Network autocorrelation modeling
Dittrich, D.; Leenders, R.T.A.J.; Mulder, J.

Published in:
Sociological Methods and Research

Document version:
Publisher's PDF, also known as Version of record

DOI:
10.1177/0049124117729712

Publication date:
2019

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright, please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Network Autocorrelation Modeling: A Bayes Factor Approach for Testing (Multiple) Precise and Interval Hypotheses

Dino Dittrich¹, Roger Th. A. J. Leenders², and Joris Mulder¹

Abstract
Currently available (classical) testing procedures for the network autocorrelation can only be used for falsifying a precise null hypothesis of no network effect. Classical methods can be neither used for quantifying evidence for the null nor for testing multiple hypotheses simultaneously. This article presents flexible Bayes factor testing procedures that do not have these limitations. We propose Bayes factors based on an empirical and a uniform prior for the network effect, respectively, first. Next, we develop a fractional Bayes factor where a default prior is automatically constructed. Simulation results suggest that the first two Bayes factors show superior performance and are the Bayes factors we recommend. We apply the recommended Bayes factors to three data sets from the literature and compare the results to those coming from classical analyses using p values. R code for efficient computation of the Bayes factors is provided.

¹ Department of Methodology and Statistics, Tilburg University, Tilburg, the Netherlands
² Department of Organization Studies, Tilburg University, Tilburg, the Netherlands

Corresponding Author:
Dino Dittrich, Department of Methodology and Statistics, Tilburg University, Postbus 90153, 5000 LE Tilburg, the Netherlands.
Email: dino.dittrich@gmail.com
Keywords

network autocorrelation model, hypothesis testing, Bayes factor, fractional Bayes factor, informative prior

The network autocorrelation model (Ord 1975) has been extensively used to represent theories of social influence throughout recent decades. It allows researchers to quantify the strength of a peer effect in a network for a given theory of interpersonal influence while controlling for sociological and other covariates. The identification and magnitude of the peer, or network, effect \( \rho \), also known as the network autocorrelation, is often the focus of interest in model applications. Typically, a researcher aims to identify whether there is social influence present in the network, resulting in an inferential test of \( H_0 : \rho = 0 \) versus \( H_1 : \rho \neq 0 \). Subsequently, if the null hypothesis is rejected, the researcher then concludes that there is evidence for some degree of social influence.

Even though the network autocorrelation model and this statistical approach have yielded many interesting and theoretically useful findings, more intricate hypothesis tests are often more informative. For example, when a researcher is interested in testing whether the degree of social influence is zero, small, medium, or large, a more informative test would be \( H_0 : \rho = 0 \) versus \( H_1 : 0 < \rho \leq .25 \) versus \( H_2 : .25 < \rho \leq .5 \) versus \( H_3 : .5 < \rho < 1 \). This not only allows a researcher to conclude whether there is evidence for a nonzero amount of network autocorrelation in the network, but it grants the researcher the opportunity to simultaneously test several strengths of social influence against each other as well. This article focuses on Bayesian hypothesis testing procedures for such multiple precise and interval hypotheses for the network autocorrelation.

The standard approach to testing a network effect is null hypothesis significance testing. Classical null hypothesis significance tests such as the Wald test, the likelihood ratio test, or the Lagrange multiplier test are based on different test statistics, summary values constructed from the sample, which have asymptotically known distributions under the null hypothesis (Leenders 1995). Then, assuming the null hypothesis to be true, the probability of observing a test statistic at least as extreme as the observed test statistic is calculated. This probability is called the “\( p \) value.” Subsequently, the \( p \) value is compared to a prespecified significance level \( \alpha \) which is usually set to .05 (Weakliem 2004). If the \( p \) value is smaller than \( \alpha \), the null hypothesis is rejected. In this case, one would conclude that there is enough
evidence in the data to reject the null hypothesis of no network effect. If the $p$ value is larger than $\alpha$, there is not enough evidence in the data to reject the null.

Classical null hypothesis significance testing in the network autocorrelation model has a number of disadvantages. First, the procedure cannot be used to provide evidence in favor of the null hypothesis (Wetzels and Wagenmakers 2012); it can only falsify the null hypothesis. If the $p$ value is larger than $\alpha$, this ultimately implies a state of ignorance where the null can be neither rejected nor supported by the data. For example, if the estimated autocorrelation parameter is $\hat{\rho} = .16$ with a two-sided $p$ value of .08 (so the null hypothesis that $\rho = 0$ would not be rejected based on $\alpha = .05$), this does not mean that the null hypothesis is “accepted”; it merely means that judgment regarding the rejection of this particular hypothesis is suspended and that no degree of belief in the hypothesis has been determined. On the other hand, if the $p$ value is smaller than $\alpha$, it is still not possible to say how much evidence there is against the null (and certainly not how much evidence there is in favor of $\rho$ being .16 indeed), only that there is enough evidence to reject the null that $\rho = 0$ based on a chosen significance level. Second, classical null hypothesis significance tests are not consistent. If the null is true and the sample size grows to infinity, there is still a probability of $\alpha$ (typically .05) of drawing the incorrect conclusion that the null is false. This is undesirable, as one should be able to draw more accurate conclusions with growing sample size. Third, $p$ values in the network autocorrelation model depend heavily on asymptotic theory and consequently, the type I error probability is not controlled for in an accurate manner in the case of small networks (Dittrich, Leenders, and Mulder 2017). A final important issue in the context of this article is that $p$ values cannot be adequately used when testing multiple competing hypotheses against each other (Shaffer 1995). Instead, one can only test each hypothesis against the null which does not answer the question which hypothesis, out of a set of precise and interval hypotheses, is most supported by the data.

