Sociological Methodology. 2011; 41:299–345.

Butts CT, Acton RM, Marcum CS. Interorganizational Collaboration In the Hurricane Katrina Response. Journal of Social Structure. 2012; 13:1–37.

Butts CT, Petrescu-Prahova M, Cross BR. Responder Communication Networks in the World Trade Center Disaster: Implications for Modeling of Communication Within Emergency Settings. Journal of Mathematical Sociology. 2007; 31:121–147.

Caimo A, Friel N. Bayesian inference for exponential random graph models. Social Networks. 2011; 33:41–55.

Chang J. lda: Collapsed Gibbs sampling methods for topic models. r package version 1.3.1. 2011

Chatterjee S, Diaconis P. Estimating and understanding exponential random graph models. Annals of Statistics. 2013 in press.

Chib S, Jeliazkov I. Marginal Likelihood from the Metropolis–Hastings Output. Journal of the American Statistical Association. 2001; 96:270–281.

Comfort LK, Ko K, Zagorecki A. Coordination in Rapidly Evolving Disaster Response Systems: The Role of Information. American Behavioral Scientist. 2004; 48:295–313.

Diaconis P, Chatterjee S, Sly A. Random Graphs with a Given Degree Sequence. Annals of Applied Probability. 2011; 21:1400–1435.

Erdös P, Rényi A. On random graphs. Publicationes Mathematicae. 1959; 6:290–297.

Frank O, Strauss D. Markov graphs. Journal of the American Statistical Association. 1986; 81:832–842.

Girvan M, Newman M. Community structure in social and biological networks. Proceedings of the National Academy of Sciences. 2002; Vol. 99:7821–7826.

Haas, J.; Drabek, T. Complex Organizations: A Sociological Perspective. Macmillan; 1973.

Handcock, M. Tech. rep., Center for Statistics and the Social Sciences. University of Washington; 2003. Assessing degeneracy in statistical models of social networks. http://www.csss.washington.edu/Papers.

Handcock MS, Hunter DR, Butts CT, Goodreau SM, Morris M, Krivitsky P. R package ergm version 2.2-2: A Package to Fit Simulate and Diagnose Exponential-Family Models for Networks. 2010 http://CRAN.R-project.org/package=hergm.

Handcock MS, Raftery AE, Tantrum JM. Model-Based Clustering for Social Networks. Journal of the Royal Statistical Society, Series A. 2007; 170:301–354. with discussion.

Hoff P. Bilinear Mixed-Effects Models for Dyadic Data. Journal of the American Statistical Association. 2005; 100:286–295.

Hoff PD, Raftery AE, Handcock MS. Latent Space Approaches to Social Network Analysis. Journal of the American Statistical Association. 2002; 97:1090–1098.

Holland PW, Leinhardt S. An Exponential Family Of Probability Distributions For Directed Graphs. Journal of the American Statistical Association. 1981; 76:33–65.

Hunter DR. Curved exponential family models for social networks. Social Networks. 2007; 29:216–230. [PubMed: 18311321]

Hunter DR, Goodreau SM, Handcock MS. Goodness of fit of social network models. Journal of the American Statistical Association. 2008; 103:248–258.

Hunter DR, Handcock MS. Inference in curved exponential family models for networks. Journal of Computational and Graphical Statistics. 2006; 15:565–583.

Ishwaran H, James LF. Gibbs Sampling Methods for Stick-breaking Priors. Journal of the American Statistical Association. 2001; 96:161–173.

Jonasson J. The random triangle model. Journal of Applied Probability. 1999; 36:852–876.

Jones JH, Handcock M. Likelihood-Based Inference for Stochastic Models of Sexual Network Formation. Population Biology. 2004; 65:413–422.

Kass RE, Raftery AE. Bayes factors. Journal of the American Statistical Association. 1995; 90:773–795.

Kleibergen F, Kleijn R, Tinbergen RP. The Bayesian Score Statistic. Tinbergen Institute Discussion Paper. 2000 TI 00–035/4.

Kolaczyk, ED. Statistical Analysis of Network Data: Methods and Models. Springer; 2009.

Koschade S. A Social Network Analysis of Jemaah Islamiyah: The Applications to Counter-Terrorism and Intelligence. Studies in Conflict and Terrorism. 2006; 29:559–575.

Koskinen JH, Robins GL, Pattison PE. Analysing exponential random graph (p-star) models with missing data using Bayesian data augmentation. Statistical Methodology. 2010; 7:366–384.

Krackhardt, D. Graph Theoretical Dimensions of Informal Organizations. Hillsdale, NJ: Lawrence Erlbaum; 1994. p. 89-111.

Krivitsky P, Handcock MS, Raftery AE, Hoff P. Representing Degree Distributions, Clustering, and Homophily in Social Networks With Latent Cluster Random Effects Models. Social Networks. 2009; 31:204–213. [PubMed: 20191087]

Krivitsky, PN.; Handcock, MS. R package version 2.4-2. 2012. latentnet: Latent position and cluster models for statistical networks.

Lehmann, EL.; Romano, JP. Testing Statistical Hypotheses. 3rd ed.. New York: Springer; 2005.

Leskovec J, Lang K, Dasgupta A, Mahoney M. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. CoRR. 2008 abs/0810.1355.

Liu, JS. Monte Carlo Strategies in Scientific Computing. New York: Springer; 2008.

MacKay, D. Information Theory, Inference and Learning Algorithms. Cambridge University Press; 2003.

