Check for updates

Received: 20 March 2025 | Accepted: 30 July 2025

DOI: 10.1111/bmsp.70007

the british psychological society
promoting excellence in psychology

TUTORIAL

A tutorial on Bayesian model averaging for exponential random graph models

Ihnwhi Heo[1] | **Jan-Willem Simons**[2] | **Haiyan Liu**[1]

[1]University of California, Merced, Merced, California, USA

[2]Utrecht University, Utrecht, The Netherlands

Correspondence
Ihnwhi Heo, Department of Psychological Sciences, University of California, Merced, 5200 N. Lake Road, Merced, CA 95343, USA.
Email: ihnwhi.heo@gmail.com

Abstract

The use of exponential random graph models (ERGMs) is becoming prevalent in psychology due to their ability to explain and predict the formation of edges between vertices in a network. Valid inference with ERGMs requires correctly specifying endogenous and exogenous effects as network statistics, guided by theory, to represent the network-generating process while ensuring key effects shaping network topology are not omitted. However, specifying a comprehensive model is challenging, particularly when relying on a single model. Despite this, most applied research continues to use a single ERGM, raising two concerns: Selecting misspecified models compromises valid statistical inference, and single-model inference ignores uncertainty in model selection. One approach to addressing these issues is Bayesian model averaging (BMA), which evaluates multiple candidate models, accounts for uncertainty in parameter estimation and model selection, and is more robust to model misspecification than single-model inference. This tutorial provides a guide to implementing BMA for ERGMs. We illustrate its application using data from a college friendship network, with a supplementary example based on the Florentine marriage network; both focus on averaging exogenous covariate effects. We demonstrate how BMA incorporates theoretical considerations and addresses modelling challenges in ERGMs, with annotated R code provided for replication and extension.

KEYWORDS

Bayes factor, Bayesian model averaging, exponential random graph model, model uncertainty, posterior model probability, prior sensitivity, social network analysis

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2025 The Author(s). *British Journal of Mathematical and Statistical Psychology* published by John Wiley & Sons Ltd on behalf of British Psychological Society.

1 | INTRODUCTION

1.1 | Relational data and social network analysis

Individuals, each possessing unique psychological and behavioural attributes, engage in interactions within various social and behavioural contexts. These interactions give rise to structured patterns, which manifest as social networks known as systems of relationships among individuals (e.g., Allan & Phillipson, 2017; Basu & Sen, 2021; Grund & Densley, 2012, 2015; Grunspan et al., 2014; Liu et al., 2018, 2021; McElroy et al., 2019; Pachucki & Breiger, 2010). When such interactions are recorded, they are represented as relational data, which differ from non-relational data. While non-relational data consist of measurements that describe individual characteristics in isolation, relational data capture the presence, directionality, and strength of connections among individuals. As such, relational data provide a powerful lens for understanding individuals both in terms of their own attributes and within the broader social environments that shape their behaviour and experiences.

When research questions involve relational data, social network analysis offers a set of statistical tools for modelling and interpreting the structure and dynamics of social interactions. Social network analysis enables the investigation of social processes by capturing and analysing patterns of relationships within networks (Borgatti et al., 2006; Cranmer et al., 2020; Scott, 2017; Wasserman & Faust, 1994). In recent years, social network analysis has gained growing attention in psychology, as it provides a novel perspective through which classical psychological questions can be revisited and reinterpreted (e.g., Broda et al., 2023; Burt et al., 2013; Gilman et al., 2022; Massidda et al., 2016). Given its versatility and broad applicability, it is important to critically examine the methodological foundations and modelling strategies available within the social network analysis framework to fully leverage its potential in psychological research.

1.2 | Modelling social networks and our focal model

A range of statistical methods has been developed to study the structure and dynamics of social networks, spanning both cross-sectional and longitudinal data. The present study focuses on cross-sectional settings, where many of these methods are commonly applied in psychological research (but see Broda et al., 2023, for approaches to longitudinal social network modelling in psychology). Among these, latent space models (Hoff et al., 2002; Liu et al., 2021; Liu & Zhang, 2021) propose that each individual (or actor, more generally) in a network occupies a position in an unobserved, low-dimensional social space. The distance between two individuals in this latent space predicts the likelihood of a connection in the observed (manifest) network: The probability of a tie decreases as the distance between them increases. These spatial positions are interpreted as reflecting social standing or similarity to others. The overall configuration of positions provides meaningful insight into the network's structure. Therefore, latent space models are effective for uncovering hidden dimensions of social similarity or proximity, but they are less focused on modelling explicit structural configurations of networks.

Stochastic block models (Anderson et al., 1992; Holland et al., 1983; Lee & Wilkinson, 2019; Sweet, 2015), representing another class of network models, group nodes into latent blocks (i.e., unobserved subgroups of actors) under the assumption that the probability of a tie depends primarily on the block membership. These models capture variation in the propensity of two actors being connected based on their respective block memberships. Specifically, stochastic block models assume that the probability of a tie differs systematically between within-block and between-block pairs. They are commonly used for community detection or identifying clusters in networks, placing less emphasis on modelling local dependence structures. While traditional stochastic block models do not include covariates, recent extensions have incorporated covariate information into the modelling framework to enhance their interpretability and flexibility (e.g., Sweet, 2015).

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In this tutorial, we focus on exponential random graph models (ERGMs; Cranmer et al., 2020; Lusher et al., 2013; Robins et al., 2007; Snijders et al., 2006; Wasserman & Pattison, 1996), a widely used class of statistical models for network data. Compared to latent space models or stochastic block models, ERGMs are suited to model local structural configurations and the probabilistic processes underlying network formation. Specifically, ERGMs provide a statistical framework for estimating the likelihood of edge formation (i.e., ties representing relationships) between vertices (i.e., nodes representing individuals) in a network. For instance, in adolescent peer networks (e.g., who befriends whom in a classroom), student characteristics can be used to formulate hypotheses about which ties are more likely to form, with ERGMs providing a method to test these theoretical expectations (Jiao et al., 2017). While ERGMs have long been widely used in empirical sociological research, they are increasingly applied to empirical social network data in psychology as well (see, e.g., Flakus et al., 2021; Holler & Schüßler, 2024; Jiao et al., 2017; Kwiatkowska & Rogoza, 2019; Schüßler et al., 2025; Tejada-Gallardo et al., 2023).

Importantly, ERGMs account for both endogenous dependencies—including network structures such as edges, triadic closure, and degree distributions—and exogenous covariates, such as individual node-level attributes like gender or ethnicity. Here, the 'edges' term models the overall network density by capturing the presence or absence of ties between pairs of nodes; the terms capturing triadic closure reflect the network's tendency toward transitivity by modelling the likelihood that a given set of three nodes forms a closed triangle; and degree-related terms represent the number of connections each node has, reflecting popularity or activity patterns within the network. By incorporating both endogenous and exogenous effects, ERGMs enable researchers to model and test multiple theoretical predictions about the underlying generative mechanisms driving social network formation (Cranmer et al., 2020; Light & Moody, 2020).

An essential step in fitting any ERGM in an applied research setting is to select variables based on theoretical considerations that explain the formation of ties in the network, similar to the variable selection process in linear regression analysis. However, unlike standard regression models, specifying ERGMs requires accounting for both the structural properties of the network and the attributes of individuals within it, making model specification more complex. A subtle yet crucial aspect of model specification is determining which terms (i.e., factors hypothesized to drive network formation) should be included. As we shall elaborate, this aspect is often underappreciated but carries substantial methodological implications. In the following sections, we first discuss the challenges associated with ERGM (mis)specification and model uncertainty, followed by an introduction to Bayesian model averaging as an alternative modelling approach.

1.3 | Challenges of ERGM (mis)specification and model uncertainty

Valid inference using ERGMs rests on a crucial assumption: The model is correctly specified and contains no omitted variables (Bull et al., 1994). Satisfying this assumption requires including all variables that predict network formation in the analysis. This ensures that, for any two networks with the same number of nodes and identical values for the specified network statistics (i.e., endogenous and exogenous effects), the probability of observing each network is equal (Cranmer et al., 2020). Differences in probabilities would then indicate the presence of unmodelled effects influencing the network structure. Consequently, it is essential to include all theoretically relevant endogenous and exogenous variables in ERGMs to control for pertinent factors and draw accurate conclusions about the generative processes underlying the observed network structures. Failure to meet this assumption results in model misspecification, which can distort statistical inference due to omitted variable bias (Bull et al., 1994).

In practice, the assumption that an ERGM is correctly specified with no omitted variables—and thus provides valid inference—is often unrealistic. Achieving such a specification is particularly challenging because accurately capturing the true network-generating process requires the precise and deliberate inclusion of effects from both endogenous dependencies and exogenous covariates inherent in complex social networks. Depending on theoretical or statistical considerations, a wide range of

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

endogenous terms can be incorporated or even constructed in ERGMs. Moreover, endogenous dependencies and exogenous covariates often (are thought to theoretically) interact, further complicating the accurate representation of the observed network structure. When the dynamic temporal nature of social networks is taken into account (e.g., Leifeld et al., 2018), an additional layer of complexity is introduced, making valid inference with ERGMs even more challenging (see, e.g., Koskinen et al., 2015; Krivitsky & Goodreau, 2019, for separable temporal ERGMs and longitudinal ERGMs).

The task of accurately capturing the true network-generating process presents two major challenges when researchers apply ERGMs to relational data. First, there is a non-negligible possibility that researchers will select misspecified ERGMs. While extremely misspecified models are degenerative—making parameters inestimable because the estimation equations are only satisfied under (almost) complete or (very) sparse graphs (Cranmer et al., 2020; Lusher et al., 2013; Snijders et al., 2006)—more subtle degrees of model misspecification can be hard to detect. As such, single-model selection can compromise valid statistical inference while remaining unnoticed by researchers. Second, most applied ERGM studies rely on a single ERGM specification chosen based on theoretical considerations. Methodological issues stemming from single-model inference have been widely discussed (Hinne et al., 2020; Hoeting et al., 1999; Montgomery & Nyhan, 2010). In short, proceeding with statistical analyses using a single ERGM ignores the uncertainty inherent in model selection. In other words, researchers relying on a single model overlook the possibility that alternative ERGM specifications may better capture the generative processes underlying network formation.

We note that model uncertainty can arise at various decision points in the ERGM specification process. First, even under the assumption of a single true network-generating process, identifying the terms that best represent this process is a nontrivial task that requires careful theoretical and empirical justification. Second, when the research objective involves comparing multiple competing network-generating processes, the main challenge lies in correctly specifying these alternatives and accounting for the uncertainty surrounding them in order to identify the most plausible generative mechanism. Third, regardless of whether a single or multiple true network-generating processes exist and are correctly specified, determining which control variables to include, how to specify them, and how to quantify the uncertainty associated with these choices remains crucial. Given the breadth of model specification choices faced by applied researchers, concerns about single-model ERGM inference become evident in many applied settings. These challenges underscore the need for alternative approaches that are robust to model misspecification and that explicitly account for model uncertainty. Addressing these concerns can improve the reliability of inferences drawn from relational data—an issue that remains largely overlooked in the extant ERGM literature on psychological applications.

1.4 | Bayesian model averaging: An alternative approach

The fit of ERGMs is typically evaluated using goodness-of-fit measures for each hypothetical model fitted to the data (Hunter et al., 2008). The purpose of assessing goodness-of-fit is to determine whether the observed network deviates from the distribution of network statistics—particularly those not explicitly modeled (e.g., degree distribution, geodesic distances, shared partner distributions, and other structural properties)—derived from simulated networks implied by the model. If the goodness-of-fit assessment is unsatisfactory, researchers may iteratively add or remove parameters to improve model fit. However, this process can be time-consuming and may lead to overfitting, producing a model that reflects sample-specific noise rather than generalizable structure. In particular, as researchers iteratively modify the model based on observed goodness-of-fit statistics, they may inadvertently tune the model to idiosyncratic aspects of the dataset. This phenomenon is known as capitalization on chance and occurs when the iterative evaluation process exploits random variation in the data, resulting in a model that fits the sample well but may not generalize beyond it. In addition, this approach overlooks the possibility that alternative model specifications may offer equally or more plausible explanations for the observed social network.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

To address the outlined methodological issues with single model inference for ERGMs—namely, violations of correct model specification and the ignoring of model uncertainty—we point to Bayesian model averaging (BMA) as a promising alternative (Hinne et al., 2020; Hoeting et al., 1999; Raftery et al., 1997). BMA is a multi-model inferential method (van den Bergh et al., 2021) that explains or predicts a phenomenon by averaging parameter estimates across a set of candidate models, weighted by their posterior model probabilities. By averaging over this model space, BMA naturally incorporates model selection uncertainty into the inference process, reducing the risk of selecting a single misspecified model. Furthermore, the advantages of Bayesian statistics (e.g., van de Schoot et al., 2021) enable researchers to account for uncertainty in parameter estimation. To implement BMA in the context of ERGMs, researchers can specify multiple models with different endogenous dependencies or exogenous covariates and average over them to obtain model-averaged parameter estimates. As Hinne et al. (2020) highlighted, this approach can reduce overconfidence in model selection, acknowledge the associated uncertainty, and provide robust estimates even in the presence of model misspecification.

1.5 | Bridging the gap: Applying BMA to ERGMs in psychological research

The benefits of BMA for modelling ERGMs are clear and compelling. While the potential for BMA in ERGMs has been acknowledged in Caimo and Friel (2011), its application in psychological research remains limited. We believe that one main reason for this underutilization is the lack of methodological guidelines and practical demonstrations on how to apply BMA in psychology, which hinders its adoption among applied researchers.

To bridge this gap and further engage the psychological research community in using BMA to enhance the validity of ERGM inference, this tutorial has two objectives: (1) to outline the methodological foundations of BMA for ERGMs and (2) to demonstrate its application through the analysis of empirical social network data. In our empirical demonstration, we provide annotated R code and discuss key theoretical and practical considerations that mirror the types of decisions applied researchers typically face. By bridging the methodological foundations of BMA with its application in psychological research, this tutorial provides a step-by-step guide to promote the benefits of BMA in ERGMs for psychological researchers.

1.6 | Structure of this tutorial

We organize the remainder of this tutorial as follows. First, we introduce Bayesian inference for ERGMs, covering the basic model definition, Bayesian parameter estimation, and Bayesian model selection techniques. Next, we outline the theoretical background of BMA, including Bayesian model-averaged parameter estimates and the inclusion Bayes factor. We then address practical considerations for implementing BMA in ERGMs, namely, the selection of candidate models and the specification of prior model probabilities. Following this, we present an empirical analysis using real social network data to demonstrate the real-world applicability of BMA in ERGMs. Our analysis also examines the sensitivity of results to different prior model probabilities. Finally, we summarize key findings and provide directions for future methodological research.

2 | BAYESIAN EXPONENTIAL RANDOM GRAPH MODELLING

In the analysis of relational data, researchers often seek to explain the formation of complex patterns of interaction among individuals within a given social context. These interactions are typically understood to arise from both the characteristics of the actors themselves (i.e., exogenous factors) and the structure of

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

relationships between them (i.e., endogenous factors). For example, researchers might investigate whether individuals A and B are more likely to become friends if they share a common characteristic, such as gender. Alternatively, they may examine whether A is more likely to reciprocate a friendship initiated by B, or whether individuals A and B are more likely to become friends due to the connection to a mutual friend, C.

The ERGM is a useful analytical tool for examining and predicting the probabilistic formation of edges in a network while incorporating covariates from both endogenous and exogenous sources. For more detailed discussions on the model definition of the ERGM, readers are encouraged to consult Lusher et al. (2013), Cranmer et al. (2020), and Robins et al. (2007). Here, we provide a brief definition of the model, which is central to understanding the application of BMA to ERGMs.

2.1 | The exponential random graph model

A random network $\boldsymbol{Y}$ comprises a set of n nodes and $n \times n$ dyads. The network is often represented as an adjacency matrix, where the entry in the i'th row and j'th column, denoted as $\{Y_{ij}; i = 1, \ldots, n; j = 1, \ldots, n\}$, represents the relation between nodes (or vertices) i and j. In a network of binary relationships, $Y_{ij} = 1$ if there is a link from node i to node j, and 0 otherwise. The diagonal elements Y_{ii} are set to 0 because nodes are typically assumed not to form edges with themselves (e.g., in the case of friendship). For undirected relations, where the connection has no direction, $Y_{ij} = Y_{ji}$, making the adjacency matrix Y symmetric.