In this article, we propose Bayes factor tests (Jeffreys 1961; Kass and Raftery 1995; Mulder and Wagenmakers 2016) as an alternative approach to classical null hypothesis significance testing in the network autocorrelation model. The Bayes factor is a Bayesian hypothesis testing criterion that circumvents the aforementioned issues with null hypothesis significance testing. First, in contrast to classical null hypothesis significance testing, it allows the researcher to evaluate and quantify the relative evidence in the data in favor of the null, or any other, hypothesis against another hypothesis (Kuha 2004). These hypotheses can be precise hypotheses, for example,
$H_0 : \rho = 0$, or interval hypotheses, for example, $H_1 : 0 < \rho < 1$. For instance, a Bayes factor of $B_{01} = 5$ implies that it is five times more likely to observe the data under $H_0$ than under a specific alternative hypothesis $H_1$. Second, Bayes factors are consistent, that is, if $H_0$ is the true model, the Bayes factor $B_{01}$ tends to infinity as the sample size goes to infinity (Casella et al. 2009). In other words, the larger the sample size, the more do the data support one hypothesis over another. Third, the Bayes factor provides “exact” inference without the need for asymptotic approximations (De Oliveira and Song 2008). Lastly, the Bayes factor can be straightforwardly extended to test more than two hypotheses against each other, for example, $H_0 : \rho = 0$ versus $H_1 : -.25 < \rho < 0$ versus $H_2 : 0 < \rho \leq .25$ versus $H_3 : .25 < \rho \leq .5$ versus $H_4 : .5 < \rho < 1$ (Raftery, Madigan, and Hoeting 1997). This feature is of particular relevance in the network autocorrelation model, as in many network studies, researchers do not doubt that social influence occurs but are interested in testing competing theories about its strength. In summary, these advantageous properties explain the increasing usage of Bayes factors in social science research, such as in linear regression models (Braeken, Mulder, and Wood 2015), analysis of variance (Klugkist, Laudy, and Hoijtink 2005), repeated measures (Mulder et al. 2009), or structural equation modeling (Gu et al. 2014).

So far, only two Bayes factors for the network autocorrelation model have been developed in the literature. Hepple (1995) proposed a Bayes factor for testing competing connectivity matrices against each other, while LeSage and Parent (2007) provided Bayes factors for testing different explanatory variables. We have neither found any Bayes factors for the standard one-sided test $H_0 : \rho = 0$ versus $H_1 : 0 < \rho < 1$ nor for any multiple hypothesis tests. This is surprising, as the network autocorrelation parameter is at the heart of the model and testing for network effects is of crucial importance for network scientists when testing for and understanding theories of social influence. In sum, the objective of this article is to provide methodology that

- makes it possible to test multiple competing hypotheses regarding $\rho$ against each other and to precisely quantify the amount of evidence in favor of any of the hypotheses tested (including $H_0$),
- works for any combination of precise and/or interval hypotheses, and
- overcomes the problems with classical null hypothesis significance testing of $\rho$.

We also provide ready-to-use R code to make the methodology easily applicable for applied researchers.
In order to compute the Bayes factor, the so-called prior distributions, or simply priors, for the unknown model parameters have to be specified under each hypothesis. These priors quantify which values for the parameters are most likely before observing the data. For the testing problems considered in this work, the prior for the network autocorrelation $\rho$ under the alternative(s) is most important. In this article, we develop and explore several Bayes factors for testing the network effect: first, a Bayes factor based on an empirical informative prior which stems from an extensive literature review of empirical applications of the network autocorrelation model; second, a Bayes factor based on a uniform prior that assumes every value of $\rho$ to be equally likely a priori; third, a so-called fractional Bayes factor (O’Hagan 1995) which can be computed without needing to formulate a proper, that is, integrable, prior distribution for $\rho$ based on one’s prior beliefs. Subsequently, we conduct a simulation study to investigate the numerical properties of and differences between the proposed Bayes factors and then use the Bayes factors to reanalyze three data sets from the literature. Finally, we give efficient R code for the computation of the Bayes factors.

The article is organized as follows: In the second section, we discuss the network autocorrelation model in more detail and continue with a short introduction to Bayesian hypothesis testing in the third section. In the fourth section, we motivate several prior choices for the network autocorrelation parameter $\rho$. We assess the numerical performance of the Bayes factors in a simulation study in the fifth section and highlight their practical use with three examples in the sixth section. The seventh section concludes.

**The Network Autocorrelation Model**

Most social phenomena are embedded within networks of interdependencies. Building from a standard linear regression model, the network autocorrelation model effectively incorporates such interdependencies between individuals. In the model, the network structure is explicitly used to account for network influence and to estimate the magnitude of this influence. More specifically, the network influence is considered as a model parameter, the so-called network autocorrelation $\rho$, and operates directly on the outcome variable. Formally, the network autocorrelation model is expressed as

$$
y = \rho W y + X \beta + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I_g),$$

(1)

where $y$ is a $(g \times 1)$-vector of values of a dependent variable for the $g$ network actors, $X$ is a $(g \times k)$-matrix of values for the actors on $k$ covariates, $\beta$ is a $(k \times 1)$-vector of regression coefficients, $I_g$ denotes the
identity matrix, and $\mathbf{e}$ is a ($g \times 1$)-vector containing independent and identically normally distributed error terms with zero mean and variance of $\sigma^2$. Furthermore, $\mathbf{W}$ is a given ($g \times g$)-connectivity matrix representing social ties in a network, with $W_{ij}(i,j \in \{1,\ldots,g\})$ denoting the degree of influence of actor $j$ on actor $i$. Finally, the network autocorrelation $\rho$ is the key parameter of the model and quantifies the social influence for given $\mathbf{y}$, $\mathbf{W}$, and $\mathbf{X}$. We denote the resulting set of model parameters as $\mathbf{\theta} := (\rho, \sigma^2, \mathbf{\beta})$. Subsequently, we will repeatedly rely on the model’s likelihood, given by (e.g., Doreian 1980)

$$f(\mathbf{y}|\rho, \sigma^2, \mathbf{\beta}) = |\text{det}(A_{\rho})|(2\pi \sigma^2)^{-\frac{g}{2}}\exp\left(-\frac{1}{2\sigma^2}(A_{\rho}\mathbf{y} - X\mathbf{\beta})^T(A_{\rho}\mathbf{y} - X\mathbf{\beta})\right),$$