McGrory CA, Titterington DM. Variational approximations in Bayesian model selection for finite mixture distributions. Computational Statistics and Data Analysis. 2006; 51:5352–5367.

Meng X-L. Posterior Predictive p-Values. Annals of Statistics. 1994; 22:1142–1160.

Newman MEJ, Strogatz SH, Watts DJ. Random graphs with arbitrary degree distributions and their applications. Physical Reviews. 2001; 64

Nowicki K, Snijders TAB. Estimation and Prediction for Stochastic Block-structures. Journal of the American Statistical Association. 2001; 96:1077–1087.

Pearson K. On the criterion that a given system of observations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Philosophical Magazine. 1900; 50:157–175.

Petrescu-Prahova M, Butts C. Emergent Coordinators in the World Trade Center Disaster. International Journal of Mass Emergencies and Disasters. 2008; 28:133–168.

Rao CR. Large Sample Tests of Statistical Hypotheses Concerning Several Parameters with Applications to Problems of Estimation. Proceedings of the Cambridge Philosophical Society. 1948; 44:50–57.

Rao CR, Poti SJ. On locally most powerful tests when alternatives are one sided. Sankhy . 1946; 7:439.

Richardson S, Green PJ. On Bayesian Analysis of Mixtures with an unknown number of components. Journal of the Royal Statistical Society, Series B. 1997; 59:731–792.

Rinaldo A, Fienberg SE, Zhou Y. On the geometry of discrete exponential families with application to exponential random graph models. Electronic Journal of Statistics. 2009; 3:446–484.

Salter-Townshend M, Murphy TB. Variational Bayesian inference for the latent position cluster model for network data. Computational Statistics & Data Analysis. 2013; 57:661–671.

Schweinberger M. Instability, Sensitivity, and Degeneracy of Discrete Exponential Families. Journal of the American Statistical Association. 2011; 106:1361–1370. [PubMed: 22844170]

Schweinberger, M.; Snijders, TAB. Settings in Social Networks: A Measurement Model. Chap. 10. In: Stolzenberg, RM., editor. Sociological Methodology. Vol. Vol. 33. Boston & Oxford: Basil Blackwell; 2003. p. 307-341.

Snijders TAB. Markov chain Monte Carlo Estimation of exponential random graph models. Journal of Social Structure. 2002; 3:1–40.

Snijders TAB, Nowicki K. Manual for BLOCKS version 1.8. 2007 http://www.gmw.rug.nl/stocnet/.

Snijders TAB, Pattison PE, Robins GL, Handcock MS. New specifications for exponential random graph models. Sociological Methodology. 2006; 36:99–153.

Stephens M. Dealing with label-switching in mixture models. Journal of the Royal Statistical Society, Series B. 2000; 62:795–809.

Strauss D. On a general class of models for interaction. SIAM Review. 1986; 28:513–527.

Tallberg C. A Bayesian approach to modeling stochastic blockstructures with covariates. Journal of Mathematical Sociology. 2005; 29:1–23.

Tierney, KJ.; Trainor, J. Research Progress and Accomplishments 2003–2004. Buffalo, NY: State University of New York at Buffalo, Multidisciplinary Center for Earthquake Engineering Research; 2004. Networks and Resilience in the World Trade Center Disaster; p. 157-172.

Tierney L. Markov Chains for Exploring Posterior Distributions. Annals of Statistics. 1994; 22:1701–1728.

Topper CM, Carley K. A Structural Perspective on the Emergence of Network Organizations. Journal of Mathematical Sociology. 1999; 24:67–96.

van Duijn MAJ, Snijders TAB, Zijlstra BJH. P2: a random effects model with covariates for directed graphs. Statistica Neerlandica. 2004; 58:234–254.

Wasserman, S.; Faust, K. Social Network Analysis Methods and Applications. Cambridge: Cambridge University Press; 1994.

Highlights

- Network of collaborations between 717 organisations responding to September 11, 2001

- Model tailored to collaboration network, modelling degrees, covariates, transitivity