ERGMs model the probability distribution of $\boldsymbol{Y}$ as follows:

$$p(\boldsymbol{Y} = \boldsymbol{y}|\boldsymbol{\theta}) = \frac{\exp\{\boldsymbol{\theta}^T \cdot s(\boldsymbol{y})\}}{\mathcal{Z}(\boldsymbol{\theta})} = \frac{\exp\{\boldsymbol{\theta}^T \cdot s(\boldsymbol{y})\}}{\sum_{\boldsymbol{y} \in \boldsymbol{Y}} \exp\{\boldsymbol{\theta}^T \cdot s(\boldsymbol{y})\}}, \tag{1}$$

where $\boldsymbol{y}$ is a realization of the random network $\boldsymbol{Y}$, $s(\boldsymbol{y})$ is a known vector of network statistics computed on the network—such as centrality measures (i.e., indicators of node importance, such as how many ties a node has) and the number of triangles (i.e., closed triplets of nodes where each node is connected to the other two)—$\mathcal{Z}(\boldsymbol{\theta})$ is the normalizing constant, and $\boldsymbol{\theta}$ is a vector of model parameters that describe the dependence of $p(\boldsymbol{y}|\boldsymbol{\theta})$ on $s(\boldsymbol{y})$.

The parameter estimation of ERGMs is analytically challenging due to the intractability of the normalizing constant $\mathcal{Z}(\boldsymbol{\theta})$ and the issue of model degeneracy (Handcock et al., 2003; Rinaldo et al., 2009; Snijders et al., 2006). Specifically, evaluating the normalizing constant for all but trivially small networks complicates the handling of ERGMs. Model degeneracy is another common issue in ERGMs. It refers to a situation in which the probability distribution specified by an ERGM is concentrated on a small number of graph topologies, such as empty or complete graphs (Handcock et al., 2003; Lusher et al., 2013; Rinaldo et al., 2009; Snijders et al., 2006). Graph topologies refer to distinct structural configurations of a network, characterized by patterns of node connectivity, including which nodes are connected, the density and arrangement of ties, and structural features such as clusters or hubs. Model degeneracy can lead to poor model fit and often renders the estimation process unstable or infeasible because the probability distribution assigns near-zero probability to most other possible networks in the sample space. Classic model degeneracy, often caused by factors such as transitive closure, arises from the specification chosen by researchers. This sensitivity to model misspecification is a critical challenge in ERGMs: When degeneracy occurs, the estimation process fails, as no update to a network can be made, rendering the model inestimable.

Classical estimators for finding parameter solutions of ERGMs include maximum pseudo-likelihood estimation (Besag, 1974; Strauss & Ikeda, 1990) and Monte Carlo maximum likelihood estimation (Geyer & Thompson, 1992). However, Caimo et al. (2022) demonstrated that these classical methods produce poor estimates when ERGMs are near-degenerate. In particular, pseudo-likelihood estimators have been shown to be biased (van Duijn et al., 2009).

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Compared to classical methods, the Bayesian estimation framework offers the advantage of circumventing the need to calculate the intractable normalizing constant. Specifically, Bayesian estimation based on the exchange algorithm has demonstrated good performance by producing simulated networks from the posterior distribution whose topologies closely match the observed data (see Caimo et al., 2022). We explain the mechanics of the exchange algorithm in more detail in the subsection on Bayesian ERGM parameter estimation. In addition, the Bayesian method presented by Caimo et al. (2022) has shown that the Markov chain converges to the posterior even when the parameters are initially set in a degenerate region. In the following subsections, we describe the Bayesian framework for ERGMs with respect to parameter estimation and model selection (Caimo et al., 2022; Caimo & Friel, 2011).

2.2 | Bayesian ERGM parameter estimation

A key tenet of Bayesian inference is the incorporation of a prior distribution with the likelihood of the data to obtain the posterior distribution of the parameter of interest based on Bayes' theorem:

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{y})} \propto p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}). \tag{2}$$

The Bayesian paradigm combines prior knowledge about the parameters of interest (i.e., the prior distribution, $p(\boldsymbol{\theta})$) with information obtained from the data (i.e., the likelihood, $p(\boldsymbol{y}|\boldsymbol{\theta})$) to produce updated knowledge in the form of the posterior distribution, $p(\boldsymbol{\theta}|\boldsymbol{y})$). The term $p(\boldsymbol{y})$ in the denominator is the marginal likelihood, which serves as the normalizing constant to ensure that the posterior density is proper. As a result, the posterior distribution is proportional to the product of the likelihood and the prior distribution.

The analytic derivation of the posterior distribution is computationally challenging, and this intractability is further complicated in the Bayesian estimation of ERGMs due to the normalizing constant in Equation (1) and the marginal likelihood in Equation (2). To address this intractable computational challenge, Caimo et al. (2022) proposed the exchange algorithm to sample from the following augmented distribution:

$$p(\boldsymbol{\theta}', \boldsymbol{y}', \boldsymbol{\theta}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})h(\boldsymbol{\theta}'|\boldsymbol{\theta})p(\boldsymbol{y}'|\boldsymbol{\theta}'). \tag{3}$$

The idea of the exchange algorithm is to sample parameter values from an augmented distribution $p(\boldsymbol{\theta}', \boldsymbol{y}', \boldsymbol{\theta}|\boldsymbol{y})$, where $h(\boldsymbol{\theta}'|\boldsymbol{\theta})$ is a proposal distribution, commonly specified as a multivariate normal distribution (e.g., Koskinen et al., 2010) for the augmented variable $\boldsymbol{\theta}$, and $p(\boldsymbol{y}'|\boldsymbol{\theta}')$ refers to the likelihood that the simulated $\boldsymbol{y}'$ are defined and have the same exponential family of densities as $p(\boldsymbol{y}|\boldsymbol{\theta})$.

The exchange algorithm resembles the Metropolis–Hastings algorithm but is adapted for the doubly intractable settings of Bayesian ERGMs, due to the intractability of direct sampling from the posterior distribution and the presence of an intractable normalizing constant in the likelihood. Steps of the exchange algorithm to obtain samples from the marginal distribution for $\boldsymbol{\theta}$ in Equation (3) are described as follows (for a full illustration, see Caimo et al., 2022):

1. Update $(\boldsymbol{\theta}', \boldsymbol{y}')$ by drawing:
   (i) $\boldsymbol{\theta}'$ from $h(\cdot|\boldsymbol{\theta})$, and
   (ii) $\boldsymbol{y}'$ from $p(\cdot|\boldsymbol{\theta}')$.
2. Propose to move from $\boldsymbol{\theta}$ to $\boldsymbol{\theta}'$ with acceptance probability $\mathscr{A}$:

$$\mathscr{A} = \min\left(1, \frac{\exp\{\boldsymbol{\theta}^{\top}s(\boldsymbol{y}')\}p(\boldsymbol{\theta}')h(\boldsymbol{\theta}|\boldsymbol{\theta}')\exp\{\boldsymbol{\theta}'^{\top}s(\boldsymbol{y})\}}{\exp\{\boldsymbol{\theta}^{\top}s(\boldsymbol{y})\}p(\boldsymbol{\theta})h(\boldsymbol{\theta}'|\boldsymbol{\theta})\exp\{\boldsymbol{\theta}'^{\top}s(\boldsymbol{y}')\}} \times \frac{\mathscr{Z}(\boldsymbol{\theta})\mathscr{Z}(\boldsymbol{\theta}')}{\mathscr{Z}(\boldsymbol{\theta})\mathscr{Z}(\boldsymbol{\theta}')}\right),$$

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

where the intractable normalizing constants $\mathcal{Z}(\boldsymbol{\theta})$ and $\mathcal{Z}(\boldsymbol{\theta}')$ cancel out. In the exchange step, the observed data $\boldsymbol{y}$ are evaluated under the proposed parameter $\boldsymbol{\theta}'$, while the simulated data $\boldsymbol{y}'$ are evaluated under the current parameter $\boldsymbol{\theta}$. The compatibility of $\boldsymbol{\theta}$ with $\boldsymbol{y}'$ is assessed through the ratio $\frac{\exp\{\boldsymbol{\theta}^\top s(\boldsymbol{y}')\}}{\exp\{\boldsymbol{\theta}^\top s(\boldsymbol{y})\}}$, and similarly, the compatibility of $\boldsymbol{\theta}'$ with $\boldsymbol{y}$ is evaluated using $\frac{\exp\{\boldsymbol{\theta}'^\top s(\boldsymbol{y})\}}{\exp\{\boldsymbol{\theta}'^\top s(\boldsymbol{y}')\}}$. Supposing the proposal distribution $h(\cdot)$ is symmetric, the acceptance probability $\mathcal{A}$ simplifies to:

$$\mathcal{A} = \min\left(1,\ \frac{\exp\{\boldsymbol{\theta}^\top s(\boldsymbol{y}')\}p(\boldsymbol{\theta}')\exp\{\boldsymbol{\theta}'^\top s(\boldsymbol{y})\}}{\exp\{\boldsymbol{\theta}^\top s(\boldsymbol{y})\}p(\boldsymbol{\theta})\exp\{\boldsymbol{\theta}'^\top s(\boldsymbol{y}')\}}\right).$$

3. Accept the exchange move if $\boldsymbol{y}'$ closely approximates $\boldsymbol{y}$ in terms of the summary statistics, which are sufficient statistics and thus preserve the information needed to approximate the true posterior density.

To illustrate these steps more intuitively, the exchange algorithm draws $\boldsymbol{y}'$ from the distribution $p(\boldsymbol{y}'|\boldsymbol{\theta}')$ using a Markov chain Monte Carlo (MCMC) simulation, since direct sampling is not possible due to the intractable normalizing constant in the ERGM likelihood (Caimo et al., 2022). At each MCMC iteration, a network is simulated given a sample parameter value, using an MCMC algorithm. During this iterative process, the probability of the proposed graph is compared to that of the observed network, and the algorithm decides whether to accept or reject the proposed network. After this step, a change is made to the proposed network, for instance, by creating a new edge or dropping an existing edge. In this tutorial, we use the R package `Bergm` (Caimo & Friel, 2013, 2014), which offers functionality for approximating the posterior distribution via the exchange algorithm. The package improves the mixing of the MCMC algorithm by implementing a parallel adaptive direction sampler (Roberts & Gilks, 1994).

An important feature of Bayesian inference is the ability to incorporate prior distributions to model uncertainty about parameters (i.e., uncertainty in their values). In `Bergm`, prior distributions can be specified for the coefficients of endogenous and exogenous effects using the mean vector and variance-covariance matrix of a multivariate normal distribution. The mean vector determines the prior location (i.e., the expected value before seeing data), while the covariance matrix determines the prior spread. In this way, researchers can control the degree of prior informativeness and incorporate domain knowledge: Tighter variances reflect stronger prior certainty, while larger variances reflect weaker, more diffuse prior beliefs (for details and examples, see Caimo & Friel, 2013, 2014). Let $\boldsymbol{\beta}_D$ be a D-dimensional vector of the coefficients of endogenous and exogenous terms. The prior distribution for these ERGM parameters can be specified in the following way:

$$\boldsymbol{\beta}_D \sim \mathcal{MVN}(\boldsymbol{\mu}_D, \boldsymbol{\Sigma}_D), \tag{4}$$

where the default prior specification for $\boldsymbol{\beta}_D$ is the multivariate normal distribution with the D-dimensional mean vector of 0 and the diagonal variance–covariance matrix with elements of 10^2. In practice, diagonal elements of the $\boldsymbol{\Sigma}_D$ matrix are specified to be weakly informative (Gelman et al., 2008), a choice that has been shown to be necessary when ERGMs are complex or contain missing values (Krause et al., 2020). Accordingly, we account for these practical considerations by employing weakly informative prior settings, as will be detailed in the 'Empirical Analysis of a College Friendship Network' section.

Finally, the convergence and mixing of the MCMC chains used to approximate the posterior distribution can be assessed using standard diagnostic tools (Caimo et al., 2022). In `Bergm`, the MCMC diagnostics include the marginalized density plot, the trace plot, and the autocorrelation plot (Caimo & Friel, 2013, 2014). Researchers can then summarize the results using the posterior mean, posterior standard deviation, native standard error, time-series standard error, and Bayesian p-values.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.3 | Bayesian ERGM model selection

The posterior distribution in Equation (2) provides information about the parameters of interest while accounting for parameter uncertainties. To compare several candidate models, however, Bayes' theorem needs to be extended to the model level to obtain the posterior model probability:

$$p(\mathcal{M}_h|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\mathcal{M}_h) \times p(\mathcal{M}_h)}{\sum_{k=1}^{K} \left[p(\boldsymbol{y}|\mathcal{M}_k) \times p(\mathcal{M}_k)\right]}. \tag{5}$$

In Equation (5), the posterior model probability, $p(\mathcal{M}_h|\boldsymbol{y})$, represents the relative probability assigned to model $\mathcal{M}_h$ after observing the data, while the prior model probability, $p(\mathcal{M}_h)$, represents the relative probability assigned to the model before any data were observed. Here, the subscript h indexes an arbitrary ERGM in the candidate model set. The posterior model probability of a given model is computed by dividing the product of its marginal likelihood $p(\boldsymbol{y}|\mathcal{M}_h)$ and its prior model probability by the sum of such products across all models in the candidate set. As such, the posterior model probability reflects the relative support for that model, balancing how well it explains the observed data (via the marginal likelihood, which serves as the model evidence, as explained below) and how plausible it was considered a priori, relative to all other models under consideration.

Another important term is the marginal likelihood, $p(\boldsymbol{y}|\mathcal{M}_h)$, which serves as an indicator of predictive performance or model evidence in the context of Bayesian model comparison.[1] The marginal likelihood is computed as follows:

$$p(\boldsymbol{y}|\mathcal{M}_h) = \int p(\boldsymbol{y}|\boldsymbol{\theta}, \mathcal{M}_h) \times p(\boldsymbol{\theta}|\mathcal{M}_h) d\boldsymbol{\theta}. \tag{6}$$

In `Bergm`, the marginal likelihood can be obtained using either the Chib and Jeliazkov method (CJ; Bouranis et al., 2018; Chib & Jeliazkov, 2001) or the power posterior method (Friel et al., 2014; Friel & Pettitt, 2008).

The CJ method (Bouranis et al., 2018; Chib & Jeliazkov, 2001) estimates the marginal likelihood by evaluating the likelihood, prior, and posterior at a chosen parameter value, typically one with high posterior density to improve numerical accuracy. This approach uses output from the Metropolis–Hastings algorithm, where the parameter vector is updated in full conditional blocks. The core idea is to break down the posterior distribution into a sequence of conditional distributions, each of which can be estimated using the MCMC samples. By combining these estimates, the method provides a practical way to compute the marginal likelihood without needing to know the full normalizing constant of the posterior. On the other hand, the power posterior method (Friel et al., 2014; Friel & Pettitt, 2008) estimates the marginal likelihood by constructing a transition from the prior to the posterior distribution. This is achieved by raising the likelihood to a fractional power, referred to as the inverse temperature. When the inverse temperature is zero, the distribution corresponds to the prior, and when it is one, it corresponds to the posterior. In between, the likelihood is gradually introduced into the distribution, tempering its influence. The marginal likelihood is then approximated by integrating the expected deviance for each model across this continuum of tempered distributions. In practice, the range of inverse temperatures is discretized, and numerical integration, such as the trapezoidal rule, is used to estimate the marginal likelihood.

If we take the posterior model probabilities of two models and divide one by the other, the posterior odds (i.e., the relative plausibility of one model compared to the other after observing data) are expressed as the product of the ratio of their marginal likelihoods and the prior model odds (i.e., the relative plausibility of one model compared to the other model before observing data):

[1]Compared to the marginal likelihood notation used for Bayesian parameter estimation in Equation (2), the conditional notation here specifies the marginal likelihood given a specific model of interest.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\frac{p(\mathcal{M}_1|\boldsymbol{y})}{p(\mathcal{M}_2|\boldsymbol{y})} = \frac{p(\boldsymbol{y}|\mathcal{M}_1)}{p(\boldsymbol{y}|\mathcal{M}_2)} \times \frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)}. \tag{7}$$

The first term on the right side of Equation (7) is the Bayes factor (Kass & Raftery, 1995), which quantifies each model's evidence based on the observed data. More specifically, the Bayes factor evaluates the likelihood of the data being observed under two competing models. When the prior model odds are equal to each other (i.e., $p(\mathcal{M}_1) = p(\mathcal{M}_2)$), the Bayes factor is simply the posterior model odds.