where $A_{\rho} := I_g - \rho \mathbf{W}$. To ensure that the model is well-defined, there are restrictions on the region of support for $\rho$. Typically, this region is chosen as the interval containing $\rho = 0$ for which $A_{\rho}$ is nonsingular (Hepple 1995; LeSage and Parent 2007; Smith 2009). In this case, the corresponding admissible interval for $\rho$ is given by $(\lambda^{-1}_g, \lambda^{-1}_1)$ where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_g$ are the ordered eigenvalues of $\mathbf{W}$ (Hepple 1995). For row-standardized connectivity matrices $\mathbf{W}$, that is, where each row sum is one, it holds that $\lambda_1 = 1$ (Anselin 1982). Without loss of generality, we restrict ourselves to such commonly used row-standardized connectivity matrices in the remainder. Hence, the model’s parameter space becomes $\Theta := \Theta_\rho \times \Theta_{\sigma^2} \times \Theta_{\mathbf{\beta}} = (\lambda^{-1}_g, 1) \times (0, \infty) \times \mathbb{R}^k$.

Throughout the literature, the model has also been named as mixed regressive–autoregressive model (Ord 1975), spatial effects model (Doreian 1980), network effects model (Marsden and Friedkin 1993), or spatial lag model (Anselin 2002), and it has been applied in many different fields such as sociology (Duke 1993; Kirk and Papachristos 2011; Mizruchi, Stearns, and Marquis 2006), political science (Beck, Gleditsch, and Beardsley 2006; Gimpel and Schuknecht 2003), criminology (Baller et al. 2001; Tita and Radil 2011), or geography (McMillen 2010; Mur, López, and Angulo 2008).

**Bayesian Hypothesis Testing**

In many network studies, researchers have expectations about the magnitude of the network effect. An interesting research question is whether the network effect can be classified as zero, small, medium, or large. These expectations can be translated to a set of multiple hypotheses on the network autocorrelation by setting $H_0 : \rho = 0$, $H_1 : 0 < \rho \leq .25$, $H_2 : .25 < \rho \leq .5$, and $H_3 : .5 < \rho < 1$, and the question to be
answered is which of these hypotheses is most plausible. In general, such a test is much more insightful than the standard test of no network effect versus “some” (positive) network effect, $H_0 : \rho = 0$ versus $H_1 : 0 < \rho < 1$. In order to illustrate this, consider a situation in which the estimated network autocorrelation parameter is $\hat{\rho} = .16$, with a 95 percent confidence interval for $\rho$ of $(-.06; .37)$, and a one-sided $p$ value of .08. Using the standard significance level of $\alpha = .05$, we would not reject the null hypothesis that $\rho = 0$ and would conclude that there is no statistically significant network effect present in the data. Based on the confidence interval, however, we do have quite a lot of confidence that the network effect may be positive. Hence, based on these classical outcomes, it is very difficult to state how plausible it is that the true network effect is zero, small, medium, or large, which was the initial research question of the network scientist.

The Bayes factor is a Bayesian hypothesis testing criterion that resolves this issue by providing a means to directly quantify how plausible each hypothesis is after observing the data. Suppose that we are interested in testing $T \geq 2$ hypotheses, $H_0, H_1, H_2, \ldots, H_{T-1}$. First, in Bayesian hypothesis testing, prior probabilities have to be assigned to both the model parameters under each hypothesis and to the hypotheses themselves. We denote these latter prior hypotheses probabilities by $p(H_0), p(H_1), \ldots, p(H_{T-1})$, with $\sum_{i=0}^{T-1} p(H_i) = 1$, which reflect how plausible we believe each hypothesis to be (relative to each other) before observing the data. There are multiple ways to assign prior hypotheses probabilities, for example, by assuming equal prior probabilities (reflecting prior ignorance), that is, $p(H_0) = \cdots = p(H_{T-1}) = 1/T$ (Hepple 1995; LeSage and Parent 2007), or by formulating specific prior probabilities for the various hypotheses. We will discuss procedures for eliciting prior probabilities of interval hypotheses in the fourth section.

Next, after observing the data $y$, Bayes’s theorem is applied to update the prior expectations with the information contained in the data. The resulting posterior probabilities of the hypotheses, $p(H_i | y)$, can then be written as:

$$p(H_i | y) = \frac{p(y | H_i) p(H_i)}{p(y)} = \frac{p(y | H_i) p(H_i)}{\sum_{j=0}^{T-1} p(y | H_j) p(H_j)}, \quad i \in \{0, 1, \ldots, T-1\}.$$  

(2)

These posterior probabilities quantify how probable each hypothesis is after observing the data, thus the quantity that researchers are typically
interested in. The higher the posterior probability of a particular hypothesis, the stronger the evidence in the data in favor of that hypothesis.

The term \( p(y|H_i) \) in equation (3) is called the marginal likelihood of the data \( y \) under the hypothesis \( H_i \) and denotes the probability that the data were observed under this specific hypothesis \( H_i \). It is computed by integrating the product of the model’s likelihood function and the prior distribution for the model parameters under \( H_i \). Hence, the marginal likelihood can be seen as a weighted likelihood over the parameter space under \( H_i \), with the prior under \( H_i \) acting as a weight function. In formal notation,

\[
p(y|H_i) = \int_{\Theta_i} f(y|\theta_i, H_i)p(\theta_i|H_i)d\theta_i, \tag{3}
\]

where \( \theta_i \) are the model parameters under \( H_i \), \( p(\theta_i|H_i) \) their prior density, and \( \Theta_i \) the corresponding parameter space. For instance, for \( H_0 : \rho = 0 \) and \( H_1 : 0 < \rho < 1 \) in the network autocorrelation model (1), it follows that \( \theta_0 = (\sigma^2, \beta), \Theta_0 = \mathbb{R}^+ \times \mathbb{R}^k \), and \( \theta_1 = (\rho, \sigma^2, \beta), \Theta_1 = (0, 1) \times \mathbb{R}^+ \times \mathbb{R}^k \).