- Bayesian score test of transitivity

- Bayesian model estimation and model selection

- Empirical results: model-based clustering; network redundancy and vulnerability.
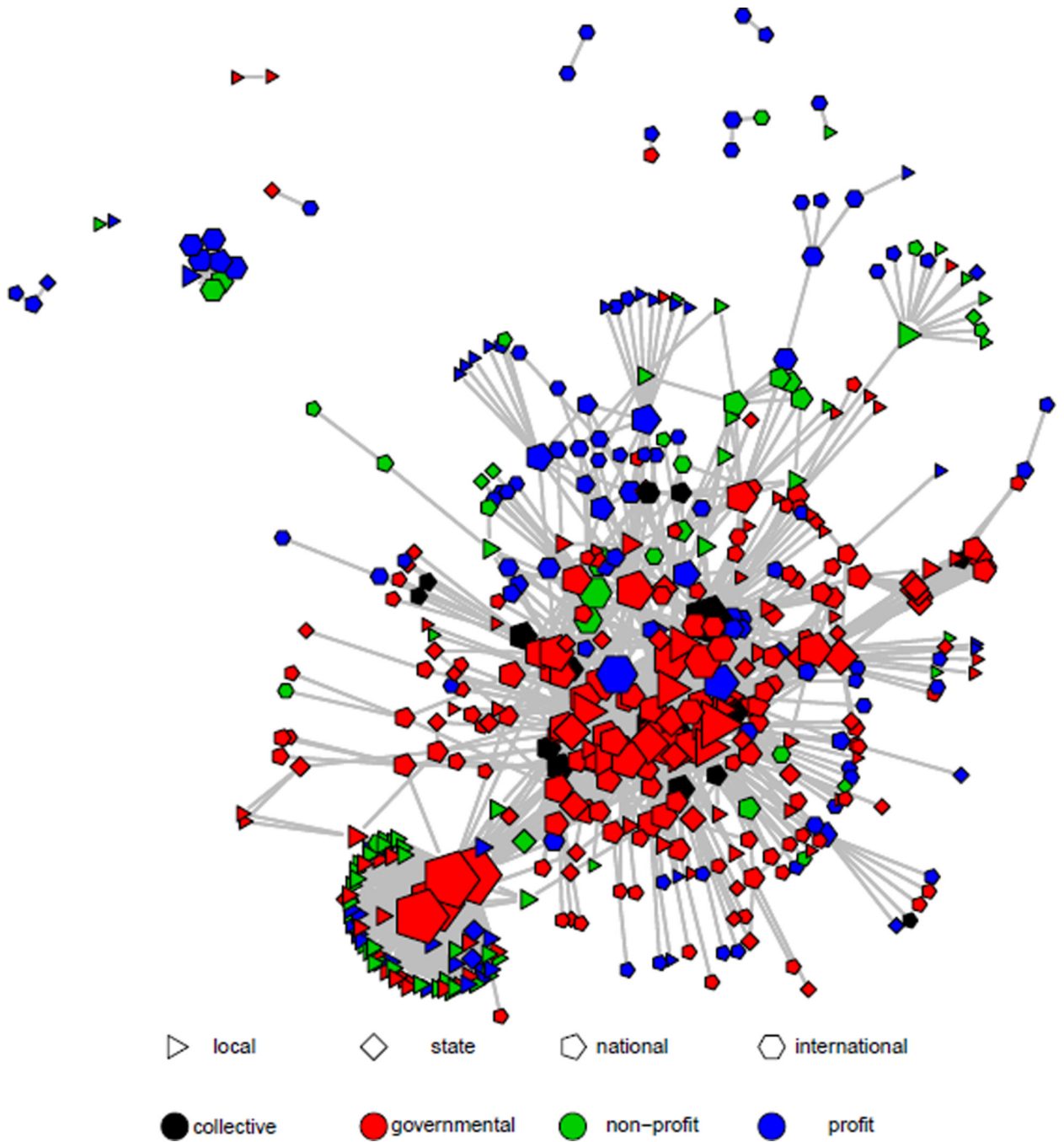
Figure 1.
Inter-organisational network of 717 organisations, excluding 226 organisations which do not