3 | BAYESIAN MODEL AVERAGING FOR ERGMS

We have outlined the basic model definition of the ERGM and explained how Bayesian inference incorporates prior information for estimating parameters (Equation (2)) and selecting models (Equation (7)). Let us now consider a scenario in which researchers aim to simultaneously account for uncertainty in their ERGM parameters of interest and the multiple models under consideration. In this case, researchers can average parameter estimates across candidate models based on their posterior model probabilities using the BMA paradigm (Hoeting et al., 1999; Raftery et al., 1997). The conceptual details and philosophical foundations of BMA are thoroughly discussed in works such as Hoeting et al. (1999), Raftery et al. (1997), Hinne et al. (2020), Montgomery and Nyhan (2010), and van den Bergh et al. (2021). This section provides a summary of the theoretical background of BMA, its application to exponential random graph modelling, and the practical considerations researchers face in its implementation.

3.1 | Theoretical background

BMA can be used to estimate parameters or predict future observations by averaging over multiple candidate models. To distinguish the parameters being averaged across models from the full set of model-specific parameters $\boldsymbol{\theta}$, we use $\boldsymbol{\gamma}$ to represent the collection of model-averaged parameters in BMA. For estimation purposes, $\boldsymbol{\gamma}$ is obtained as follows:

$$p(\boldsymbol{\gamma}|\boldsymbol{y}) = \sum_{h=1}^{H} \left[p(\boldsymbol{\gamma}|\boldsymbol{y}, \mathcal{M}_h) \times p(\mathcal{M}_h|\boldsymbol{y}) \right], \tag{8}$$

meaning the model-averaged distribution of $\boldsymbol{\gamma}$ is obtained by first sampling one model from a total of H candidate models based on their respective posterior model probabilities, and then drawing one parameter value from the posterior distribution of the sampled model. Repeating this procedure results in the model-averaged distribution. Similar to the posterior distribution of a single model, researchers can summarize the model-averaged distribution using summary statistics such as the posterior mean, posterior standard deviation, and credible interval.

Although ERGMs specify the likelihood using a nonlinear exponential-family form—analogous to logistic regression models in that ERGMs model the log-odds of tie formation—BMA remains theoretically valid in this setting. In our approach, we implement formal BMA by averaging the full posterior distributions of the parameter vector $\boldsymbol{\gamma}$ across a set of candidate models. The parameter vector $\boldsymbol{\gamma}$ is defined over a shared parameter space, and each model is weighted by its posterior model probability, as shown in Equation (8). For models that exclude a particular network statistic, the corresponding coefficient is fixed at zero. This procedure results in a marginal model-averaged posterior distribution that accounts for model uncertainty. The BMA framework has been well established for other nonlinear models, such as generalized linear models (Hoeting et al., 1999; Li & Clyde, 2018). Therefore, despite the nonlinear nature of ERGMs, the resulting model-averaged posterior distributions remain valid and interpretable for inference.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

For the sake of predicting future observations, $\boldsymbol{\gamma}$ is replaced with $\hat{\boldsymbol{y}}$ to obtain the model-averaged predictive distribution:

$$p(\hat{\boldsymbol{y}}|\boldsymbol{y}) = \sum_{b=1}^{H} \left[p(\hat{\boldsymbol{y}}|\boldsymbol{y}, \mathcal{M}_b) \times p(\mathcal{M}_b|\boldsymbol{y}) \right]. \tag{9}$$

Here, $\hat{\boldsymbol{y}}$ is a prediction of $\boldsymbol{y}$, and $p(\hat{\boldsymbol{y}}|\boldsymbol{y}, \mathcal{M}_b)$ is the posterior predictive distribution. This formulation represents the Bayesian model-averaged predictive distribution, in which predictions from individual candidate models are integrated according to their posterior model probabilities. Specifically, posterior samples of model parameters are drawn for each model $\mathcal{M}_b$, and predictive draws of $\hat{\boldsymbol{y}}$ are subsequently generated conditional on those samples. Although we introduce this predictive formulation for completeness, our primary focus remains on the estimation aspect of BMA for ERGMs. Readers interested in the predictive aspect of BMA may consult Kaplan (2021).

When a number of candidate models are considered, the relevance of including a particular parameter of interest can be quantified by summing the posterior model probabilities of the candidate models that include it:

$$p(\text{incl}_{\boldsymbol{\gamma}}|\boldsymbol{y}) = \sum_{\mathcal{M}_b:\, \mathcal{M}_b \ni \boldsymbol{\gamma}} p(\mathcal{M}_b|\boldsymbol{y}). \tag{10}$$

This is termed the posterior inclusion probability, and it indicates the strength of evidence for including a specific parameter, based on the support provided by the data. It is also possible to quantify the evidence for excluding a specific parameter by summing the posterior model probabilities of the candidate models that exclude it. This is referred to as the posterior exclusion probability:

$$p(\text{excl}_{\boldsymbol{\gamma}}|\boldsymbol{y}) = \sum_{\mathcal{M}_b:\, \mathcal{M}_b \ni \dot{\boldsymbol{\gamma}}} p(\mathcal{M}_b|\boldsymbol{y}). \tag{11}$$

Using Equations (10) and (11), researchers can evaluate the plausibility of the observed network under ERGMs that include specific parameters compared to those that exclude them. This assessment helps determine whether an endogenous or exogenous predictor should be incorporated into the model:

$$\frac{p(\text{incl}_{\boldsymbol{\gamma}}|\boldsymbol{y})}{p(\text{excl}_{\boldsymbol{\gamma}}|\boldsymbol{y})} = \frac{p(\boldsymbol{y}|\text{incl}_{\boldsymbol{\gamma}})}{p(\boldsymbol{y}|\text{excl}_{\boldsymbol{\gamma}})} \times \frac{p(\text{incl}_{\boldsymbol{\gamma}})}{p(\text{excl}_{\boldsymbol{\gamma}})}. \tag{12}$$

On the left-hand side of Equation (12) are the posterior inclusion odds, while the first and second terms on the right-hand side refer to the inclusion Bayes factor and the prior inclusion odds, respectively. The inclusion Bayes factor represents the update from the prior inclusion odds to the posterior inclusion odds, indicating the support of the data for including a certain predictor compared to not including it, regardless of specific model selections. The prior inclusion odds consist of the prior inclusion probability in the numerator and the prior exclusion probability in the denominator, each calculated by summing the prior model probabilities that either include or exclude a certain parameter of interest.

3.2 | Practical considerations

The practical implementation of BMA for ERGMs requires addressing two considerations: selecting the pool of candidate models and assigning prior probabilities to these models.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.2.1 | How many candidate models?

In principle, all possible candidate models should be considered for BMA (e.g., Hinne et al., 2020; van den Bergh et al., 2021). For instance, in previous applications of BMA to psychometric network models (Hinne et al., 2020; Sekulovski et al., 2024), all possible network structures, defined by the presence or absence of connections between nodes, were evaluated. However, for ERGMs, pooling the full set of candidate models is neither straightforward nor realistic. This is because there is an effectively unlimited number of endogenous terms that can be included, and even when working with a relatively small set of nodal attributes, researchers can specify different effects to represent them, resulting in a vast model space. We therefore emphasize that theoretical considerations should take precedence over mere enumeration when identifying the pool of candidate models. That is, researchers should define the space of candidate models based on parameter configurations that are relevant to their theoretical interests, while excluding those that are not. Models lacking theoretical justification should be excluded before using BMA to assess the uncertainty surrounding the inclusion of models that are theoretically relevant (see Montgomery & Nyhan, 2010).

As for statistical considerations, it is important to note that Bayesian estimation of ERGMs can be halted if some candidate models in the candidate pool are radically misspecified and susceptible to model degeneracy (Handcock et al., 2003). In a similar vein, estimation issues may arise due to optimization challenges or collinearity, as reported in cases involving multiple endogenous terms (Duxbury, 2021; Lusher et al., 2013). These models should be carefully inspected by researchers to determine whether they are theoretically relevant; otherwise, they should be excluded from the analysis. Additionally, ERGMs cannot be fitted as empty models, so candidate models should at least include two terms, one of which is typically the 'edges' term for network density.

Considering these points, if a researcher selects p endogenous and/or exogenous terms to test hypotheses based on theory, the total number of candidate models in a pool would be 2^p. However, if this pool includes an empty model, a model containing only a density term, or candidate models with parameter configurations that do not accurately represent the theoretical model of interest, such models should be excluded. The final pool must include a sufficient number of models that reflect theoretical considerations, covering all plausible specifications while maintaining theoretical rigor and computational efficiency.

3.2.2 | How to decide prior model probabilities?

Assigning prior probabilities to each candidate model is another key consideration in BMA. As Hinne et al. (2020) noted, the results of BMA can depend on the choice of the prior model probabilities, making careful assignment crucial. In the absence of prior knowledge regarding the plausibility of the candidate models before observing data, it is typically assumed that all candidate models are equally likely a priori. This approach, often referred to as assigning a 'uniform prior' over the model space, ensures that each model is given equal probability. For instance, assuming there are H (e.g., 10) candidate models, each is assigned an equal probability of $\frac{1}{H}$ (e.g., $\frac{1}{10}$). This neutral choice of assigning equal probabilities due to limited prior knowledge is described by Hoeting et al. (1999).

If some candidate models are more plausible a priori, these models should receive more weight than others. To illustrate, consider the use of multiple ERGMs to model a directed friendship network. If theoretical considerations or empirical evidence from previous studies support models with reciprocity terms, candidate models that include reciprocity terms should receive higher prior probabilities, while those without them should receive lower probabilities. The degree to which each weight is adjusted depends on the researchers' decisions. In such cases, it is advisable to test different prior model probability settings and examine the sensitivity of the results. The prior model probabilities over the candidate

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

models must sum to 1. Given this requirement, researchers can allocate differing degrees of prior belief based on expert knowledge, findings from previous empirical studies, or substantive theory.

4 | EMPIRICAL ANALYSIS OF A COLLEGE FRIENDSHIP NETWORK

Having outlined the theoretical and practical considerations of using BMA for ERGMs, we now demonstrate its application through an analysis of an empirical friendship network in a college setting. The data were collected in 2017 by the Lab for Big Data Methodology at the University of Notre Dame and consist of friendship ties among fourth-year college students, along with nodal-level characteristics including demographic, behavioural, and psychological variables. This dataset has previously been used in other forms of social network models, such as latent space models (Liu et al., 2021; Liu & Zhang, 2021), and is particularly relevant to psychological research due to its inclusion of psychological trait measures. For clarity of presentation, interpretation, and visualization, we focus on a subsample of 90 students from the full dataset in the present analysis. This subsample was selected because it exhibited the clearest class-based clustering pattern, making it informative for illustrating the application of social network modelling and BMA for ERGMs. The data and R code used to produce the results are available on the Open Science Framework (OSF) at https://osf.io/g9eq4/.

4.1 | Research scenario

As emphasized in the 'Practical Considerations' subsection, theoretically grounded models should be prioritized over those solely selected for computational feasibility. In this section, we present a hypothetical research scenario to illustrate the application of our approach. Specifically, we consider a case in which a researcher aims to explain and predict the structure of college friendships based on substantive theory and findings from previous empirical studies.

First, the researcher incorporates a structural effect to capture the tendency for triadic closure in the college friendship network—a common feature of social networks in which actors connected to a mutual third party are more likely to form ties with each other. In the context of the hypothetical scenario, this tendency is interpreted through the lens of balance theory (e.g., Cai et al., 2024; Heider, 1958; Pircalabelu & Claeskens, 2016), which postulates that social actors strive for cognitive consistency in their relational structures. Specifically, when students A and B, and B and C, are connected, the likelihood of a tie forming between A and C increases. The absence of such a tie produces cognitive dissonance, an uncomfortable psychological state that actors are motivated to resolve by forming a new connection, thereby forming and completing a balanced triad. To formalize this mechanism in the ERGM, the researcher includes the geometrically weighted edgewise shared partner (GWESP; see Hunter, 2007; Snijders et al., 2006) term.

The GWESP term is a widely used specification for modelling triadic closure. It captures the tendency for a tie to become more likely between two actors as they share more mutual partners, reflecting the principle that 'friends of friends' are more likely to become connected. Compared to other standard triad count terms, the GWESP term is often preferred because it mitigates issues such as model oversaturation and instability. This is achieved through a decay parameter that reduces the influence of ties involving a very large number of shared partners. As a result, the contribution of each additional shared partner decreases progressively, preventing triangle counts from escalating uncontrollably and improving model stability. For further details, we refer readers to Cranmer et al. (2020), Hunter (2007), and Snijders et al. (2006).

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In addition to the endogenous variables, the researcher considers the role of three exogenous covariates in tie formation within the friendship network: two personality traits from the Big Five—extraversion and openness to experience—and a behavioural indicator of smoking status (i.e., whether the student smokes or not).

Starting with the two personality traits, the researcher includes 'extraversion' and 'openness' as exogenous covariates guided by social psychological theory and prior empirical research. Extraverted individuals are generally more social, talkative, assertive, and active than those with lower extraversion (Roccas et al., 2002). Accordingly, extraverted college students are hypothesized to initiate more interactions, increasing their likelihood of forming friendship ties within the network (Rubin et al., 2006; Selfhout et al., 2010; Wagner et al., 2014). Similarly, college students high in openness to experience tend to be more intellectual, imaginative, sensitive, and open-minded than those low in openness (Roccas et al., 2002). These qualities may facilitate engagement with diverse peers and encourage exploration of unfamiliar social contexts, thereby fostering a broader set of connections (Wagner et al., 2014). Both personality effects are modeled using 'nodecov' terms, which test whether individuals with higher values on these traits are more likely to be involved in a greater number of friendship ties.

The third exogenous covariate, smoking, reflects a visible behavioural marker that is theorized to function as a signal of social status within college peer culture. The researcher includes this effect based on sociological theories of peer status hierarchies and prior empirical research. Individuals who smoke are often perceived as more rebellious, autonomous, or socially confident, which may enhance their social status (in terms of popularity) within peer networks (Lakon et al., 2015). Although the relationship between smoking and status is likely reciprocal, several studies suggest that smoking can function as a status signal, enhancing individuals' appeal for friendship selection within peer groups (Lakon et al., 2015; Schaefer et al., 2012). Accordingly, smokers may attract more friendship ties because they occupy a high-status position within the group. This status-related mechanism is modeled using a 'nodecov' term, which captures whether individuals who smoke are more likely to be involved in a greater number of friendship ties.

Importantly, the researcher remains uncertain about the theoretical importance of each exogenous effect for understanding the true network-generating process. To address this uncertainty, the researcher evaluates model uncertainty associated with including these three exogenous covariates—extraversion, openness, and smoking—within the ERGM framework for modelling friendship formation. Models that exclude all three covariates are not considered, as the focus is on assessing the relative contribution of each covariate rather than their collective absence.

Lastly, the researcher includes gender and whether the students attended the same college class as control variables in the ERGM. These effects are specified using the 'nodematch' term, as an operationalization of the homophily principle, that is, as a stable individual psychological preference for forming friendships with others who are similar to oneself (Kossinets & Watts, 2009; McPherson et al., 2001). This preference is theorized to stem from a motivation to reduce subjective uncertainty in social interactions (Hogg, 2000). Similarity in attributes, values, and experiences enhances interpersonal predictability, facilitates smoother communication, and fosters mutual understanding (Festinger, 1957). Gender-based homophily in friendship networks has been shown to be particularly robust, with same-gender friendships being more common than would be expected at random (McPherson et al., 2001). Likewise, class-based homophily reflects the tendency for students who attended the same class to form friendship ties, as increased opportunities for interaction often lead to greater familiarity and connection. Specifically, this expectation derives from Feld's theory on the focused organization of social ties (Feld, 1981), which emphasizes that social foci like college classrooms structure interaction and foster relationship formation. Including these controls through 'nodematch' terms allows the model to account for background-driven, similarity-based friendship formation mechanisms that might otherwise confound the interpretation of the theoretical focal endogenous and exogenous effects.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

This hypothetical research scenario reflects a typical ERGM specification process in psychological research. The researcher selects theoretically relevant endogenous, exogenous, and control variables, represents them using appropriate model terms to capture the underlying network-generating process, and evaluates these models using BMA for ERGMs to address potential model misspecification and model uncertainty.