When performing pairwise model comparisons between two hypotheses \( H_i \) and \( H_j \), we can consider the ratio of the corresponding two posterior probabilities of the hypotheses. In this case, the normalizing constant \( p(y) \) in equation (3) cancels out, and we can write

\[
\frac{p(H_i|y)}{p(H_j|y)} = \frac{p(y|H_i)}{p(y|H_j)} \times \frac{p(H_i)}{p(H_j)} := BF_{ij} \times \frac{p(H_i)}{p(H_j)}. \tag{4}
\]

The term \( p(H_i)/p(H_j) \) in equation (5) is called the prior odds of the two hypotheses and quantifies how much more, or less, likely the researcher expected \( H_i \) to be compared to \( H_j \) before observing the data \( y \). When a researcher does not a priori believe one being more likely than the other, the prior odds can be set equal to one. The term on the left-hand side of equation (5), \( p(H_i|y)/p(H_j|y) \), is known as the posterior odds, reflecting how much more (if larger than one) or less (if smaller than one) likely \( H_i \) is than \( H_j \) after taking the observed data into account. For example, if the posterior odds are five, this means that \( H_i \) is five times more likely than \( H_j \) for this data set. From equation (4), we can see that the posterior odds can be written as the prior odds multiplied by the Bayes factor, \( BF_{ij} \), which is defined as the ratio of the marginal likelihoods. Hence, the Bayes factor indicates to which extent the data change the prior odds to the posterior odds. Note that the Bayes factor can be used for quantifying the relative evidence between hypotheses without needing to specify how plausible the hypotheses are before
observing the data. If both models are considered as equally likely a priori, that is, \( p(H_i) = p(H_j) \), the Bayes factor equals the posterior odds.

The Bayes factor is a measure of relative evidence; it quantifies the amount of evidence in the data in favor of one hypothesis relative to another hypothesis. Jeffreys (1961) proposed a classification scheme to group Bayes factors into different categories (see Table 1). For example, there is “substantial” evidence for \( H_i \) when the Bayes factor exceeds 3 and, equivalently, substantial evidence for \( H_j \) when the Bayes factor is less than 1/3. These labels provide some rough guidelines when speaking of relative evidence in favor of a hypothesis, but the interpretation should ultimately depend on the context of the research question (Kass and Raftery 1995).

### Table 1. The Evidence Categories for the Bayes Factor (BF<sub>ij</sub>) as given by Jeffreys (1961).

<table>
<thead>
<tr>
<th>BF&lt;sub&gt;ij&lt;/sub&gt;</th>
<th>ln(BF&lt;sub&gt;ij&lt;/sub&gt;)</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;100</td>
<td>&gt;4.61</td>
<td>Decisive evidence for ( H_i )</td>
</tr>
<tr>
<td>30–100</td>
<td>3.40 to .61</td>
<td>Very strong evidence for ( H_i )</td>
</tr>
<tr>
<td>10–30</td>
<td>2.30 to 3.40</td>
<td>Strong evidence for ( H_i )</td>
</tr>
<tr>
<td>3–10</td>
<td>1.10 to 2.30</td>
<td>Substantial evidence for ( H_i )</td>
</tr>
<tr>
<td>1–3</td>
<td>0 to 1.10</td>
<td>Not worth more than a bare mention</td>
</tr>
<tr>
<td>1/3–1</td>
<td>−1.10 to 0</td>
<td>Not worth more than a bare mention</td>
</tr>
<tr>
<td>1/10–1/3</td>
<td>−2.30 to −1.10</td>
<td>Substantial evidence for ( H_j )</td>
</tr>
<tr>
<td>1/30–1/10</td>
<td>−3.40 to −2.30</td>
<td>Strong evidence for ( H_j )</td>
</tr>
<tr>
<td>1/100–1/30</td>
<td>−4.61 to −3.40</td>
<td>Very strong evidence for ( H_j )</td>
</tr>
<tr>
<td>&lt;1/100</td>
<td>&lt;−4.61</td>
<td>Decisive evidence for ( H_j )</td>
</tr>
</tbody>
</table>

Bayes Factor Tests for the Network Autocorrelation Parameter

In this section, we propose three Bayes factor tests when considering precise hypotheses, \( H_{\text{precise}} : \rho = c \), and interval hypotheses, \( H_{\text{interval}} : a_1 < \rho < a_2 \), for the network autocorrelation parameter. One of the most important steps in Bayesian hypothesis testing is the prior specification of the model parameters. In the network autocorrelation model (equation 1), a prior for \( \rho \) must be specified under an interval hypothesis, while no prior for \( \rho \) needs to be formulated under a precise hypothesis, as \( \rho \) is not a free parameter in this case. Despite its importance, prior specification of the network effect \( \rho \) has been largely neglected in the scarce literature on Bayesian hypothesis testing in the network autocorrelation model. The previous works
of Hepple (1995) and Han and Lee (2013) are exclusively based on a uniform prior for $\rho$ when testing the plausibility of different connectivity matrices and spatial model specifications, respectively. LeSage and Parent (2007) additionally employed a Beta prior for $\rho$ in a variable selection problem, while Sewell (2017) used a vague normal prior for $\rho$ when analyzing egocentric data. As these authors did not consider testing $\rho$ in particular, the prior choice was also less important in those contexts. In our setting, however, the prior for the network effect $\rho$ under the alternative(s) should be carefully chosen, as the Bayes factor can be sensitive to the prior for the tested parameter (Kass and Raftery 1995; Liu and Aitkin 2008; Sinharay and Stern 2002).