collaborate with others. The shape of the symbol represents the scale of operations, the
colour represents the type of organisation, and the size scales with the number of
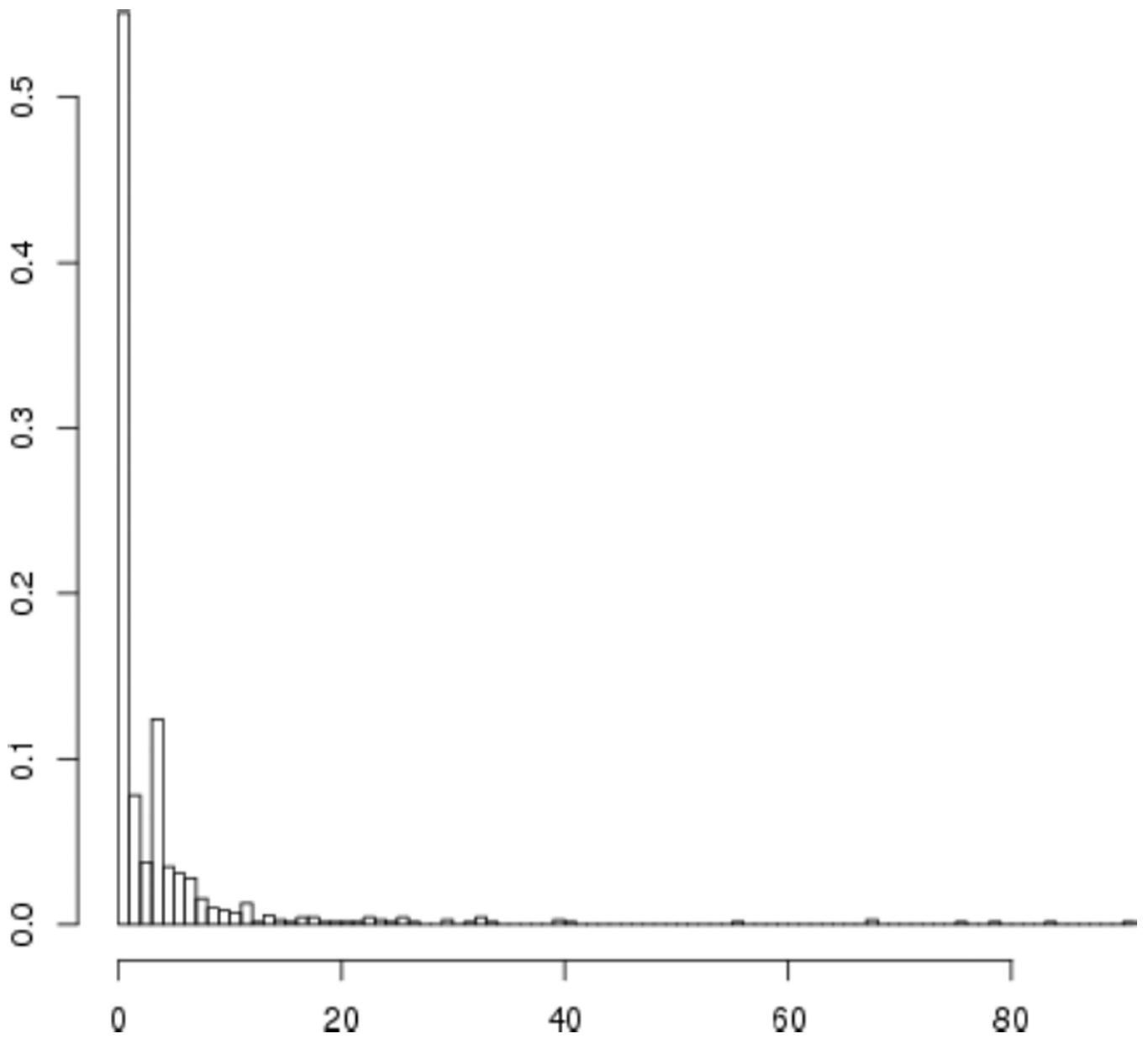collaborations involved