4.2 | Data description

Having formulated the theoretical expectations, the researcher begins by examining the descriptive statistics of the data, concentrating first on the network characteristics and then on the nodal attributes. The college friendship network is modeled as an undirected graph, where nodes represent individual students and edges denote mutual friendship ties. In the original dataset, friendship ties are valued on a scale from 0 to 5. The specific meaning of each level on the friendship scale is provided in Table 1. For the purpose of ERGM analysis, these valued ties were dichotomized into a binary format: Ties with values of 3 ('The person is a friend of mine.') or 4 ('The person is one of my best friends.') were recoded as 1, indicating the presence of a friendship, while ties with values of 0, 1 or 2 (ranging from 0 = 'I have never heard the name' to 2 = 'I have met the person a few times, but he/she is not a friend of mine',) were retained as 0. This binarization thus assumes that all retained ties reflect equally significant friendship. The following syntax demonstrates how the necessary packages and data were loaded for the analysis.

```
# Load necessary packages
library(Bergm)      # Fit Bayesian ERGMs
library(igraph)     # Network descriptives and visualization
library(tidyverse)  # Data management
library(parallel)   # Parallel processing

# Load the friendship network data
FriendNet <- read.csv("FriendNet.csv", header = FALSE)
FriendNet <- as.matrix(FriendNet)
FriendNet <- as.network(FriendNet, directed = FALSE, loops = FALSE)

# Load and clean the attribute data
FriendAtt <- read.csv("FriendAtt.csv")
FriendAtt <- FriendAtt[,-1]  # remove ID column

# Assign attributes to network
FriendNet %v% "gender" <- FriendAtt[, "gender"]
FriendNet %v% "class" <- FriendAtt[, "class"]
FriendNet %v% "smoke" <- FriendAtt[, "smoke"]
FriendNet %v% "ext" <- FriendAtt[, "ext"]
FriendNet %v% "opn" <- FriendAtt[, "opn"]
```

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Five-point scale for coding relationships in the college friendship network.

Level	Interpretation
0	I have never heard the name
1	I heard about the person but had no personal interaction with her/him
2	I have met the person a few times, but he/she is not a friend of mine
3	The person is a friend of mine
4	The person is one of my best friends

4.2.1 | Network characteristics

The first step to understanding the structure of the college friendship network is to examine its network characteristics. This is done using the following syntax:

```
# Convert to igraph object
FriendNet_igraph <- intergraph::asIgraph(FriendNet)

# Create a dataframe of vertex attributes
vertex_attributes <- data.frame(
  gender = V(FriendNet_igraph)$gender,
  class = V(FriendNet_igraph)$class,
  smoke = V(FriendNet_igraph)$smoke,
  ext = V(FriendNet_igraph)$ext,
  opn = V(FriendNet_igraph)$opn
)

# Compute degree centrality
vertex_attributes$DegreeCentrality <- degree(FriendNet_igraph, mode = "all
    ")

# Display and sort by degree centrality
vertex_attributes %>%
  arrange(desc(DegreeCentrality)) %>%
  print()

# General summary information
summary(FriendNet)
```

The college friendship network consists of 90 nodes (i.e., college students) and 1160 edges (i.e., friendships), with no missing ties. The network density, defined as the proportion of realized edges to the total number of potential edges, is approximately .290, indicating a moderately connected network. Degree centrality, which quantifies the number of direct connections a node has, is particularly relevant for identifying influential individuals within the network. Students with high degree centrality may serve as social hubs, playing important roles in fostering communication, maintaining group cohesion, and facilitating the flow of information within the network. Visual representations of the network and degree centrality are provided in Figures 1 and 2.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

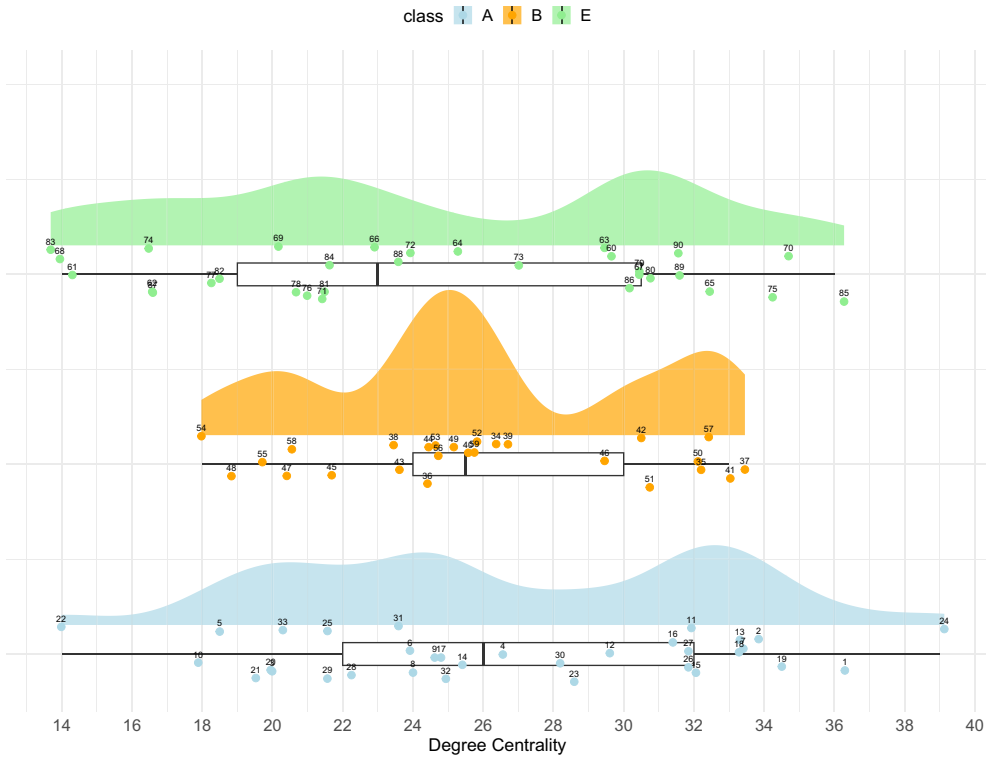

FIGURE 1 College friendship network of students included in the empirical analysis, visualized by class and degree centrality.

Figure 1 presents a raincloud plot of degree centrality by college class membership. Each distribution combines a raw data jitter plot to represent individual students, a density estimate to show the distribution of degree centrality values, and a boxplot to indicate variability, altogether allowing for both individual-level and group-level inspection of connectivity. Students in Class *A* (bottom; colored light blue) generally have higher degree centrality based on the median (i.e., the second quartile of the boxplot). Students in Class *B* (middle; orange) have a lower median, and students in Class *E* (top; light green) show the lowest median centrality. Students in Classes *A* and *E* show similar levels of variability (as indicated by the range from minimum to maximum in the boxplot), whereas Class *B* displays less variability in degree centrality. This suggests that in Class *B*, students are more similar to one another in terms of their number of connections, while Classes *A* and *E* exhibit a wider range of social connectivity, indicating a more unequal distribution of social ties within those classes. In each class, students 24, 37, and 85 exhibit the highest degree centrality, indicating that they serve as social hubs within their respective classes. At a minimum, all students are connected, as there are no isolates in the network.

Figure 2 visually represents the college friendship network of students. Nodes are colored according to class membership, and their sizes are scaled to reflect each student's degree centrality, with larger nodes indicating greater connectivity. The visualization reveals noticeable clustering by class, particularly among students in Class *A* (located at the bottom right and colored light blue), Class *B* (located on the left and colored orange), and Class *E* (located at the upper right and colored light green), suggesting that class membership plays an important role in shaping friendship formation. Notably, the students with the highest degree centrality in each class—students 24 (Class *A*), 37 (Class *B*), and 85 (Class *E*)—are also connected to peers outside their own class, indicating their potential role as cross-class social hubs. These students are likely to play central roles in facilitating information flow and fostering social cohesion within and across class boundaries. In contrast, some students occupy more peripheral positions in the network, such as student 22 in Class *A*, student 54 in Class *B*, and

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

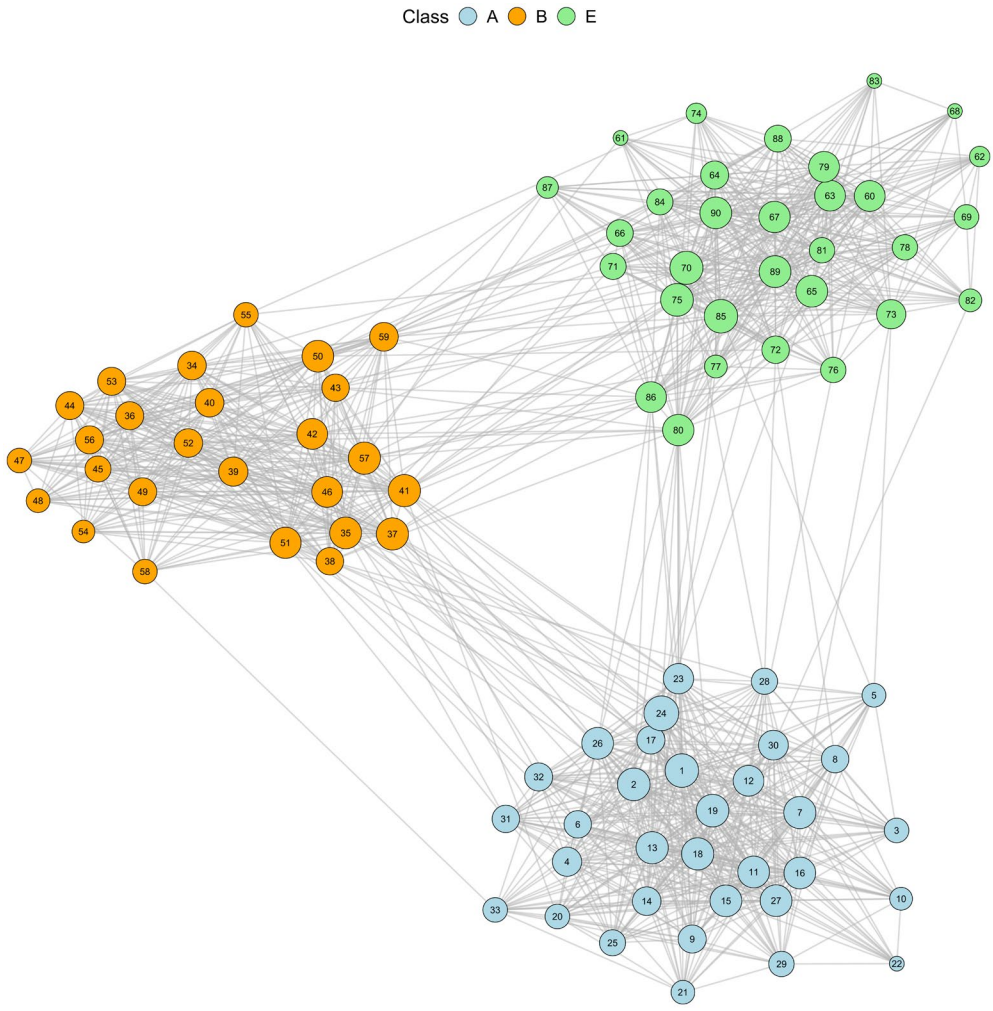

FIGURE 2 Network visualization of college friendship.

student 83 in Class *E*. These students may be less integrated into the broader social structure of the cohort, possibly reflecting either individual preferences or structural barriers to connection.

4.2.2 | Nodal attributes

The nodal attributes, in turn, provide psychological, behavioural, demographic, and group-based information about each student in the college friendship network. Specifically, the dataset includes two Big Five psychological personality traits—extraversion and openness—smoking status (behaviour), gender (demographic), and classroom membership (group-based). There were no missing values in the nodal attributes.[2] Descriptive statistics for these variables are generated using the following syntax.

[2]As should now be clear, there were no missing values in either the tie data or the nodal attribute data. Nonetheless, we acknowledge that missing data can pose challenges for researchers. In Bayesian analyses of ERGMs, missing values can be handled through data augmentation, whereby missing entries are 'filled in' as part of the estimation process (Koskinen et al., 2010). Readers interested in methods for handling missing data in ERGMs are referred to Koskinen et al. (2010) and Krause et al. (2020) for further information.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

```r
# Summary statistics
## Continuous variables
vertex_summary <- vertex_attributes %>%
  summarise(
    Ext_Mean = mean(ext, na.rm = TRUE),
    Ext_SD = sd(ext, na.rm = TRUE),
    Ext_Min = min(ext, na.rm = TRUE),
    Ext_Max = max(ext, na.rm = TRUE),
    Opn_Mean = mean(opn, na.rm = TRUE),
    Opn_SD = sd(opn, na.rm = TRUE),
    Opn_Min = min(opn, na.rm = TRUE),
    Opn_Max = max(opn, na.rm = TRUE)
  ) %>%
  print()

# Label categorical variables
vertex_attributes <- vertex_attributes %>%
  mutate(
    Gender = factor(gender, levels = c(0, 1), labels = c("Male", "Female")
   ),
    Smoke = factor(smoke, levels = c(0, 1), labels = c("Smoke_No", "Smoke_
   Yes")),
    Class = as.factor(class)
  )

# Categorical variable counts
vertex_attributes %>%
  count(Smoke) %>%
  rename(Count = n)

vertex_attributes %>%
  count(Gender) %>%
  rename(Count = n) %>%
  print()

vertex_attributes %>%
  count(Class) %>%
  rename(Count = n) %>%
  print()
```

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Extraversion

The dataset includes two dimensions of the Big Five personality traits. The first is extraversion, measured using four items from the 20-item Mini-IPIP Scale for the Big Five personality factors (Donnellan et al., 2006). Factor scores for extraversion (variable name: `ext`) were derived from these items following procedures outlined in Liu et al. (2021, 2018). Scores range from −2.381 to 1.269, with a mean of −.591. Higher scores indicate greater levels of extraversion.

Openness to experience

The second personality trait is openness to experience, measured using four items related to openness from the same Mini-IPIP Scale (Donnellan et al., 2006). Factor scores for openness (variable name: `opn`) were also derived following the method used in Liu et al. (2021, 2018). Openness scores ranged from −.926 to .829, with a mean of .016. Higher values correspond to higher levels of openness.

Smoking status

A health-related behaviour variable included in the dataset is smoking status of the respondent (variable name: `smoke`). Of the 90 students, 70 reported that they do not smoke (coded as 0), while the remaining 20 identified as smokers (coded as 1).

Gender

One demographic variable included in the dataset is student gender (variable name: `gender`). Among the 90 students, 39 identified as male (coded as 0) and 51 as female (coded as 1).

Class

The dataset also includes information about each student's classroom in college (variable name: `class`). As such, this variable reflects each student's group affiliation within the college context. Students in the same college class participated in many shared activities, which served as a foundation for forming friendships. Among the 90 students included in the analysis, 33 were from Class *A*, 26 from Class *B*, and 31 from Class *E*.

4.3 | Analytic details

To estimate the parameters of the candidate ERGMs, we specified endogenous terms, exogenous terms, and control variables. When fitting all models in the candidate pool, we followed standard ERGM practice by including the 'edges' term in every candidate model to represent the network density (Cranmer et al., 2020). In the code block, this is implemented via the `edges` specification in the `fixed_terms` object. In addition, the GWESP term was included as an endogenous structural term using `gwesp`. For this term, we used a decay parameter fixed at .5 to model the tendency toward triad closure (Caimo & Friel, 2014; Pircalabelu & Claeskens, 2016). The three exogenous nodal covariates (extraversion, openness, and smoking status) were evaluated using different configurations across candidate models. These covariates reflect, respectively, the student's level of extraversion, level of openness to experience, and whether they smoke. Each covariate was included using the `nodecov` term: `nodecov('ext')`, `nodecov('opn')`, and `nodecov('smoke')`. The two control variables, gender and class membership, were included via `nodematch('gender')` and `nodematch('class')`. The `nodematch` term captures homophily effects (i.e., the tendency for similar individuals to form ties) by modelling the increased likelihood of tie formation between students who share the same attribute, such as gender or class.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Unlike the fixed terms, the covariates varied across models. To enumerate all combinations, these nodal covariates were specified as `optional_terms` in the code. We considered only models that included at least one of the three covariates. Models containing only the endogenous terms and control variables (i.e., `edges`, `gwesp`, `gender`, and `class`) were excluded to reflect a scenario in which the researcher is specifically interested in the covariate effects. This process resulted in a total of seven candidate models (`model_formulas`).