Prior expectations about the network autocorrelation parameter can be formulated based on a researcher’s beliefs or stem from previous empirical evidence from the literature. On the other hand, if the available prior information is weak or a researcher prefers to avoid adding prior information to an analysis, so-called noninformative priors are often used. Such noninformative priors are typically improper and are supposed to be completely dominated by the data (Gelman et al. 2013). In the following, we first present an empirical informative prior for $\rho$, second, a vague proper one, and third, an improper prior for the network effect. We combine these different marginal prior distributions for $\rho$ with the standard noninformative prior for the nuisance parameters $(\sigma^2, \beta), p(\sigma^2, \beta) \propto 1/\sigma^2$, assuming all parameters to be a priori independent (Hepple 1979, 1995; Holloway, Shankar, and Rahman 2002; LeSage 1997). However, note that the exact choice of the prior for the nuisance parameters hardly has an effect on the Bayes factor, as long as this prior is relatively vague (Kass and Raftery 1995).

The Bayes Factor Based on an Empirical Prior

In a review of published empirical applications of the network autocorrelation model, Dittrich et al. (2017) showed that medium network effects, for example, $\rho \approx .3$, are much more likely to be found in real-world networks than larger effects, for example, $\rho \approx .8$, or negative effects, for example, $\rho \approx -.2$. Dittrich et al. also showed that the distribution of empirically observed network effects is well-approximated by a normal distribution centered at .36 with a standard deviation of .19. Unless a new study considers a case that is fundamentally different from the networks studied in the literature at large to date, a network autocorrelation in a new study is likely to come from a population distribution for $\rho$ that resembles this normal distribution. This yields the following empirically motivated prior
\begin{align*}
p^E(\rho|\text{H}_{\text{interval}}) & \sim \mathcal{N}(0.36, 0.19^2)(a_1, a_2),
\end{align*}
\tag{5}
which is the aforementioned normal distribution with mean 0.36 and standard deviation 0.19, truncated to the corresponding parameter space of \(\rho\) under an interval hypothesis. The marginal likelihoods based on this empirical prior for \(\rho\) under precise and under interval hypotheses are given by (cf. Hepple 1995):
\begin{align*}
p^E(y|\text{H}_{\text{precise}}) = \pi \frac{g-k}{2} \Gamma \left( \frac{g-k}{2} \right) \sqrt{\frac{|(X^T X)^{-1}|}{A_c y^T A_c^T M A_c y}} \frac{g-k}{2},
\end{align*}
\tag{6}
\begin{align*}
p^E(y|\text{H}_{\text{interval}})
&= \pi \frac{g-k}{2} \Gamma \left( \frac{g-k}{2} \right) \sqrt{\frac{|(X^T X)^{-1}|}{A_p y^T A_p^T M A_p y}} \frac{g-k}{2} \\
&= \sqrt{2\pi} \times 0.19^2 \left( \Phi \left( \frac{a_2 - 0.36}{0.19} \right) - \Phi \left( \frac{a_1 - 0.36}{0.19} \right) \right) \\
&\quad \int_{a_1}^{a_2} \exp \left( - \frac{(\rho - 0.36)^2}{0.19^2} \right) |A_p y^T A_p^T M A_p y| \frac{g-k}{2} d\rho,
\end{align*}
\tag{7}
where \(M := I_g - X(X^T X)^{-1} X^T\) and \(\Phi(\cdot)\) denotes the cumulative distribution function of the standard normal distribution.

The unidimensional integral in equation (7) does not have a closed-form solution and has to be evaluated numerically. This can be done by relying on standard numerical methods, for example, Simpson’s rule (Atkinson 1989), and we present R code therefor in Online Appendix A, allowing researchers to use the Bayes factor without having to deal with the formulae themselves. Subsequently, the desired Bayes factors are obtained through equation (4) by using the marginal likelihoods of the precise and interval hypotheses under consideration.

The unconstrained version of the empirical prior in equation (5), that is, \(p(\rho) \sim \mathcal{N}(0.36, 0.19^2)\), can also be employed for determining prior probabilities of interval hypotheses. In this approach, the prior probabilities of the hypotheses are based on the probabilities of the tested model parameter
falling in the respective intervals under a proper unconstrained prior. As an example, consider the two interval hypotheses $H_1$:

\[ 0.25 < \rho < 0.5 \]

and $H_2$:

\[ 0.5 < \rho < 1 \]

The probabilities of $\rho$ falling in the intervals $(0.25, 0.5)$ and $(0.5, 1)$ under the unconstrained empirical prior are equal to 0.49 and 0.23, respectively (see Figure 1). These probabilities give a quantification of the plausibility of the hypotheses under the unconstrained empirical prior. Hence, the corresponding prior odds, $p(H_1)/p(H_2)$, of the two hypotheses in this example is 2.12 ($\approx 0.49/0.23$). In other words, under the empirical prior and before considering the data, a value of $\rho$ inside the interval $(0.25, 0.5]$ is 2.12 times as likely as $\rho$ being inside $(0.5, 1)$. Thus, if we assume that either $H_1$ or $H_2$ is true and that their prior odds corresponds to 2.12, the prior probabilities of the hypotheses are $p(H_1) = 0.68 \left( \approx 0.49/(0.49 + 0.23) \right)$ and $p(H_2) = 0.32 \left( \approx 0.23/(0.49 + 0.23) \right)$. This seems reasonable as medium effects ($H_1$) are generally more plausible than large effects ($H_2$) in social network research (Dittrich et al. 2017). To the best of our knowledge, using an unconstrained prior for specifying prior odds of interval hypotheses is novel in the literature.