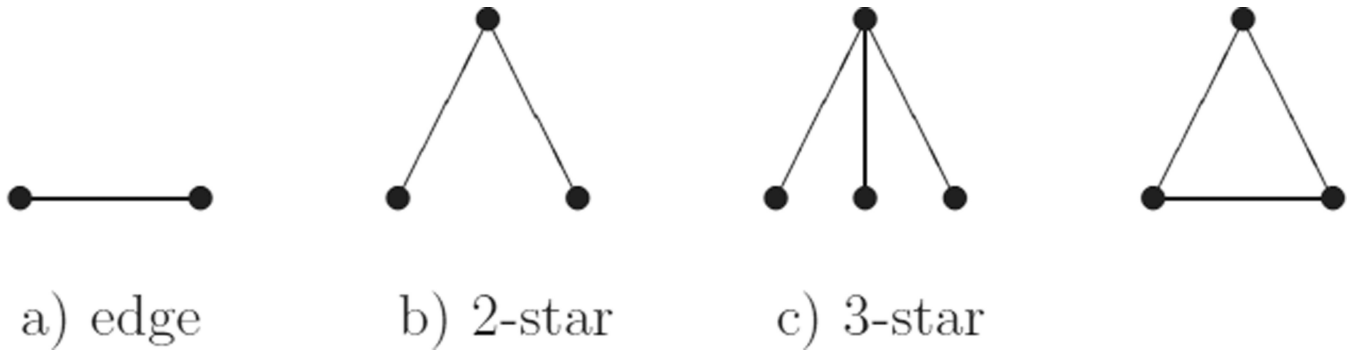
Figure 2.
Network statistics of interest, where circles represent organisations and lines represent collaborations between organisations
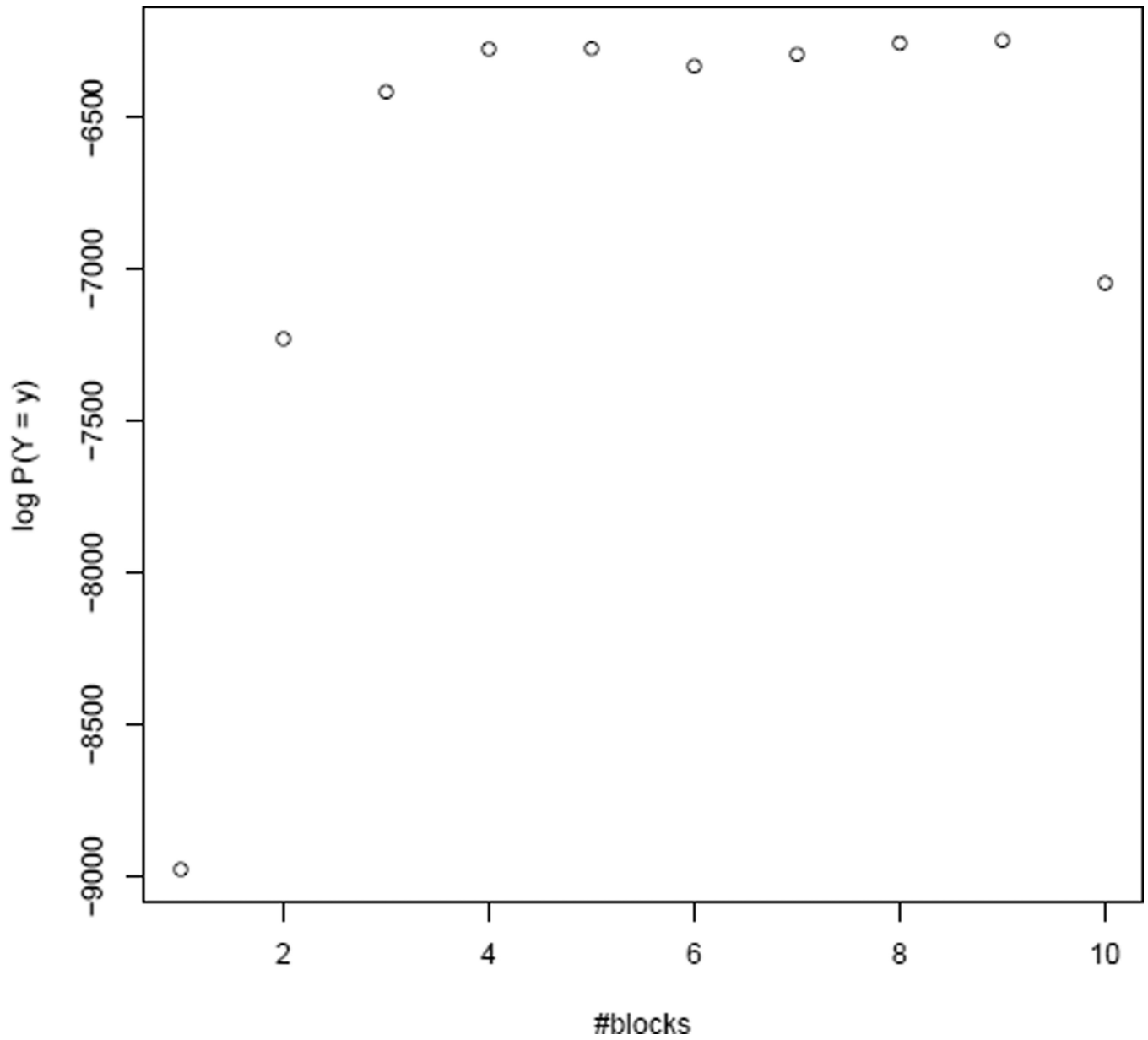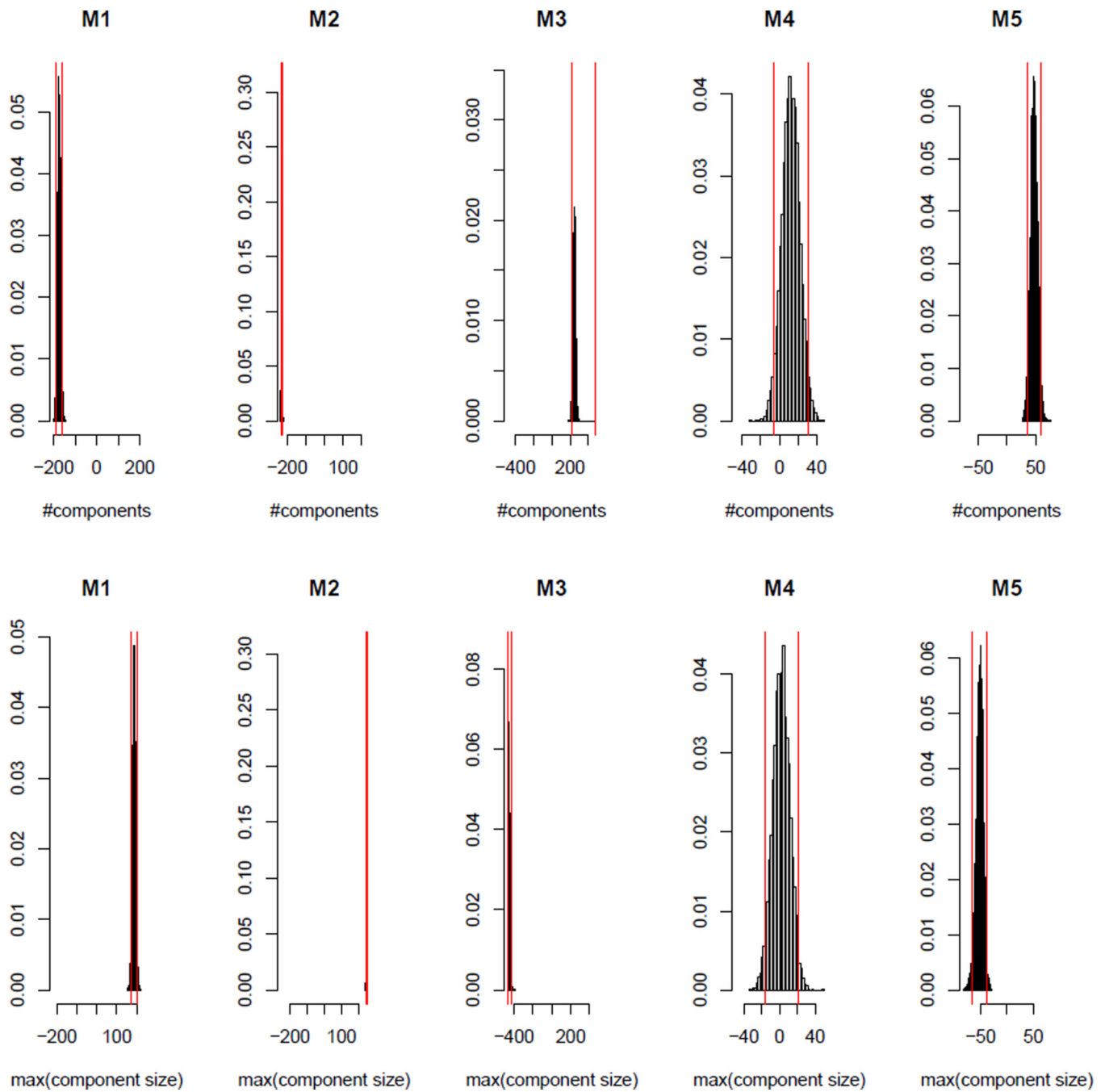
Figure 3.