To formalize the structure of each candidate model, we denote the probabilistic formulation as follows:

$$\mathcal{M}_b: p(\boldsymbol{Y}=\boldsymbol{y}|\boldsymbol{\theta}_b) \propto \exp\{\theta_{b,1}\cdot \texttt{edges}(\boldsymbol{y}) + \theta_{b,2}\cdot \texttt{gwesp}(\boldsymbol{y}) + \theta_{b,3}\cdot \texttt{class}(\boldsymbol{y}) + \theta_{b,4}\cdot \texttt{gender}(\boldsymbol{y}) + \sum_{k\in\mathcal{S}_b} \theta_{b,k}\cdot \texttt{nodecov}_k(\boldsymbol{y})\},$$

where $\mathcal{S}_b \subseteq \{\texttt{smoke}, \texttt{ext}, \texttt{opn}\}$ indexes the optional nodal covariates included in model $\mathcal{M}_b$, with the constraint that $\mathcal{S}_b \neq \emptyset$. The R code below implements this model specification logic by (1) defining the fixed endogenous terms and control variables (edges, gwesp, class, and gender), (2) enumerating all possible combinations of the exogenous nodal covariates, and (3) constructing the corresponding model formulas.

```
# Fixed terms
fixed_terms <-
  "edges + gwesp(0.5, fixed = TRUE) + nodematch('class') + nodematch('
   gender')"

# Optional terms
optional_terms <- c(
  "nodecov('smoke')",
  "nodecov('ext')",
  "nodecov('opn')"
)

# Generate all combinations of optional terms
all_combinations <- lapply(1:length(optional_terms), function(k) {
  combn(optional_terms, k, simplify = FALSE)
}) %>% unlist(recursive = FALSE)

# Create model formulas
model_formulas <- lapply(all_combinations, function(terms) {
  as.formula(paste("FriendNet ~", fixed_terms, "+", paste(terms, collapse
   = " + ")))
})
```

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

For this set, we assigned equal prior model probabilities (i.e., $p(\mathcal{M}_1) = p(\mathcal{M}_2) = \ldots = p(\mathcal{M}_7) = \frac{1}{7}$), assuming no prior preference for any particular model (i.e., a uniform prior).

We used the R `Bergm` package (Caimo & Friel, 2013, 2014) to fit the models. The main function used was `bergm()`. For the MCMC algorithm, the number of chains was set to twice the number of parameters being estimated (e.g., 10 chains for a model with 5 parameters), which is the default setting in the `Bergm` package. Each chain included 4000 posterior draws (`main.iters = 4000`), after discarding the first 1000 iterations as burn-in (`burn.in = 1000`), and used 2000 auxiliary iterations (`aux.iters = 2000`) with a default scaling factor of .6 (`gamma = .6`) for the parallel adaptive direction sampler (Caimo & Friel, 2013, 2014). The auxiliary iterations refer to the number of MCMC steps used in the internal likelihood simulation (i.e., the auxiliary chain), which is required by the exchange algorithm. Since the ERGM likelihood is doubly intractable, this internal MCMC chain is used to simulate a network for each proposed parameter value. For further technical details, see Caimo et al. (2022); for software implementation, see Caimo and Friel (2014).

The prior distribution for model coefficients was specified as a multivariate normal with a zero mean vector (`M.prior`) and a diagonal covariance matrix with elements set to 4 (`S.prior`), following recommendations for weakly informative priors (Krause et al., 2020). The `V.proposal` parameter controls the variance of the multivariate normal proposal distribution $h(\boldsymbol{\theta}'|\boldsymbol{\theta})$ and was set to the `Bergm` package default of .0025. The number of posterior draws was determined by assessing convergence and ensuring parameter stability across chain lengths. The R code below fits each candidate model using the `bergm()` function. Parallel processing is used to expedite computation by utilizing all but two cores.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

```
# Detect the number of cores and create a cluster
cl <- makeCluster(detectCores() - 2)  # Use all but two cores

# Load required packages in each worker node
clusterEvalQ(cl, library(Bergm))
clusterEvalQ(cl, library(statnet))

# Export necessary objects to worker nodes
clusterExport(cl, varlist = c("model_formulas", "FriendNet"))

fit_model <- function(formula) {
  # Count the number of terms in the model
  num_terms <- length(attr(terms(formula), "term.labels"))

  # Dynamically create priors
  M.prior <- rep(0, num_terms)  # Mean vector
  S.prior <- diag(4, num_terms) # Covariance matrix

  # Fit the model
  bergm(
    formula = formula,
    prior.mean = M.prior,
    prior.sigma = S.prior,
    main.iters = 4000,
    burn.in = 1000,
    aux.iters = 2000,
    gamma = 0.6,
    V.proposal = 0.0025
  )
}

# Fit models in parallel
model_results <- parLapply(cl, model_formulas, fit_model)

# Stop the cluster
stopCluster(cl)
```

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

To evaluate model evidence for the purpose of model averaging, we computed the marginal likelihood separately from the posterior sampling process, as the current implementation of the `Bergm` package does not support simultaneous estimation. To obtain the marginal likelihood, we used the CJ method based on adjusted pseudolikelihood, implemented via the `evidenceCJ()` function (Bouranis et al., 2018; Caimo & Friel, 2013, 2014). This required an additional MCMC chain of 4000 iterations (`main.iters = 4000`), including a burn-in of 1000 iterations (`burn.in = 1000`) and 2000 auxiliary iterations (`aux.iters = 2000`). We used 3000 post-burn-in samples from this chain to calculate the log marginal likelihood (`num.samples = 3000`). For the `V.proposal` and `ladder` arguments, we used values of 1.5 and 30, respectively, following the default settings. The `ladder` argument controls the adjustment of the ERGM pseudolikelihood (see Bouranis et al., 2018). Prior distributions for the parameters were based on the multivariate normal distribution with a mean vector of zeros (`M.prior`) and a diagonal covariance matrix with elements of 4 (`S.prior`). As demonstrated previously, setting 4 as the elements of the diagonal covariance matrix is intended to specify a weakly informative prior that is generally preferred for ERGMs (Krause et al., 2020). The R code below computes the log marginal likelihoods for each candidate model using the `evidenceCJ()` function. As in the model fitting stage, parallel processing is used to accelerate computation by distributing the tasks across multiple cores.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

```
# Detect the number of cores and create a cluster
cl <- makeCluster(detectCores() - 2)  # Use all but two cores

# Export necessary objects to worker nodes
clusterEvalQ(cl, library(Bergm))  # Ensure required package is loaded
clusterExport(cl, varlist = c("model_formulas", "FriendNet"))  # Export
    candidate models

# Define the function to compute log marginal likelihood
compute_log_marginal <- function(formula) {
  # Count the number of terms in the model
  num_terms <- length(attr(terms(formula), "term.labels"))

  # Dynamically create priors
  M.prior <- rep(0, num_terms)        # Mean vector
  S.prior <- diag(4, num_terms)     # Covariance matrix (updated
   dynamically)

  # Fit the model and compute evidence
  result <- evidenceCJ(
    formula = formula,
    prior.mean = M.prior,
    prior.sigma = S.prior,
    main.iters = 4000,
    burn.in = 1000,
    aux.iters = 2000,
    num.samples = 3000,
    V.proposal = 1.5,
    ladder = 30
  )

  # Extract and return log marginal likelihood
  return(result$log.evidence)
}

# Compute log marginal likelihoods in parallel
log_marginal_likelihoods <- parLapply(cl, model_formulas, compute_log_
    marginal)

# Stop the cluster
stopCluster(cl)
```

Using parameter estimates and log marginal likelihood estimates, we compute Bayesian model-averaged parameter estimates following Equation (8). To illustrate, we simulate from the model-averaged posterior distribution $p(\boldsymbol{\gamma}|\boldsymbol{y})$ by first sampling a model $\mathcal{M}_h$ from the candidate set $\{\mathcal{M}_1, \ldots, \mathcal{M}_H\}$ in

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

proportion to its posterior model probability $p(\mathcal{M}_b|\boldsymbol{y})$, and then drawing the corresponding parameter estimate γ from the posterior samples under the selected model:

$$\gamma^{(i)} \sim p(\boldsymbol{\gamma}|\boldsymbol{y}, \mathcal{M}_b), \quad \text{where} \quad \mathcal{M}_b \sim p(\mathcal{M}_b|\boldsymbol{y}).$$

This simulation procedure approximates the right-hand side of Equation (8) through repeated resampling and is shown in the R code below. Specifically, for a parameter of interest (indexed by `param_name`), the function repeatedly samples one model according to its posterior probability (i.e., `post_prob`) and checks whether the sampled model includes the target parameter (i.e., whether `param_name` appears in the posterior sample matrix's column names). If the parameter is present, the function randomly draws one value from its posterior distribution; if it is not included in the sampled model, a value of zero is recorded. Repeating this process generates posterior draws from the model-averaged distribution of the parameter.

```r
# General function for Bayesian model averaging
bma_posterior <- function(num.iter, model_estimates, param_name, post_prob
    ) {
  # Initialize output matrix
  param_matrix <- matrix(NA, num.iter, 1)

  # Sample model indices according to posterior model probabilities
  sampled_models <- sample(seq_along(model_estimates), num.iter, replace =
    TRUE, prob = post_prob)

  for (i in seq_len(num.iter)) {
    mod <- sampled_models[i]  # model index
    mat <- model_estimates[[mod]]  # posterior sample matrix for the
   selected model

    if (!is.null(mat) && param_name %in% colnames(mat)) {
      # If the parameter exists in this model, sample one value from that
   column
      col_idx <- which(colnames(mat) == param_name)
      row_idx <- sample(1:nrow(mat), 1)
      param_matrix[i, 1] <- mat[row_idx, col_idx]
    } else {
      # If the parameter is not in this model, assign 0
      param_matrix[i, 1] <- 0
    }
  }

  return(param_matrix)
}
```

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4.4 | BMA results

In presenting the results, we briefly describe the computational aspects of the BMA procedure. For this empirical application using a subsample of 90 students, the full estimation process, comprising candidate model estimation and log marginal likelihood computation, took approximately 13 min and 9 s. These analyses were conducted on a machine equipped with an Apple M1 chip (16 GB RAM), macOS Sequoia Version 15.5, utilizing 6 cores for parallel processing. While the runtime reported here is representative, it may vary depending on other factors (e.g., the number of iterations specified for posterior simulation and marginal likelihood estimation). Nevertheless, the results demonstrate that the proposed method is computationally feasible in practical settings.

We note that, from this point on, we use the term attribute effects to refer to the 'nodecov' effects of extraversion, openness, and smoking status. First, none of the candidate models faced any estimability issues (e.g., multicollinearity, optimization, or degeneracy). Table 2 subsequently presents all seven candidate models in terms of their prior model probabilities, posterior model probabilities, and Bayes factors. The first column specifies which terms are included in each candidate model (Models 1 through 7, denoted M1–M7). The second column lists the prior model probabilities, calculated as $\frac{1}{7}$, reflecting equal prior weight for each candidate model. The third column provides the posterior model probabilities obtained after BMA. The fourth column reports the Bayes factors.

Inspection of Table 2 shows that the first three models (M1, M2, and M3) accounted for 97% of the total posterior model probability, indicating much stronger support from the data compared to the remaining candidate models. Notably, the posterior probability of the best model (M1) was .600, suggesting that while it was clearly favoured, it did not overwhelmingly dominate the model space. Model 1, which includes all three attribute effects, outperformed all other candidate models in terms of posterior model probability.

When only the openness to experience effect was excluded in Model 2, the posterior probability dropped but still remained at a reasonable level (.333). However, in subsequent models that excluded the extraversion or smoking effects, the posterior probability decreased substantially. For example, in Model 3, which excludes both the openness and smoking effects, the posterior probability dropped to .041.

The Bayes factor (BF_{1r}) represents the relative predictive performance of the best model compared to the model in row *r*. A similar pattern to that observed in the posterior probabilities emerges here. For example, the Bayes factor $BF_{12} = 1.803$ indicates that Model 1 (which includes all three attribute effects) was 1.803 times better than Model 2 (which includes the extraversion and smoking but not the openness effect) in explaining the data. In contrast, $BF_{14} = 22.433$ for Model 4—which includes the

TABLE 2 Prior/posterior model probabilities and Bayes factors.

Model	$p(\mathcal{M})$	$p(\mathcal{M} \mid \text{data})$	BF_{1r}
M1: edges + gwesp + extraversion + openness + smoke + gender + class	.143	.600	1.000
M2: edges + gwesp + extraversion + smoke + gender + class	.143	.333	1.803
M3: edges + gwesp + extraversion + gender + class	.143	.041	14.683
M4: edges + gwesp + extraversion + openness + gender + class	.143	.027	22.433
M5: edges + gwesp + openness + gender + class	.143	6.399×10^{-10}	9.372×10^{8}
M6: edges + gwesp + smoke + gender + class	.143	3.472×10^{-12}	1.727×10^{11}
M7: edges + gwesp + openness + smoke + gender + class	.143	1.240×10^{-20}	4.838×10^{19}

Note: gwesp refers to the geometrically weighted edgewise shared partner. extraversion and openness are personality trait scores reflecting students' levels of extraversion and openness to experience, respectively. smoke is a binary indicator for smoking status (0 = does not smoke, 1 = smokes). The effects of these three exogenous covariates were modeled using the nodecov term. gender is a binary indicator for student gender (0 = male, 1 = female). class indicates classroom membership. The effects of these two control variables were modeled using the nodematch term.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

extraversion and openness but not the smoking effect—shows that Model 1 explained the data 22.433 times better than Model 4.

Table 3 displays the evidence of parameter inclusion and summary statistics based on the model-averaged posterior distribution. According to the posterior inclusion probabilities, all three attribute effects showed stronger evidence for inclusion than expected under the prior, with posterior probabilities exceeding the prior value of .571. The extraversion effect had the highest posterior inclusion probability (1.000) and an extremely large inclusion Bayes factor of 1.166×10^9, indicating that, across all candidate models, those including this effect were vastly more likely than those excluding it. This finding therefore provides decisive support for the inclusion of extraversion in predicting friendship formation in the college network. The openness effect showed weaker evidence, with a posterior inclusion probability of .626 and an inclusion Bayes factor of 1.258. Although the posterior inclusion probability exceeded the prior, the Bayes factor suggests minimal support for including this attribute effect. The smoking status effect had a posterior inclusion probability of .932 and an inclusion Bayes factor of 10.348, providing substantial support for its inclusion. Overall, while all three attribute effects were supported to some extent, the results suggest that the extraversion and smoking status effects are more predictive of social tie formation in the network than the openness effect.

Summary statistics in Table 3 provide information about the size and direction of the parameters associated with the three attribute effects on friendship in the college network. The coefficient of .325 for extraversion indicates the mean log-odds of friendship formation associated with being more extroverted, after averaging over candidate models and controlling for other covariates. The exponentiated value, $e^{0.325} \approx 1.384$, indicates that with each unit increase in extraversion, the likelihood of forming a friendship increased by approximately 38.4%, on average, after adjusting for other variables. The coefficient for openness was −.017. The exponentiated value, $e^{-0.017} \approx 0.983$, implies a slight decrease (about 1.7%) in the likelihood of forming a friendship tie with each unit increase in openness, holding other terms constant. Finally, the coefficient for smoking status was .173. The exponentiated value, $e^{0.173} \approx 1.189$, suggests that, on average, students who smoke were about 18.9% more likely to form friendship ties compared to non-smokers, holding all else constant.