In the remainder of this article, we rely on the following method to assign prior model probabilities when testing one precise null hypothesis and $T - 1$ interval hypotheses. As there are $T$ hypotheses in total, we set a prior probability of $1/T$ to the precise null hypothesis $H_0$. Subsequently, the remaining probability of $(T - 1)/T$ is divided upon the interval hypotheses $H_1, ..., H_{T-1}$ using the prior probabilities of the intervals under a proper

![Figure 1. The probability density function of the unconstrained empirical prior for $\rho$, $p(\rho) \sim N(0.36, 0.19^2)$](image)

The areas of the shaded regions under the probability density function correspond to the probabilities of $\rho$ falling in the interval $(0.25, 0.5]$ (black) and $(0.5, 1)$ (gray) which are equal to 0.49 and 0.23, respectively.
unconstrained prior. For instance, consider $H_0: \rho = 0$, $H_1: 0 < \rho < .25$, $H_2: .25 < \rho \leq .5$, and $H_3: .5 < \rho < 1$. Here, we test five hypotheses in total, so $H_0$ is assigned a prior probability of .2. The remaining probability of $4/5 = .8$ is split between the four interval hypotheses based on the probability masses contained in the four intervals under the unconstrained empirical prior. For the hypotheses considered above, the probabilities of $\rho$ falling in the respective intervals are $p(H_1) = .03$, $p(H_2) = .25$, $p(H_3) = .49$, and $p(H_4) = .23$. As we have already set $p(H_0) = .2$, there is a total probability of .8 left for the remaining hypotheses. Rescaling and making them add up to .8 results in prior model probabilities of $H_1$, $H_2$, $H_3$, and $H_4$ of .02, .20, .39, and .19. Finally, when combining the prior probabilities of the hypotheses and the marginal likelihoods in equations (6) and (7), the corresponding posterior model probabilities can be calculated via equation (2).

The Bayes Factor Based on a Uniform Prior

The uniform prior treats all possible network effects under the alternative(s) as equally likely before observing the data $y$, resulting in a flat prior distribution for $\rho$. As the region of support for $\rho$ is bounded, the uniform prior for the network autocorrelation is a vague proper prior. Hence, it is less informative than the empirical prior but also does not represent complete prior ignorance which is typically expressed by using improper priors. The uniform prior under an interval hypothesis can be written as:

$$p^U(\rho|H_{\text{interval}}) \sim \mathcal{U}(a_1, a_2),$$

where $\mathcal{U}(a_1, a_2)$ denotes the uniform distribution on $(a_1, a_2)$. The marginal likelihood under $H_{\text{precise}}$ remains the same as in equation (6), and the marginal likelihood under an interval hypothesis $H_{\text{interval}}$ in combination with a uniform prior for $\rho$ is now given by:

$$p^U(y|H_{\text{interval}}) = \pi^{-\frac{g-k}{2}} \Gamma \left( \frac{g-k}{2} \right) \sqrt{|(X^TX)^{-1}|} \int p^U(\rho|H_{\text{interval}})|A_\rho|y^T A_\rho^T M A_\rho y^{-\frac{g-k}{2}} d\rho$$

$$= \pi^{-\frac{g-k}{2}} \Gamma \left( \frac{g-k}{2} \right) \sqrt{|(X^TX)^{-1}|} \frac{1}{a_2 - a_1} \int_{a_1}^{a_2} |A_\rho|y^T A_\rho^T M A_\rho y^{-\frac{g-k}{2}} d\rho.$$
Again, the Bayes factor is computed as the ratio of two marginal likelihoods under the uniform prior for $\rho$. Similarly as for the empirical prior, an unconstrained uniform prior can also be used for specifying prior model probabilities. In the uniform prior setting, this implies that the prior odds of two interval hypotheses is equal to the ratio of the lengths of the intervals. Posterior probabilities of the hypotheses can then be obtained via equation (2) and by plugging in the marginal likelihoods and prior probabilities of the hypotheses based on the uniform prior.

**The Fractional Bayes Factor**

Finally, there may be situations where a researcher does not have any prior beliefs about possible network effects if the null would be false or where a researcher prefers not to specify a proper prior based on external knowledge. Such prior ignorance is usually reflected by employing improper priors. However, if improper priors for the tested parameters are imposed, the Bayes factor depends on unknown normalizing constants and is not well-defined (O’Hagan 1995). To that end, the fractional Bayes factor methodology was originally proposed by O’Hagan (1995) as a way to bypass this issue. In the fractional Bayes factor, the main idea is to split the information in the data into two different fractions that sum up to one. The first (small) fraction, denoted by $b$, is used for updating the initial improper prior into a proper default prior (Mulder 2014b); the second fraction, $1 - b$, is taken for evaluating the hypotheses under investigation based on the default prior. In mathematical notation, this comes down to rewriting the marginal likelihood in equation (3) as:

$$p(y|H_i, b) = \frac{\int_{\Theta_i} f(y|\theta_i, H_i)p^N(\theta_i|H_i)d\theta_i}{\int_{\Theta_i} f(y|\theta_i, H_i)^b p^N(\theta_i|H_i)d\theta_i}$$

$$= \int_{\Theta_i} f(y|\theta_i, H_i)^{1-b} p(\theta_i|H_i, y^b)d\theta_i,$$

where $p^N(\theta_i|H_i)$ denotes a noninformative improper prior density for $\theta_i$ under $H_i$ and $p(\theta_i|H_i, y^b) = f(y|\theta_i, H_i)^b p^N(\theta_i|H_i)/\int_{\Theta_i} f(y|\theta_i, H_i)^b p^N(\theta_i|H_i)d\theta_i$ is the updated, proper default prior. Thus, in order to compute the marginal likelihood in the network autocorrelation model in the fractional Bayes factor approach, one needs to choose a noninformative improper prior for $\rho$ and to specify the fraction $b$. As improper prior for $\rho$ we use,
\[ p^N(\rho | \text{H}_{\text{interval}}) = (1 - \rho)^{-1} \chi_{(0,1)}(\rho). \]  

(10)

This prior approximates the model’s independence Jeffreys prior very well (Online Appendix C), which has been shown to outperform the standard uniform prior for \( \rho \) in Bayesian estimation of the model (Dittrich et al. 2017). At the same time, the independence Jeffreys prior itself also imposes an a priori dependence structure between the model parameters (Dittrich et al. 2017), which makes it difficult to compare its inferences to those based on the previously proposed normal and uniform marginal priors for \( \rho \). For this reason, we rely on the marginal improper prior for \( \rho \) in equation (10) as it relaxes this a priori dependence. Note that in the fractional Bayes factor approach, we consider \( \text{H}_{\text{interval}} : 0 < \rho < 1 \) as the only alternative interval hypothesis. Else, if a researcher had more precise expectations about the magnitude of the network effect, for example, \( .25 < \rho < .5 \), it seems much more sensible to use a proper prior for \( \rho \) instead.