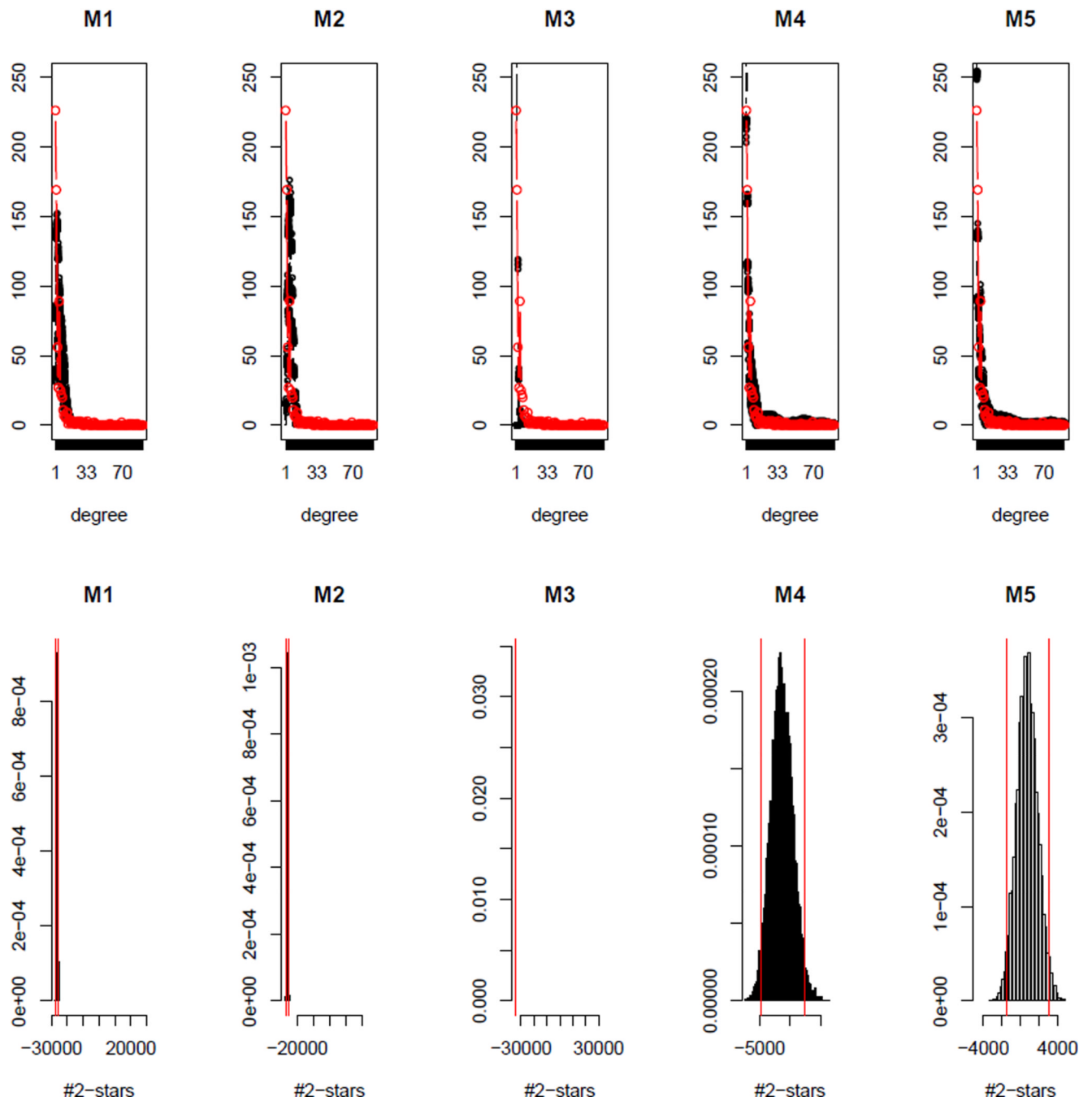Approximations of log $P(\mathbf{Y} = \mathbf{y})$ under the Bayesian model with 1, …, 10 blocks
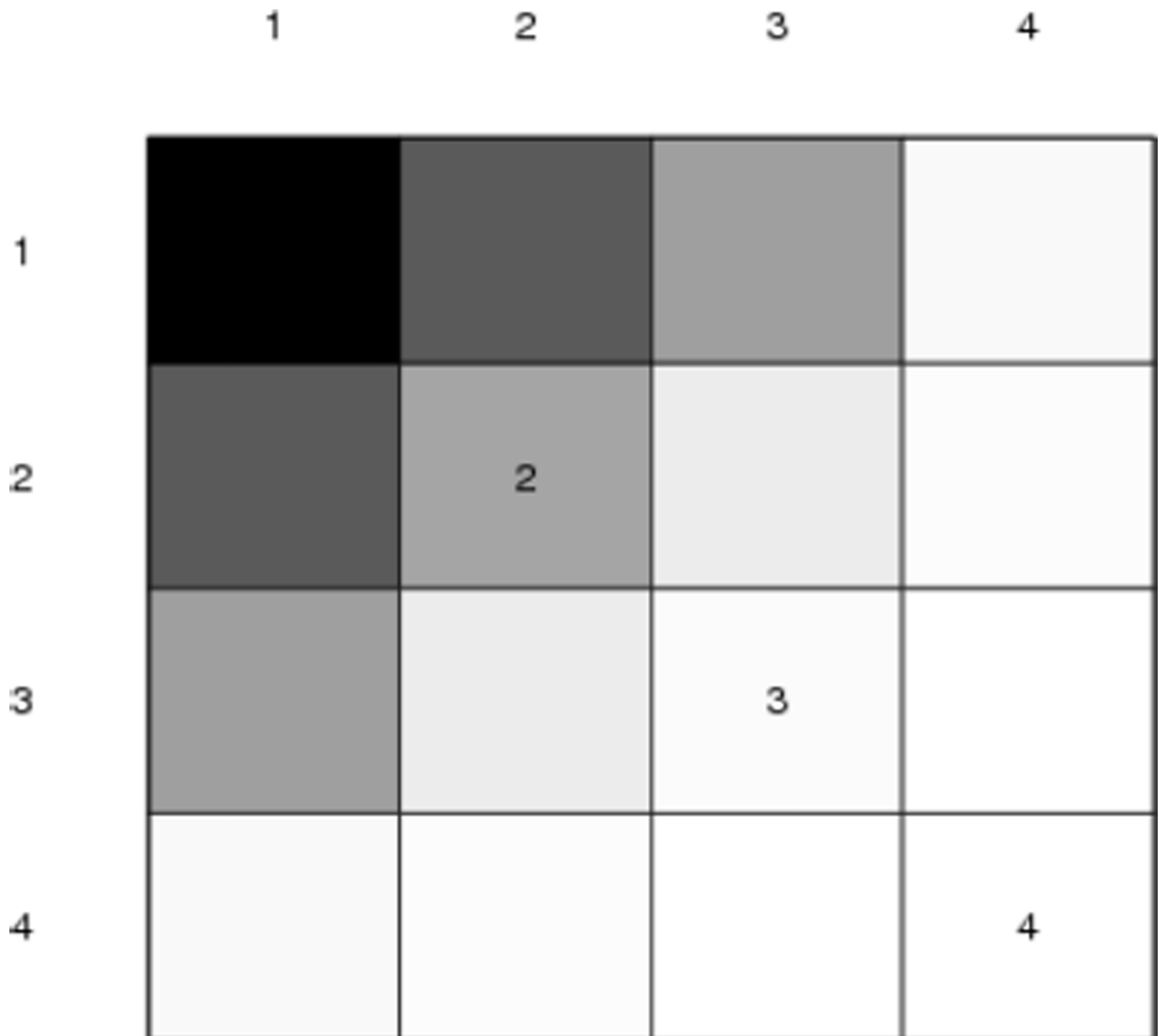
Figure 4.
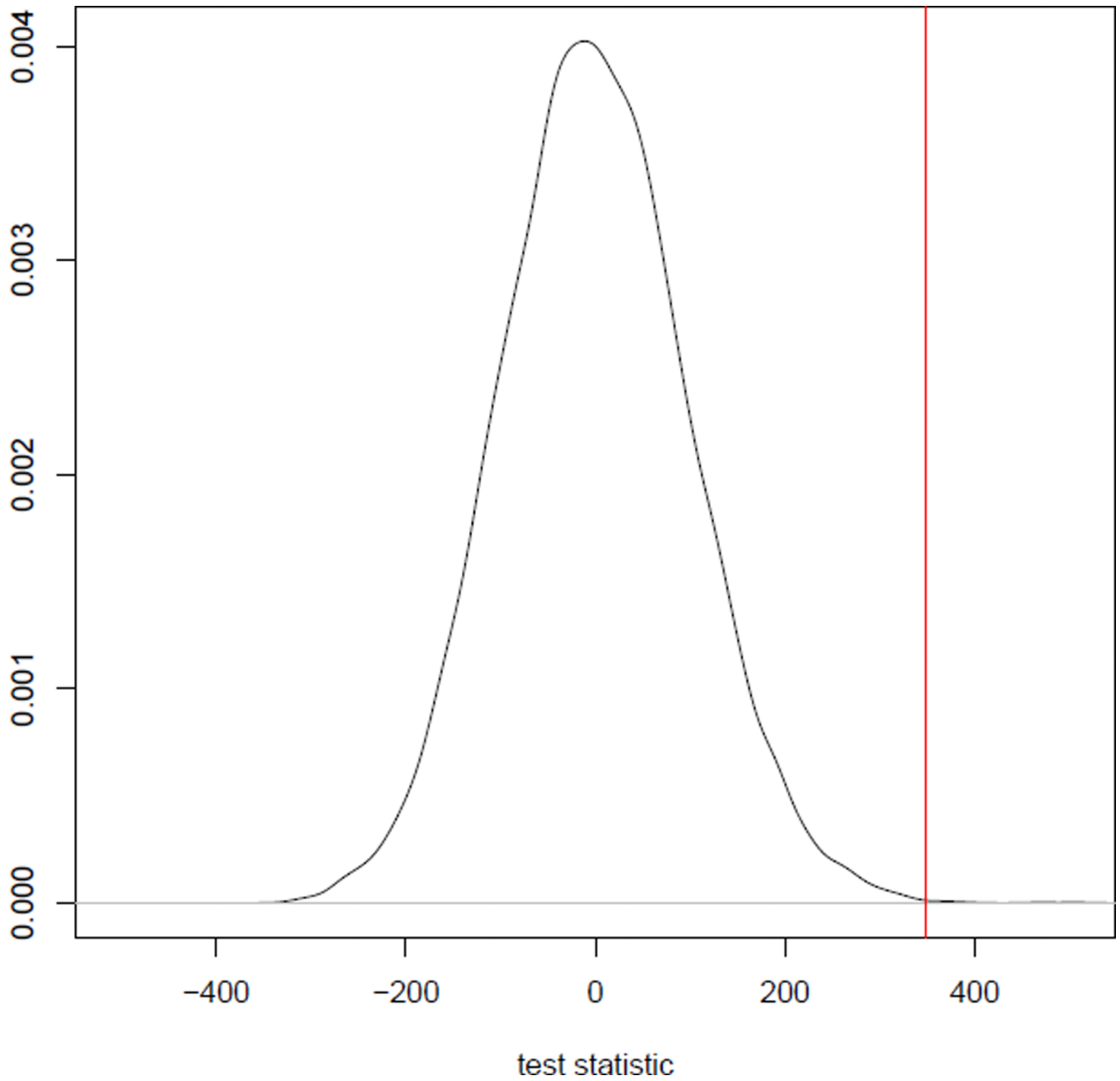Degree distribution of inter-organisational network