In sum, the mean coefficients in Table 3 indicate positive associations for extraversion and smoking status with friendship formation, and a small negative association for openness, with extraversion showing the strongest effect, followed by smoking status and openness. The strength of the evidence also varies across the three attribute effects. Extraversion and smoking status both have posterior inclusion probabilities close to 1 and large inclusion Bayes factors (1.166×10^9 and 10.348, respectively), suggesting substantial to decisive evidence for inclusion. In contrast, openness has a smaller coefficient and a weaker inclusion Bayes factor (1.258), indicating only minimal support for inclusion. This pattern reflects differential levels of uncertainty about each attribute's relevance when averaged across the candidate models. We note that including a 'null' candidate model with no attribute effects can serve as a useful baseline when the research question considers the possibility that none of the attributes are relevant. However, in our empirical analysis, we excluded such a model because our focus was to evaluate which of the three attributes might matter, under the assumption that at least one was theoretically meaningful.

TABLE 3 Summary of model-averaged posterior distribution and evidence of parameter inclusion.

Parameter	*M*	SD	*p*(incl)	*p*(incl|*y*)	BF_{incl}
Network ~ extraversion	.325	.108	.571	1.000	1.166×10^9
Network ~ openness	−.017	.152	.571	.626	1.258
Network ~ smoke	.173	.166	.571	.932	10.348

Note: *M* and SD refer to the mean and standard deviation of the model-averaged posterior distribution for the parameter, respectively. *p*(incl) refers to prior inclusion probability. *p*(incl|***y***) refers to posterior inclusion probability. BF_{incl} refers to the inclusion Bayes factor. network refers to the college friendship network. extraversion and openness are personality trait scores reflecting students' levels of extraversion and openness to experience, respectively. smoke is a binary indicator for smoking status (0 = does not smoke, 1 = smokes). The effects of these three exogenous covariates were modelled using the nodecov term.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 Parameter estimates for each candidate model.

Parameter	M1	M2	M3	M4	M5	M6	M7
Edges	−5.697	−5.629	−5.567	−5.554	−6.074	−6.150	−6.176
Gwesp	1.321	1.291	1.319	1.318	1.437	1.408	1.409
Extraversion	.323	.325	.335	.347			
Openness	−.028			−.057	.220		.230
Smoke	.189	.195				.249	.263
Gender	.949	.925	.927	.931	.909	.936	.924
Class	4.853	4.823	4.812	4.820	4.689	4.708	4.732

Note: M1 through M7 correspond to the candidate model specifications listed in Table 2. gwesp refers to the geometrically weighted edgewise shared partner. extraversion and openness are personality trait scores reflecting students' levels of extraversion and openness to experience, respectively. smoke is a binary indicator for smoking status (0 = does not smoke, 1 = smokes). The effects of these three exogenous covariates were modeled using the nodecov term. gender is a binary indicator for student gender (0 = male, 1 = female). class indicates classroom membership. The effects of these two control variables were modeled using the nodematch term.

4.4.1 | Estimation results for candidate models

For comparison purposes, parameter estimates from each candidate model are presented in Table 4. According to the best-fitting model (M1 in Table 2), the estimated coefficients for the extraversion, openness, and smoking status are .323, −.028, and .189, respectively. These values are close to the model-averaged estimates from BMA (.325, −.017, and .173; see Table 3), which may give the impression that model averaging offers little additional benefit. However, it is important to recognize that the BMA estimates are computed by averaging across all candidate models, weighted by their posterior model probabilities. Given that the posterior model probability of M1 is approximately .600, the BMA estimates provide a more nuanced summary that incorporates uncertainty in model selection, rather than relying solely on a single model. This distinction becomes especially relevant when examining how parameter estimates vary across models. For example, the coefficient for openness takes on positive values in models such as M5 and M7, while the coefficient for smoking status reaches values as high as .263 in M7. These variations underscore the risk of overconfidence and potential misinterpretation when relying on a single, possibly misspecified, model.

4.5 | Sensitivity analysis of prior model probabilities

A critical aspect of implementing BMA, as noted in the 'Practical Considerations' subsection, is the specification of prior probabilities over the model space. Our initial analyses employed a uniform prior.[3] However, researchers are encouraged to test alternative prior specifications to assess the robustness of their BMA results. To examine the impact of different prior settings, we conducted a sensitivity analysis of prior model probabilities. We looked into two distinct approaches: one using a beta-binomial prior and another using a theory-informed prior.

4.5.1 | Beta-binomial prior

Assigning a uniform prior over models in the candidate pool is regarded as a neutral choice (Hoeting et al., 1999); however, this approach has an unintended consequence. When the prior is re-expressed in terms of model size (i.e., the number of included predictors), the implied prior is no longer uniform. Because there

[3]We follow a common convention in the BMA literature by first presenting results based on diffuse prior settings and subsequently examining alternative specifications. This order reflects practices adopted in previous BMA studies (see, e.g., Berkhout et al., 2024; Hinne et al., 2020; van den Bergh et al., 2021).

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

are $\binom{p}{x}$ models that include x out of p possible predictors, the uniform prior assigns more total probability to those values of x with many such combinations. As a result, the implied distribution over model size is bell-shaped, with most mass around $p/2$. As articulated in van den Bergh et al. (2021), Wilson et al. (2010), Ley and Steel (2009), and Castillo et al. (2015), this bias disproportionately favours models that include approximately half the total covariates, thereby down-weighting both sparse and dense model specifications. Such implicit preferences can distort inference, especially in psychological applications of ERGMs, where theoretical considerations may motivate either a parsimonious set of terms or richly parameterized models.

An alternative approach is the beta-binomial model prior. This prior first assigns equal mass to each possible model size (i.e., the number of included predictors, $x = 0, \ldots, p$) and then distributes that mass uniformly across the models of that size. To set the hyperparameters α and β, we followed van den Bergh et al. (2021) and repeated the BMA analysis using $\alpha = \beta = 1$. This choice ensures a more agnostic stance to model size across the candidate model space. Other variants have been proposed in the literature. For instance, Wilson et al. (2010) suggested setting $\alpha = 1$ and $\beta = \lambda p$, where λ is a user-specified parameter. As another example, Castillo et al. (2015) proposed setting $\alpha = 1$ and $\beta = p^u$, where u is specified by researchers. Both approaches place greater prior weight on models with fewer predictors.

```
# Number of predictors in each model
predictor_counts <- c(1, 1, 1, 2, 2, 2, 3)  # Number of predictors in each
    of the 7 models

# Total number of predictors (p)
p <- 3

# Alpha and Beta for the Beta-Binomial Prior
alpha <- 1
beta <- 1

# Compute the prior probability for each model
prior_prob_beta_binomial <- sapply(predictor_counts, function(x) {
  beta_binom_chunk <- choose(p, x) * beta(x + alpha, p - x + beta) / beta(
    alpha, beta)
  beta_binom_chunk / choose(2^p, x)  # Distribute uniformly within each
    group of x predictors
})

# Normalize to ensure prior probabilities sum to 1
prior_prob_beta_binomial <- prior_prob_beta_binomial / sum(prior_prob_beta
    _binomial)

# Compute posterior model probabilities using beta-binomial prior
(post_prob_beta_binomial <- as.numeric(e^logml * prior_prob_beta_binomial
    / sum(e^logml * prior_prob_beta_binomial)))
```

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The results of the sensitivity analysis under the beta-binomial prior, shown in Table 5, illustrate how this model prior influences prior and posterior inclusion probabilities, as well as inclusion Bayes factors, for the attribute covariates in the college friendship network. Under this setting, the prior inclusion probabilities for all three covariates are approximately .429, which is slightly lower than the uniform prior value of .571 used in the main analysis (Table 3). Despite the lower prior probabilities, the posterior inclusion probabilities remain largely stable for extraversion and smoking status. Specifically, the posterior inclusion probability is 1.000 for extraversion and .788 for smoking status. For openness, however, the posterior inclusion probability decreases to .407, falling below the prior inclusion probability. The inclusion Bayes factors changes accordingly under this prior specification. Extraversion continues to receive decisive support, with a Bayes factor of 4.750×10^8, down from 1.166×10^9 under the uniform prior. Although numerically reduced, this change does not alter the interpretation of the strength of the evidence. The Bayes factor for smoking decreases to 4.970, compared to 10.348 under the uniform prior, indicating that while the strength of evidence is halved, it still supports inclusion. In contrast, the evidence for openness weakens, with a Bayes factor of .916, which slightly favours its exclusion. These findings highlight that, compared to extraversion and smoking status, both of which consistently receive evidence for inclusion, openness appears more sensitive to prior specification and may not be supported in alternative prior settings.

4.5.2 | Theory-informed prior

Researchers may adjust prior model probabilities to reflect substantive and theoretical considerations, particularly in contexts where prior knowledge identifies certain covariates as theoretically important for explaining and predicting relational patterns. This theory-informed approach, by assigning higher prior weights to models containing such covariates, facilitates hypothesis testing grounded in substantive theory rather than relying solely on data-driven evidence in the BMA process. In this second example, we conducted a sensitivity analysis by hypothetically assuming that models including openness better explain and predict friendship ties. Accordingly, we assigned a prior model probability of .3 to the candidate model containing openness as a nodal covariate, while the remaining models received equal prior weights of .117. This setup implies that the model with openness received approximately three times more support at the prior stage. We selected this scenario because openness was the covariate that received less empirical support than extraversion or smoking status in the initial analysis. In this regard, the sensitivity analysis serves to assess the robustness of the findings under an alternative prior structure that emphasizes the effect of openness, even though earlier results identified extraversion and smoking status as more influential predictors.

TABLE 5 Sensitivity analysis of prior model probabilities and evidence of parameter inclusion.

Parameter	$p(\text{incl})$	$p(\text{incl}\mid \boldsymbol{y})$	BF_{incl}
Beta-binomial prior			
Network ~ extraversion	.429	1.000	4.750×10^8
Network ~ openness	.429	.407	.916
Network ~ smoke	.429	.788	4.970
Theory-Informed Prior			
Network ~ extraversion	.467	1.000	6.931×10^8
Network ~ openness	.650	.626	.903
Network ~ smoke	.467	.932	15.768

Note: $p(\text{incl})$ refers to prior inclusion probability. $p(\text{incl}\mid \boldsymbol{y})$ refers to posterior inclusion probability. BF_{incl} refers to the inclusion Bayes factor. Network refers to the college friendship network. Extraversion and openness are personality trait scores reflecting students' levels of extraversion and openness to experience, respectively. Smoke is a binary indicator for smoking status (0 = does not smoke, 1 = smokes). The effects of these three exogenous covariates were modeled using the nodecov term.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

```
# Define weights manually
# Assign a weight of 0.3 to the model with the openness term only
high_weight <- 0.3  # Prior weight for the model with the openness term
    only
remaining_weight <- (1 - high_weight) / 6  # Equal weight for the
    remaining 6 models

manual_prior_weights <- c(
  remaining_weight,  # Model 1
  remaining_weight,  # Model 2
  high_weight,       # Model 3 (openness)
  remaining_weight,  # Model 4
  remaining_weight,  # Model 5
  remaining_weight,  # Model 6
  remaining_weight   # Model 7
)

# Normalize to ensure the prior probabilities sum to 1 (optional step for
    safety)
manual_prior_weights <- manual_prior_weights / sum(manual_prior_weights)

# Compute posterior model probabilities using the manual prior
(post_prob_manual <- as.numeric(e^logml * manual_prior_weights / sum(e^
    logml * manual_prior_weights)))
```

The results of this analysis are summarized in Table 5. Under the theory-informed prior, where greater weight was assigned to the candidate model including openness, the prior inclusion probability for openness increased to .650, while the probabilities for extraversion and smoking status were set to .467. Despite this adjustment, the posterior inclusion probabilities and inclusion Bayes factors remained largely consistent with those obtained under the beta-binomial prior. Specifically, the posterior inclusion probability for openness was .626, and the corresponding Bayes factor was .903, suggesting evidence for exclusion. In contrast, extraversion and smoking status maintained strong support, with posterior inclusion probabilities of 1.000 and .932 (identical to those in Table 3) and Bayes factors of 6.931×10^8 and 15.768, respectively. These findings suggest that even when greater prior weight is assigned to openness based on substantive considerations, the data continue to favour extraversion and smoking status as more robust predictors of tie formation in the friendship network.

In summary, across both sensitivity analyses, extraversion and smoking status consistently received strong support for inclusion, as indicated by high posterior inclusion probabilities and large Bayes factors, regardless of the prior model probability settings. Openness, by contrast, showed evidence toward exclusion, with its inclusion Bayes factor falling below 1 under both the beta-binomial and theory-informed priors. These sensitivity analyses therefore suggest that the main conclusions regarding the

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

importance of extraversion and smoking status are robust to changes in prior assumptions, whereas the role of openness remains more uncertain and sensitive to prior specification.

4.6 | Theoretical implications of the results

4.6.1 | Effects of exogenous covariates

Consistent with our theoretical prediction, extraversion emerged as a robust and positive predictor of friendship in the college network. Specifically, we found that individuals higher in extraversion were more likely to maintain a greater number of friendships. This finding supports the theoretical postulate that more extraverted individuals tend to form and maintain more friendship ties due to their comparatively higher levels of sociability, talkativeness, assertiveness, and proactive engagement with peers (Rubin et al., 2006; Selfhout et al., 2010; Wagner et al., 2014). The BMA results also support the inclusion of smoking status as a relevant covariate. Specifically, we found that individuals who smoke were more likely to have more friendships than those who do not. This finding aligns with the theoretical idea that smoking functions as a visible behavioural marker that signals and enhances social status within college peer culture, thereby increasing one's attractiveness as a friend (Lakon et al., 2015; Schaefer et al., 2012).

By contrast, openness to experience was less important as a predictor of the number of friendships and even showed a small negative effect. This finding suggests that, within this particular network, openness may not be a salient factor in explaining or predicting friendship—or, if it is, that individuals higher in openness tend to maintain smaller friendship networks. Despite theoretical expectations that higher openness would lead individuals to engage with a broader range of people and social contexts—resulting in larger friendship networks—we did not observe this pattern. However, this result is consistent with prior research reporting mixed or nonsignificant effects of openness on friendship (Harris & Vazire, 2016).

One possible explanation for this finding, as proposed by Harris and Vazire (2016), is that openness to experience may be less directly associated with friendship than other Big Five traits, such as extraversion. Instead, openness may be more relevant to outcomes like an individual's position within the friendship network or the diversity of their friendship ties. Harris and Vazire (2016) also emphasize that openness is closely linked to personal values and attitudes, which are more likely to influence friendships through dyadic mechanisms such as homophily, rather than through increased social activity. This suggests that openness to experience may shape the formation and maintenance of friendships primarily when individuals share similar levels of openness.

Taken together, these results demonstrate how BMA can be used not only to assess statistical support for covariates but also to critically revisit and refine theoretical expectations. The strong evidence for extraversion and smoking status affirms their roles as socially consequential traits in structuring friendships within college networks. In contrast, the limited support for openness highlights the need to reconsider its theoretical relevance and encourages further empirical investigation into the mechanisms and contexts through which this trait may influence friendship.

4.6.2 | Effects of endogenous terms and control variables

With respect to the endogenous and control effects, inspection of the various model coefficients in Table 2 shows that the GWESP term, as well as the gender and class homophily terms, were all positive predictors of friendship across the candidate models. These results lend support to several theoretical frameworks: balance theory, which posits a tendency toward triadic closure in social networks (Heider, 1958); the homophily principle, which suggests that similarity in gender promotes friendship formation in college settings (McPherson et al., 2001); and focus theory, which argues that shared

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

contexts—such as being in the same class—structure opportunities for interaction and increase the likelihood of within-class friendships (Feld, 1981).

5 | EMPIRICAL ANALYSIS OF THE FLORENTINE MARRIAGE ALLIANCE NETWORK: A PREVIEW

In addition to the college friendship network example presented earlier, we provide a second empirical example based on the Florentine marriage alliance network, which is available via the OSF at https://osf.io/g9eq4/. This second example is included as supplementary material to avoid overloading the main article, while still showcasing additional aspects of ERGMs relevant to implementing BMA. The structure of this example parallels that of the college friendship network analysis, allowing readers to follow the analysis in a similar fashion. To preview a key conceptual insight from the second example and encourage readers to consult the full analysis, we briefly describe the behaviour of Bayes factors and their sensitivity to prior choice in the following subsection.