Typically, the fraction \( b \) in the fractional Bayes factor is chosen as the smallest value for which the updated default prior in equation (9) is proper (Berger and Mortera 1999; Mulder 2014b; O’Hagan 1995). This choice results in maximal possible use of the information in the data for hypothesis testing. If the proposed improper prior in equation (10) is combined with the standard noninformative prior for \((\sigma^2, \beta)\), \( p(\sigma^2, \beta) \propto 1/\sigma^2 \), the corresponding updated prior in equation (9) is proper if \( b > k/g \) (see Online Appendix D for a proof). The proof shows that this also holds for the updated prior in equation (9) under \( \text{H}_{\text{precise}} \). We denote the resulting choice for \( b \) as \( b_1 = (k+1)/g \). On the other hand, if misspecification of the improper prior is a concern, larger values for \( b \) may be preferred, as they can reduce the sensitivity of the fractional Bayes factor to prior misspecification (Conigliani and Anthony 2000; O’Hagan 1995). As empirical network autocorrelations are more likely to come from the estimated unconstrained population distribution for \( \rho \) in equation (5) than from a distribution that resembles the improper prior in equation (10), prior misspecification is indeed a valid concern. In this case, O’Hagan (1995) suggested to use \( b_2 = \max(k+1; \ln(g))/g \), which makes the fractional Bayes factor more robust but increases slowly with \( g \), or \( b_3 = \max(k+1; \sqrt{g})/g \), when sensitivity to misspecification of the prior is a serious concern. Finally, the marginal likelihoods under \( \text{H}_{\text{precise}} \) and \( \text{H}_{\text{interval}} \) in the fractional Bayes factor approach are given by:
\[
p(y|H_{\text{precise}}, \theta) = b^2 \pi \frac{\Gamma\left(\frac{g - k}{2}\right)}{\Gamma\left(\frac{gb - k}{2}\right)} |A_c|^{1-b} y^T A_c^T M A_c y^\frac{g(b-1)}{2}, \tag{11}
\]

\[
p(y|H_{\text{interval}}, \theta) = b^2 \pi \frac{\Gamma\left(\frac{g - k}{2}\right)}{\Gamma\left(\frac{gb - k}{2}\right)} \int p^N(\rho|H_{\text{interval}}) |A_\rho|^b y^T A_\rho^T M A_\rho y^\frac{g - k}{2} d\rho
\]

\[
= b^2 \pi \frac{\Gamma\left(\frac{g - k}{2}\right)}{\Gamma\left(\frac{gb - k}{2}\right)} \int_0^1 (1 - \rho)^{-\frac{g - k}{2}} |A_\rho|^b y^T A_\rho^T M A_\rho y^\frac{g - k}{2} d\rho. \tag{12}
\]

For more technical details about the computation of equation (12), we refer the reader to Online Appendix E. As before, the marginal likelihoods in equations (11) and (12) are used to calculate the fractional Bayes factor itself, which is again the ratio of two marginal likelihoods. If a researcher is also interested in obtaining posterior model probabilities, a default choice is to impose equal prior model probabilities, that is, \( p(H_0) = p(H_1) = .5 \), while subjective prior model probabilities could be assigned if a researcher has clear beliefs about the plausibility of the null hypothesis of a zero network effect.

**Simulation Study**

In this section, we present results of a simulation study investigating the performances of and the differences between the Bayes factors discussed in the fourth section. To that end, we first look at the one-sided test A: \( H_0 : \rho = 0 \) versus \( H_1 : 0 < \rho < 1 \). In particular, we explore which Bayes factor converges fastest to the true hypothesis and assess the sensitivity of the fractional Bayes factor to the choice of \( b \). Next, we consider a multiple hypothesis test B with a precise hypothesis and four interval hypotheses: \( H_0 : \rho = 0 \) versus \( H_1 : -.25 < \rho < 0 \) versus \( H_2 : 0 < \rho \leq .25 \) versus \( H_3 : .25 < \rho \leq .5 \) versus \( H_4 : .5 < \rho < 1 \). Our primary goal
here is to study how fast the posterior model probabilities converge to the true model for different data-generating hypotheses and to check how robust these findings are to different marginal priors for \( \rho \) and to different prior model probabilities.

**Study Design**

In order to mimic realistic networks, we consider two different approaches for obtaining connectivity matrices \( W \) in test A. First, we use the well-known small-world structure (Watts and Strogatz 1998) to generate simulated networks. Such networks are highly clustered, that is, there is a tendency that network actors create very dense subnetworks but at the same time have small path lengths, that is, there is a high probability that any two actors in the network are connected by short paths of acquaintances (Watts and Strogatz 1998). Typically, this results in networks where most actors are linked to only a few others, while some actors, also known as *hubs*, have a lot of ties. These hubs function as connectors between different subnetworks and shorten the path length between two actors in the entire network. Small-world structures can be observed in online social networks (Fu et al. 2007), scientific collaboration networks (Newman 2001), or corporate elite networks (Davis, Yoo, and Baker 2003). We obtain simulated small-world networks relying on the “watts.strogatz.game” function from the igraph package in R (Version number is 1.0.1) (Csardi and Nepusz 2006). In the underlying algorithm, a ring lattice of \( g \) actors, each connected to its \( d \) nearest neighbors by undirected ties, is constructed first. Next, with probability \( r \), each tie in the network is randomly rewired. Following Neuman and Mizruchi (2010) and Wang, Neuman, and Newman (2014), we set \( r = .1 \) that leads to highly clustered networks with low average path lengths. In our simulation study, we consider 14 network sizes \((g \in \{25, 50, 75, 100, 150, 200, 300, \ldots, 1000\})\) and two different mean degrees \((d \in \{4, 8\})\). The simulated connectivity matrices are binary, that is, \( W_{ij} = 1(i, j \in \{1, \ldots, g\}) \) if there is a tie between actor \( i \) and \( j \) and zero otherwise. Subsequently, we row-normalize the generated symmetric raw connectivity matrices.  