Figure 5.
Goodness of fit of models M1—M5 in terms of statistics S1—S2, centred at observed values, where vertical lines indicate .025 and .975 quantiles
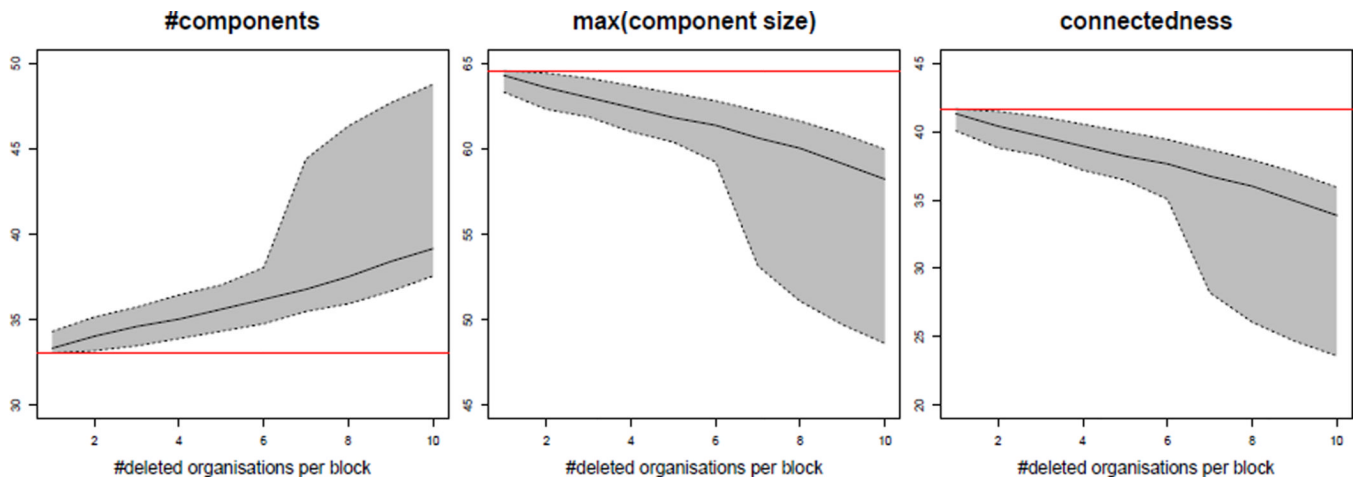
Figure 6.
Goodness of fit of models M1—M5 in terms of statistics S3—S4, centred at observed values, where vertical lines indicate .025 and .975 quantiles