5.1 | The Jeffreys–Lindley paradox and Bayes factors

An important point illustrated in the second example concerns the behaviour of Bayes factors in the application of ERGMs. In estimating the candidate models and obtaining their marginal likelihoods, we employed weakly informative priors (a mean vector of zeros and a diagonal covariance matrix with elements set to 4). Although such priors are generally recommended for ERGMs (Krause et al., 2020), they may introduce a degree of arbitrariness that can give rise to the Jeffreys–Lindley paradox (Jeffreys, 1961; Lindley, 1957; Wagenmakers & Ly, 2023). As noted by Mulder et al. (2024) in the context of ERGMs, this paradox describes the tendency of Bayes factors to increasingly favour simpler models as network size grows, even when the data contain subtle signals that support more complex alternatives. When the simpler model is in fact true, the Bayes factor can become arbitrarily large in its favour (i.e., diverge to infinity), causing the posterior probability of the simpler model to converge to one. This tendency was observed in the second example, and we illustrate how researchers can address it by introducing the unit information prior approach for ERGMs proposed by Mulder et al. (2024) and conducting a sensitivity analysis.

6 | DISCUSSION

With the growing integration of social network analysis into psychology, ERGMs have become increasingly popular tools for explaining and predicting the social network structures underlying psychological phenomena. Valid inference in ERGMs relies on the correct specification of both endogenous and exogenous terms, guided by substantive theory.[4] However, accurately capturing the true generative process behind network formation with an ERGM remains a complex and ongoing challenge in applied settings.

The primary objectives of this tutorial were to illustrate BMA for ERGMs as a multi-model inferential framework and to demonstrate its application through the theory-driven analysis of empirical social network data. This tutorial and implementation guide represent one of the first empirical applications of BMA for ERGMs and offer detailed and practical guidelines to support researchers in navigating theoretical and methodological decisions in this process. Given the advantages of BMA, we outlined how

[4]Although this tutorial has been motivated by the need to correctly specify endogenous and exogenous terms to capture the true network-generating processes, this is not the only important modelling assumption. To provide a broader overview of additional considerations, we include a discussion in Appendix 1 for interested readers.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

it can be integrated into ERGMs as a principled alternative to address issues of model misspecification and the limitations of single-model inference.

6.1 | Implications of BMA for ERGMs

Our empirical analysis highlights several benefits of using BMA for ERGMs. With weakly informative priors and uniform prior model probabilities, the best-fitting model received a posterior model probability of .600 in the college friendship network and .578 in the Florentine marriage alliance network (see Data S1). This indicates that while the best-fitting model in each example was the most plausible after observing the network structure, other models still contributed to explaining and predicting the observed networks. If only the best model were selected for single-model inference, such a decision could lead to overconfidence in its results and ignore other models that could account for additional uncertainty. If the selected model does not reflect the true network-generating process, thus constituting a misspecified model, researchers risk drawing conclusions that fail to capture the underlying data-generating mechanism. BMA mitigates this overconfidence by considering multiple candidate models and provides a more nuanced understanding of the network structure. Referring back to our empirical illustration, we explicitly considered parameter estimates not only from the best model but also from other plausible candidate models through BMA. This is an obvious advantage of using BMA that mitigates the risks associated with exponential random graph modelling in applied research.

In addition, the Bayesian model-averaged posterior distribution and the inclusion Bayes factor incorporate parameter estimates from all candidate models. The quantification of averaged effects again avoids the pitfalls of relying on estimates of a single model, which could lead to drawing inferences based on capitalization on chance. In single-model inference, parameter estimates might fluctuate with each refitting, whereas BMA averages and updates parameters based on multiple models. As such, BMA incorporates contributions of parameters under consideration into a comprehensive overview of the effects. This process, hence, naturally accounts for the information from multiple models and ensures robust and reliable inference. We also note that determining the pool of candidate models and assigning weights to the model space are critical methodological steps in BMA. As illustrated in our empirical application, we encourage researchers to engage in these processes to assess the robustness of their results.

An important limitation of this tutorial is that we applied BMA to three exogenous covariates. That is, we did not extend BMA to many other types of ERGM specification decisions that researchers commonly face. These include BMA for theoretically central endogenous effects, for control variables (whether endogenous or exogenous), and for comparing alternative operationalizations of a given theoretical construct. Although our demonstration covers only a segment of the broader model decision space, the approach is, in principle, adaptable to any set of theoretically motivated terms and associated specification choices. Indeed, BMA may be even more valuable in more complex decision contexts, where multiple a priori models are plausible. As long as no single model overwhelmingly dominates the posterior model probability, BMA provides a principled framework for accounting for model uncertainty. If one model does dominate, BMA effectively reduces to single-model selection (Hinne et al., 2020), though it remains valuable to make that dominance explicit.

6.2 | Future methodological innovations

This tutorial features a novel empirical implementation of BMA for ERGMs, marking not only a practical demonstration but also a starting point for further methodological development. In the course of conducting this tutorial, several methodological considerations emerged that we believe warrant future research. Here, we outline directions for advancing BMA in the context of ERGMs.

First, future development should be concerned with computational and algorithmic advancements. In outlining BMA for ERGMs, our analyses relied on the `Bergm` package to compute Bayes factors,

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

posterior model probabilities, and inclusion Bayes factors given a relatively small pool of candidate models. However, if researchers consider models with multiple complex endogenous dependencies (e.g., geometrically weighted terms or different decay parameters for geometrically weighted degree terms) based on substantive theories, in addition to multiple exogenous covariates, computational challenges are likely to arise. For example, averaging over these complex models may lead to estimation and convergence issues, such as poorly mixed MCMC samples, discontinued optimization, and inestimable parameters (e.g., Duxbury, 2021; Lusher et al., 2013). In addition, the `Bergm` package itself does not address model degeneracy, which occurs when MCMC samples become degenerate, particularly in larger networks or when models contain dyad-dependent network statistics. These obstacles could pose practical challenges for researchers, potentially leading to the exclusion of theoretically relevant endogenous or exogenous terms from the BMA analysis. Indeed, our empirical example did not address how researchers should manage model degeneracy when it arises in the context of BMA for ERGMs. In this regard, a valuable contribution would involve developing algorithms to detect and mitigate estimation issues, thus allowing reliable parameter estimation without convergence issues. Such algorithms could also integrate parameter estimation and model-averaging processes into a single, streamlined step for enhanced computational efficiency. In particular, tools that enable both posterior sampling and marginal likelihood estimation within a unified routine would greatly facilitate the practical implementation of BMA for ERGMs. These advancements would not only address these technical limitations but also make BMA for ERGMS more accessible to applied researchers.

Defining model-searching algorithms presents another interesting future research avenue. The current tutorial explored candidate models given a small fixed set of exogenous covariates. However, as researchers include increasing numbers of endogenous or exogenous effects and further interactions between them, the size of the pool of candidate models may become staggeringly large. Under such circumstances, manually checking the estimability of candidate models and excluding inestimable models can be challenging and inefficient. Therefore, more sophisticated model search algorithms can be devised to automatically specify candidate models and exclude those with little to no explanatory power for the observed data.

Finally, our tutorial focused on a basic class of ERGMs with undirected (i.e., symmetric) ties. However, ERGMs can be extended to accommodate more complex network types, including directed and valued networks (Cranmer et al., 2020; Desmarais & Cranmer, 2012). At present, existing software packages do not support Bayesian inference, and therefore BMA, for these ERGMs. Future methodological work could usefully focus on developing tools and frameworks that allow BMA to be applied in these extended contexts. Such developments would significantly expand the applicability of BMA and support researchers working with a wider range of network data structures.

7 | CONCLUDING REMARKS

We illustrated BMA for ERGMs and demonstrated its implementation through a theoretical and practical guide, supported by accompanying R code. This novel application of BMA to ERGMs marks an important first step, guiding applied researchers wishing to implement the method and laying the groundwork for future methodological advancements. We envision promising developments that build on this foundation, and we hope that this work inspires further innovations. Ultimately, such developments could help facilitate the broader adoption of BMA among exponential random graph modelers in psychology and related fields to take full advantage of model averaging in social network analysis.

AUTHOR CONTRIBUTIONS

Ihnwhi Heo: conceptualization; methodology; software; data curation; investigation; project administration; formal analysis; visualization; writing – original draft; writing – review and editing. **Jan-Willem Simons:** conceptualization; methodology; writing – review and editing; writing – original draft. **Haiyan Liu:** writing – review and editing; writing – original draft; data curation; conceptualization.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ACKNOWLEDGEMENTS
This research was not funded.

CONFLICT OF INTEREST STATEMENT
We have no conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT
The materials supporting this tutorial, including the code and datasets used in the college friendship and Florentine marriage alliance network analyses, are openly available via the OSF at https://osf.io/g9eq4/.

ORCID
Ihnwhi Heo https://orcid.org/0000-0002-6123-3639
Jan-Willem Simons https://orcid.org/0000-0002-3341-5320
Haiyan Liu https://orcid.org/0000-0002-4272-9399

REFERENCES
Allan, G., & Phillipson, C. (2017). *Social networks and social exclusion: Sociological and policy perspectives*. Routledge.
Anderson, C. J., Wasserman, S., & Faust, K. (1992). Building stochastic blockmodels. *Social Networks*, *14*(1–2), 137–161. https://doi.org/10.1016/0378-8733(92)90017-2
Basu, K., & Sen, A. (2021). Identifying individuals associated with organized criminal networks: A social network analysis. *Social Networks*, *64*, 42–54. https://doi.org/10.1016/j.socnet.2020.07.009
Berkhout, S. W., Haaf, J. M., Gronau, Q. F., Heck, D. W., & Wagenmakers, E.-J. (2024). A tutorial on Bayesian model-averaged meta-analysis in JASP. *Behavior Research Methods*, *56*(3), 1260–1282. https://doi.org/10.3758/s13428-023-02093-6
Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, *36*(2), 192–225. https://doi.org/10.1111/j.2517-6161.1974.tb00999.x
Borgatti, S. P., Agneessens, F., Johnson, J. C., & Everett, M. G. (2024). *Analyzing social networks*. SAGE Publications Ltd.
Borgatti, S. P., Carley, K. M., & Krackhardt, D. (2006). On the robustness of centrality measures under conditions of imperfect data. *Social Networks*, *28*(2), 124–136. https://doi.org/10.1016/j.socnet.2005.05.001
Bouranis, L., Friel, N., & Maire, F. (2018). Bayesian model selection for exponential random graph models via adjusted pseudo-likelihoods. *Journal of Computational and Graphical Statistics*, *27*(3), 516–528. https://doi.org/10.1080/10618600.2018.1448832
Broda, M. D., Granger, K., Chow, J., & Ross, E. (2023). Using social network analysis in applied psychological research: A tutorial. *Psychological Methods*, *28*(4), 791. https://doi.org/10.1037/met0000451
Bull, S. B., Greenwood, C. M., & Donner, A. (1994). Efficiency of reduced logistic regression models. *Canadian Journal of Statistics*, *22*(3), 319–334. https://doi.org/10.2307/3315595
Burt, R. S., Kilduff, M., & Tasselli, S. (2013). Social network analysis: Foundations and frontiers on advantage. *Annual Review of Psychology*, *64*, 527–547. https://doi.org/10.1146/annurev-psych-113011-143828
Cai, Y., Liu, L., & Li, Z. (2024). Modelling and monitoring social network change based on exponential random graph models. *Journal of Applied Statistics*, *51*(9), 1621–1641. https://doi.org/10.1080/02664763.2023.2230530
Caimo, A., Bouranis, L., Krause, R., & Friel, N. (2022). Statistical network analysis with Bergm. *Journal of Statistical Software*, *104*(1), 1–23. https://doi.org/10.18637/jss.v104.i01
Caimo, A., & Friel, N. (2011). Bayesian inference for exponential random graph models. *Social Networks*, *33*(1), 41–55. https://doi.org/10.1016/j.socnet.2010.09.004
Caimo, A., & Friel, N. (2013). Bayesian model selection for exponential random graph models. *Social Networks*, *35*(1), 11–24. https://doi.org/10.1016/j.socnet.2012.10.003
Caimo, A., & Friel, N. (2014). Bergm: Bayesian exponential random graphs in R. *Journal of Statistical Software*, *61*(2), 1–25. https://doi.org/10.18637/jss.v061.i02
Castillo, I., Schmidt-Hieber, J., & Van der Vaart, A. (2015). Bayesian linear regression with sparse priors. *The Annals of Statistics*, *43*, 1986–2018. https://doi.org/10.1214/15-AOS1334
Chib, S., & Jeliazkov, I. (2001). Marginal likelihood from the Metropolis–Hastings output. *Journal of the American Statistical Association*, *96*(453), 270–281. https://doi.org/10.1198/016214501750332848
Cranmer, S. J., Desmarais, B. A., & Morgan, J. W. (2020). *Inferential network analysis*. Cambridge University Press.
Desmarais, B. A., & Cranmer, S. J. (2012). Statistical inference for valued-edge networks: The generalized exponential random graph model. *PLoS One*, *7*(1), e30136. https://doi.org/10.1371/journal.pone.0030136
Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The Mini-IPIP scales: Tiny-yet-effective measures of the big five factors of personality. *Psychological Assessment*, *18*(2), 192–203. https://doi.org/10.1037/1040-3590.18.2.192
Duxbury, S. W. (2021). Diagnosing multicollinearity in exponential random graph models. *Sociological Methods & Research*, *50*(2), 491–530. https://doi.org/10.1177/0049124118782543

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Feld, S. L. (1981). The focused organization of social ties. *American Journal of Sociology*, *86*(5), 1015–1035. https://doi.org/10.1086/227352

Festinger, L. (1957). Social comparison theory. *Selective Exposure Theory*, *16*, 401.

Flakus, M., Danieluk, B., Baran, L., Kwiatkowska, K., Rogoza, R., & Schermer, J. A. (2021). Are intelligent peers liked more? Assessing peer-reported liking through the network analysis. *Personality and Individual Differences*, *177*, 110844. https://doi.org/10.1016/j.paid.2021.110844

Friel, N., Hurn, M., & Wyse, J. (2014). Improving power posterior estimation of statistical evidence. *Statistics and Computing*, *24*, 709–723. https://doi.org/10.1007/s11222-013-9397-1

Friel, N., & Pettitt, A. N. (2008). Marginal likelihood estimation via power posteriors. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *70*(3), 589–607. https://doi.org/10.1111/j.1467-9868.2007.00650.x

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, *2*, 1360–1383. https://doi.org/10.1214/08-AOAS191

Geyer, C. J., & Thompson, E. A. (1992). Constrained Monte Carlo maximum likelihood for dependent data. *Journal of the Royal Statistical Society: Series B (Methodological)*, *54*(3), 657–683. https://doi.org/10.1111/j.2517-6161.1992.tb01443.x

Gilman, R., Carboni, I., Perry, A., & Anderman, E. M. (2022). Social network analysis and its applications to school psychology: A tutorial. *School Psychology*, *37*(6), 424. https://doi.org/10.1037/spq0000529

Grund, T. U., & Densley, J. A. (2012). Ethnic heterogeneity in the activity and structure of a black street gang. *European Journal of Criminology*, *9*(4), 388–406. https://doi.org/10.1177/1477370812447738

Grund, T. U., & Densley, J. A. (2015). Ethnic homophily and triad closure: Mapping internal gang structure using exponential random graph models. *Journal of Contemporary Criminal Justice*, *31*(3), 354–370. https://doi.org/10.1177/1043986214553377

Grunspan, D. Z., Wiggins, B. L., & Goodreau, S. M. (2014). Understanding classrooms through social network analysis: A primer for social network analysis in education research. *CBE—Life Sciences Education*, *13*(2), 167–178. https://doi.org/10.1187/cbe.13-08-0162

Handcock, M. S., Robins, G., Snijders, T., Moody, J., & Besag, J. (2003). *Assessing degeneracy in statistical models of social networks*. Technical Report, Working Paper.

Harris, K., & Vazire, S. (2016). On friendship development and the big five personality traits. *Social and Personality Psychology Compass*, *10*(11), 647–667. https://doi.org/10.1111/spc3.12287

Heider, F. (1958). *The psychology of interpersonal relations*. John Wiley & Sons.

Hinne, M., Gronau, Q. F., van den Bergh, D., & Wagenmakers, E.-J. (2020). A conceptual introduction to bayesian model averaging. *Advances in Methods and Practices in Psychological Science*, *3*(2), 200–215. https://doi.org/10.1177/2515245919898657

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial (with comments by M. Clyde, David Draper and EI George, and a rejoinder by the authors). *Statistical Science*, *14*(4), 382–417. https://doi.org/10.1214/ss/1009212519

Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association*, *97*(460), 1090–1098. https://doi.org/10.1198/016214502388618906

Hogg, M. A. (2000). Subjective uncertainty reduction through self-categorization: A motivational theory of social identity processes. *European Review of Social Psychology*, *11*(1), 223–255. https://doi.org/10.1080/14792772043000040

Holland, P. W., Laskey, K. B., & Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social Networks*, *5*(2), 109–137. https://doi.org/10.1016/0378-8733(83)90021-7

Holler, C., & Schüßler, A. (2024). Gauging acceptance: A multifaceted examination of physical ability and its role for peer networks in adolescent physical education. *The Journal of Early Adolescence*, 02724316241265459. https://doi.org/10.1177/02724316241265459

Hunter, D. R. (2007). Curved exponential family models for social networks. *Social Networks*, *29*(2), 216–230. https://doi.org/10.1016/j.socnet.2006.08.005

Hunter, D. R., Goodreau, S. M., & Handcock, M. S. (2008). Goodness of fit of social network models. *Journal of the American Statistical Association*, *103*(481), 248–258. https://doi.org/10.1198/016214507000000446

Jeffreys, H. (1961). *Theory of probability*. Oxford University Press.

Jiao, C., Wang, T., Liu, J., Wu, H., Cui, F., & Peng, X. (2017). Using exponential random graph models to analyze the character of peer relationship networks and their effects on the subjective well-being of adolescents. *Frontiers in Psychology*, *8*, 583. https://doi.org/10.3389/fpsyg.2017.00583

Kaplan, D. (2021). On the quantification of model uncertainty: A Bayesian perspective. *Psychometrika*, *86*(1), 215–238. https://doi.org/10.1007/s11336-021-09754-5

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Koskinen, J., Caimo, A., & Lomi, A. (2015). Simultaneous modeling of initial conditions and time heterogeneity in dynamic networks: An application to foreign direct investments. *Network Science*, *3*(1), 58–77. https://doi.org/10.1017/nws.2015.3

Koskinen, J. H., Robins, G. L., & Pattison, P. E. (2010). Analysing exponential random graph (p-star) models with missing data using Bayesian data augmentation. *Statistical Methodology*, *7*(3), 366–384. https://doi.org/10.1016/j.stamet.2009.09.007

Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, *115*(2), 405–450. https://doi.org/10.1086/599247

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Krause, R. W., Huisman, M., Steglich, C., & Snijders, T. (2020). Missing data in cross-sectional networks–an extensive comparison of missing data treatment methods. *Social Networks*, *62*, 99–112. https://doi.org/10.1016/j.socnet.2020.02.004

Krivitsky, P. N., & Goodreau, S. M. (2019). *STERGM – Separable temporal ergms for modeling discrete relational dynamics with statnet*. Technical Report, Statnet Project. http://statnet.csde.washington.edu/workshops/SUNBELT

Krivitsky, P. N., & Kolaczyk, E. D. (2015). On the question of effective sample size in network modeling: An asymptotic inquiry. *Statistical Science*, *30*(2), 184–198. https://doi.org/10.1214/14-STS502

Kwiatkowska, M. M., & Rogoza, R. (2019). Shy teens and their peers: Shyness in respect to basic personality traits and social relations. *Journal of Research in Personality*, *79*, 130–142. https://doi.org/10.1016/j.jrp.2019.03.005

Lakon, C. M., Wang, C., Butts, C. T., Jose, R., Timberlake, D. S., & Hipp, J. R. (2015). A dynamic model of adolescent friendship networks, parental influences, and smoking. *Journal of Youth and Adolescence*, *44*, 1767–1786. https://doi.org/10.1007/s10964-014-0187-7

Lee, C., & Wilkinson, D. J. (2019). A review of stochastic block models and extensions for graph clustering. *Applied Network Science*, *4*(1), 1–50. https://doi.org/10.1007/s41109-019-0232-2

Leifeld, P., Cranmer, S. J., & Desmarais, B. A. (2018). Temporal exponential random graph models with btergm: Estimation and bootstrap confidence intervals. *Journal of Statistical Software*, *83*(6), 1–36. https://doi.org/10.18637/jss.v083.i06

Ley, E., & Steel, M. F. (2009). On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. *Journal of Applied Econometrics*, *24*(4), 651–674. https://doi.org/10.1002/jae.1057

Li, Y., & Clyde, M. A. (2018). Mixtures of g-priors in generalized linear models. *Journal of the American Statistical Association*, *113*(524), 1828–1845. https://doi.org/10.1080/01621459.2018.1469992

Light, R., & Moody, J. (2020). *The oxford handbook of social networks*. Oxford University Press.

Lindley, D. V. (1957). A statistical paradox. *Biometrika*, *44*(1/2), 187–192. https://doi.org/10.2307/2333251

Liu, H., Jin, I. H., & Zhang, Z. (2018). Structural equation modeling of social networks: Specification, estimation, and application. *Multivariate Behavioral Research*, *53*(5), 714–730. https://doi.org/10.1080/00273171.2018.1479629

Liu, H., Jin, I.-H., Zhang, Z., & Yuan, Y. (2021). Social network mediation analysis: A latent space approach. *Psychometrika*, *86*, 271–298. https://doi.org/10.1007/s11336-020-09736-z

Liu, H., & Zhang, Z. (2021). Birds of a feather flock together and opposites attract: The nonlinear relationship between personality and friendship. *Journal of Behavioral Data Science*, *1*(1), 34–52. https://doi.org/10.35566/jbds/v1n1/p3

Lusher, D., Koskinen, J., & Robins, G. (2013). *Exponential random graph models for social networks: Theory, methods, and applications*. Cambridge University Press.

Massidda, D., Carta, M. G., & Altoè, G. (2016). Integrating different factorial solutions of a psychometric tool via social network analysis. *Methodology*, *12*(3), 97–106. https://doi.org/10.1027/1614-2241/a000113

McElroy, E., McIntyre, J. C., Bentall, R. P., Wilson, T., Holt, K., Kullu, C., Nathan, R., Kerr, A., Panagaki, K., McKeown, M., Saini, P., Gabbay, M., & Corcoran, R. (2019). Mental health, deprivation, and the neighborhood social environment: A network analysis. *Clinical Psychological Science*, *7*(4), 719–734. https://doi.org/10.1177/2167702619830640

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Montgomery, J. M., & Nyhan, B. (2010). Bayesian model averaging: Theoretical developments and practical applications. *Political Analysis*, *18*(2), 245–270. https://doi.org/10.1093/pan/mpq001

Mulder, J., Friel, N., & Leifeld, P. (2024). Bayesian testing of scientific expectations under exponential random graph models. *Social Networks*, *78*, 40–53. https://doi.org/10.1016/j.socnet.2023.11.004

Pachucki, M. A., & Breiger, R. L. (2010). Cultural holes: Beyond relationality in social networks and culture. *Annual Review of Sociology*, *36*, 205–224. https://doi.org/10.1146/annurev.soc.012809.102615

Pircalabelu, E., & Claeskens, G. (2016). Focused model selection for social networks. *Social Networks*, *46*, 76–86. https://doi.org/10.1016/j.socnet.2016.03.002

Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, *92*(437), 179–191. https://doi.org/10.1080/01621459.1997.10473615

Rinaldo, A., Fienberg, S. E., & Zhou, Y. (2009). On the geometry of discrete exponential families with application to exponential random graph models. *Electronic Journal of Statistics*, *3*, 446–484.

Roberts, G., & Gilks, W. (1994). Convergence of adaptive direction sampling. *Journal of Multivariate Analysis*, *49*(2), 287–298. https://doi.org/10.1006/jmva.1994.1028

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, *29*(2), 173–191. https://doi.org/10.1016/j.socnet.2006.08.002

Roccas, S., Sagiv, L., Schwartz, S. H., & Knafo, A. (2002). The big five personality factors and personal values. *Personality and Social Psychology Bulletin*, *28*(6), 789–801. https://doi.org/10.1177/0146167202289008

Rubin, K. H., Bukowski, W., & Parker, J. G. (2006). Peer interactions, relationships, and groups. In W. Damon, M. Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology* (pp. 619–700). Wiley.

Schaefer, D. R., Haas, S. A., & Bishop, N. J. (2012). A dynamic model of us adolescents' smoking and friendship networks. *American Journal of Public Health*, *102*(6), e12–e18. https://doi.org/10.2105/AJPH.2012.300705

Schüßler, A., Holler, C., & Hill, Y. (2025). Do you prefer to collaborate with students pursuing the same goals? – A network analysis of physical education classes. *British Journal of Educational Psychology*, 1–20. https://doi.org/10.1111/bjep.12757

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Schweinberger, M., & Handcock, M. S. (2015). Local dependence in random graph models: Characterization, properties and statistical inference. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *77*(3), 647–676. https://doi.org/10.1111/rssb.12081

Scott, J. (2017). *Social network analysis*. Sage.

Sekulovski, N., Keetelaar, S., Huth, K., Wagenmakers, E.-J., van Bork, R., van den Bergh, D., & Marsman, M. (2024). Testing conditional independence in psychometric networks: An analysis of three Bayesian methods. *Multivariate Behavioral Research*, *59*(5), 913–933. https://doi.org/10.1080/00273171.2024.2345915

Selfhout, M., Burk, W., Branje, S., Denissen, J., Van Aken, M., & Meeus, W. (2010). Emerging late adolescent friendship networks and big five personality traits: A social network approach. *Journal of Personality*, *78*(2), 509–538. https://doi.org/10.1111/j.1467-6494.2010.00625.x

Snijders, T. A., Pattison, P. E., Robins, G. L., & Handcock, M. S. (2006). New specifications for exponential random graph models. *Sociological Methodology*, *36*(1), 99–153. https://doi.org/10.1111/j.1467-9531.2006.00176.x

Strauss, D., & Ikeda, M. (1990). Pseudolikelihood estimation for social networks. *Journal of the American Statistical Association*, *85*(409), 204–212. https://doi.org/10.1080/01621459.1990.10475327

Sweet, T. M. (2015). Incorporating covariates into stochastic blockmodels. *Journal of Educational and Behavioral Statistics*, *40*(6), 635–664. https://doi.org/10.3102/1076998615606110

Tejada-Gallardo, C., Blasco-Belled, A., & Alsinet, C. (2023). Does mental well-being predict being perceived as a happy peer? A longitudinal social network study. *Personality and Individual Differences*, *202*, 111988. https://doi.org/10.1016/j.paid.2022.111988

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, *1*(1), 1–26. https://doi.org/10.1038/s43586-020-00001-2

van den Bergh, D., Clyde, M. A., Gupta, A. R. K. N., de Jong, T., Gronau, Q. F., Marsman, M., Ly, A., & Wagenmakers, E.-J. (2021). A tutorial on Bayesian multi-model linear regression with BAS and JASP. *Behavior Research Methods*, *53*, 2351–2371. https://doi.org/10.3758/s13428-021-01552-2

van Duijn, M. A., Gile, K. J., & Handcock, M. S. (2009). A framework for the comparison of maximum pseudo-likelihood and maximum likelihood estimation of exponential family random graph models. *Social Networks*, *31*(1), 52–62. https://doi.org/10.1016/j.socnet.2008.10.003

Vega Yon, G. G. (2023). Power and multicollinearity in small networks: A discussion of "tale of two datasets: Representativeness and generalisability of inference for samples of networks" by Krivitsky, Coletti, and Hens. *Journal of the American Statistical Association*, *118*(544), 2228–2231. https://doi.org/10.1080/01621459.2023.2252041

Vega Yon, G. G., Slaughter, A., & de la Haye, K. (2021). Exponential random graph models for little networks. *Social Networks*, *64*, 225–238. https://doi.org/10.1016/j.socnet.2020.07.005

Wagenmakers, E.-J., & Ly, A. (2023). History and nature of the Jeffreys–Lindley paradox. *Archive for History of Exact Sciences*, *77*(1), 25–72. https://doi.org/10.1007/s00407-022-00298-3

Wagner, J., Lüdtke, O., Roberts, B. W., & Trautwein, U. (2014). Who belongs to me? Social relationship and personality characteristics in the transition to young adulthood. *European Journal of Personality*, *28*(6), 586–603. https://doi.org/10.1002/per.1974

Wang, D. J., Shi, X., McFarland, D. A., & Leskovec, J. (2012). Measurement error in network data: A re-classification. *Social Networks*, *34*(4), 396–409. https://doi.org/10.1016/j.socnet.2012.01.003

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. *8*). Cambridge University Press.

Wasserman, S., & Pattison, P. (1996). Logit models and logistic regressions for social networks: I. An introduction to markov graphs and p*. *Psychometrika*, *61*(3), 401–425. https://doi.org/10.1007/BF02294547

Wilson, M. A., Iversen, E. S., Clyde, M. A., Schmidler, S. C., & Schildkraut, J. M. (2010). Bayesian model search and multilevel inference for SNP association studies. *The Annals of Applied Statistics*, *4*(3), 1342. https://doi.org/10.1214/09-aoas322

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Heo, I., Simons, J.-W., & Liu, H. (2025). A tutorial on Bayesian model averaging for exponential random graph models. *British Journal of Mathematical and Statistical Psychology*, *00*, 1–41. https://doi.org/10.1111/bmsp.70007

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

APPENDIX 1: ADDITIONAL MODELLING CONSIDERATIONS FOR ERGMS

The current tutorial has focused on the BMA framework for ERGMs, motivated by the need to correctly specify endogenous and exogenous network statistics to ensure valid inference. In the current appendix, we introduce other modelling considerations for ERGMs that may benefit applied researchers. Specifically, we briefly discuss two aspects: (1) additional assumptions underlying ERGMs and (2) sample size requirements for reliable estimation. These discussions are intended to provide a more comprehensive perspective on ERGM implementation.

Further assumptions

In addition to the correct specification of endogenous and exogenous terms, ERGMs rely on several additional assumptions that warrant consideration in empirical applications. These include (1) the homogeneity of effects across the network, (2) the local dependence assumption, and (3) the assumption that the observed network is measured without error.

First, the homogeneity of effects assumes that the influence of network statistics—such as reciprocity or transitivity—applies uniformly across all nodes and dyads in the network (Robins et al., 2007). In real-world networks, however, individuals may differ in how they form ties. In such cases, researchers may assess this assumption by including interaction terms (e.g., between a dyadic characteristic and the reciprocity effect; see Lusher et al., 2013).

Second, the local dependence assumption posits that dependencies in the network are limited to certain edges or small, localized subgraphs (Schweinberger & Handcock, 2015). Under this assumption, the node set is partitioned into subgroups such that dependencies are allowed within each subgroup, but independence is assumed between subgroups. As noted by Schweinberger and Handcock (2015), an advantage of the local dependence is that it does not impose rigid assumptions about the structure or strength of dependence within subgraphs.

Finally, like other social network models, ERGMs typically assume that the observed network is measured without error. That is, all ties (or the absence thereof) are recorded accurately and reflect the true underlying relationships. In real-world relational data, measurement errors may arise and can potentially distort observed network configurations and network-derived statistics such as degree centrality (e.g., Borgatti et al., 2024; Wang et al., 2012). Even so, Borgatti et al. (2024) found that centrality measures tend to be robust to small amounts of error (e.g., up to 10%), suggesting that computing such measures may still be reasonable even when some degree of measurement inaccuracy is expected. Nevertheless, they emphasize the importance of striving for more accurate data collection whenever possible.

Sample size requirements

Determining the appropriate sample size for reliably estimating ERGMs remains an open and complex area of research. Most prior work on this topic has been conducted within the frequentist framework, and definitive guidelines are limited. As Krivitsky and Kolaczyk (2015) emphasize, sample size considerations depend not only on the number of observations but also on structural properties of the network, such as the number of nodes. Relatedly, Vega Yon et al. (2021) examined the power of ERGM estimation across varying sample sizes, population parameters, and network characteristics. Their simulation studies indicated that, for effect sizes ranging from .5 to 1.0, a discovery rate of approximately .75 could be achieved for the number of transitive triads in networks consisting of 30 to 50 nodes—networks considered relatively small. As such, generalizable rules remain elusive; yet, researchers may consult Vega Yon (2023) for practical guidance on sample size requirements for ERGMs.

20448317, 0, Downloaded from https://bpspsychub.onlinelibrary.wiley.com/doi/10.1111/bmsp.70007, Wiley Online Library on [18/08/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License