Figure 7.
Proportion of observed collaborations within and between blocks, represented by colour,
ranging from white (proportion 0) to black (proportion .4)

Figure 8.
Posterior predictive distribution of score-based test statistic (17) under $H_0 : \gamma = 0$, where the red line indicates the observed value

Figure 9.
Sensitivity of inter-organisational network to deleting 1, …, 10 organisations from block 1, 2, and 3 in terms of statistics S1—S3: 5%, 50%, and 95% quantiles of statistics S1—S3 in % plotted against number of deleted organisations; red lines indicate values of statistics S1—S3 based on the complete network

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Scale of operations and type of 717 organisations

	collective	governmental	non-profit	profit	total
local	3	86	67	68	224
state	5	67	6	9	87
national	23	157	32	88	300
international	1	8	20	77	106
total	32	318	125	242	717

Table 2

Model M4: Posterior medians and 95% credibility intervals (CIs) of sizes n_k and parameters θ_k of blocks $k = 1, 2, 3, 4$

k	n_k: median	n_k: 95% CI	θ_k: median	θ_k: 95% CI
1	13	(12, 14)	.074	(−.146, .308)
2	49	(34, 64)	−1.38	(−1.596, −1.104)
3	252	(231, 275)	−2.732	(−2.938, −2.492)
4	402	(376, 429)	−4.852	(−5.114, −4.554)

Table 3

Organisations which are predicted to be primary coordinators, where E: emergent coordinator; F: formal coordinator; and O: owner

organisation	coordinator	entropy
Centers for Disease Control and Prevention	F	< .001
Centers for Medicare and Medicaid Services	E	< .001
Department of Health and Human Services	F	< .001
Emergency Support Function 9—Urban Search and Rescue	F	< .001
Federal Emergency Management Agency (FEMA)	F	< .001
Health Care Financing Administration	E	< .001
New York City (NYC)	F	< .001
NYC Emergency Operations Center	F	.010
NYC Emergency Management	F	< .001
NYC Fire Department	F	.479
NYC Police Department	F	< .001
Port Authority of New York and New Jersey	O	.006
Verizon	E	< .001