Second, we also run simulations using two prominent contiguity-based spatial networks; the 49 neighborhoods in Columbus, OH, analyzed in, for example, Anselin (1988), Elhorst (2014), and Hepple (1995) and the 64 Louisiana parishes studied in Doreian (1980), Howard (1971), and Leenders (2002), among others. In general, networks based on spatial proximity do not exhibit small-world properties, as there may be no short path between two distant nodes. Thus, we rely on these two real-world networks to gain
insights into the behavior of the Bayes factors for network configurations that are different from typical small-world structures. We set $W_{ij}(i,j \in \{1,\ldots,g\})$ to one if area $i$ is adjacent to area $j$, to zero otherwise, and row-standardize the raw adjacency matrices. Graphical representations of the two real networks and an example of a simulated small-world network appear in Figure 2.

For each of the network types, we include three covariates plus an intercept term (so $k = 4$) and use three fixed levels for the network effect size ($\rho \in \{0, .25, .5\}$) to generate $y$ via $y = (I_g - \rho W)^{-1}(X\beta + \varepsilon)$ for test A. Furthermore, we also sample network effects from the estimated empirical population distribution from the estimated empirical population distribution of the network effects truncated to $(0,1)$, rather than fixing $\rho$ to a single-specific value. As the true network autocorrelation is unknown in practice, this is a more realistic setup than choosing specific values for $\rho$ a priori. Finally, we draw independent values from a standard normal distribution for the elements of $X$ (excluding the first column of $X$ which is a vector of ones), $\beta$, and $\varepsilon$. Hence, we consider 120 scenarios (14 small-world networks $\times$ 2 mean degrees $\times$ 4 sampling schemes for $\rho$ and 2 real networks $\times$ 4 sampling schemes for $\rho$) for test A in total and simulate 1,000 data sets for each scenario.

For test B, we generate network effects using the empirical prior for $\rho$, truncated to the corresponding parameter space under each hypothesis. We assign prior probabilities to each hypothesis based on the unconstrained versions of both the empirical and the uniform prior for $\rho$. We draw values for the elements of $X$, $\beta$, and $\varepsilon$ as described in the previous paragraph, so for test B we examine 140 scenarios (14 small-world networks $\times$ 2 mean degrees

Figure 2. A simulated small-world network with $g = 50$, $d = 4$, $r = .1$ (left), the network of the 49 Columbus, OH, neighborhoods (middle), and the network of the 64 Louisiana parishes (right).
× 5 data-generating hypotheses) in total, for which we generate 1,000 data sets for each.

Simulation Results

Figure 3 and Table 2 show the average weight of evidence, that is, the natural logarithm of the Bayes factor (Good 1985), for the different Bayes factors and network structures for test A: $H_0 : \rho = 0$ versus $H_1 : 0 < \rho < 1$.9

We can conclude the following from these results. First, the Bayes factor based on the empirical prior, $BF_{\text{E10}}$ (thick solid line), and the Bayes factor based on the uniform prior, $BF_{\text{U10}}$ (thin solid line), show consistent behavior, that is, the evidence for the true hypothesis is increasing with the network size. In addition, they almost always provide most evidence for the correct hypothesis, except for $\rho = .25$ and some smaller network sizes in the small-world networks. Second, the evidence for a true alternative hypothesis grows with the network size at a faster rate than the evidence for a true null as is common in other statistical models (Johnson and Rossell 2010) and for precise hypotheses in general. Third, the Bayes factor based on the empirical prior results in slightly even more evidence for a true hypothesis than the Bayes factor based on the uniform prior. Fourth, the smaller the mean degree, the bigger the evidence for a true alternative hypothesis provided by both $BF_{\text{E10}}$ and $BF_{\text{U10}}$, for fixed $g$ and $\rho$. This behavior is due to the negative bias of $\rho$ in the model for increasing network densities (Mizruchi and Neuman 2008; Neuman and Mizruchi 2010; Smith 2009). Fifth, the fractional Bayes factors based on $b_1$ and $b_2$ (dashed line and dotted line) yield very similar results. Overall, they provide less evidence for the alternative hypothesis compared to $BF_{\text{E10}}$ or $BF_{\text{U10}}$ and appear to be biased toward the null. For example, this bias is manifested from the fact that networks of approximately 300 nodes are needed before the fractional Bayes factors based on $b_1$ and $b_2$ result in evidence for the true hypothesis $H_1$ when $\rho$ equals .25 and $d = 8$ (see Figure 3). By way of comparison, the Bayes factors based on the empirical and the uniform prior already point toward evidence for the true alternative for small networks in this scenario. Sixth, the fractional Bayes factor based on $b_3$ (dot-dashed line) does not show consistent behavior when $H_0$ is true. In particular, the evidence for the true null hypothesis does not increase with the network size but remains constant or in some cases even decreases.

In order to provide more insights into the behavior of the fractional Bayes factor, we investigate its behavior more thoroughly. To that end, we observe that for small values of $b$, the updated marginal prior for $\rho$ in the fractional Bayes factor approach in equation (9) is just proper with most of its
Figure 3. The average weight of evidence based on $BF_E^{10}$ (thick solid line), $BF_U^{10}$ (thin solid line), $BF_{b1}^{10}$ (dashed line), $BF_{b2}^{10}$ (dotted line), and $BF_{b3}^{10}$ (dot-dashed line) for test A: $H_0: \rho = 0$ versus $H_1: 0 < \rho < 1$ for simulated data using generated small-world networks over 1,000 replications.
Table 2. The Average Weight of Evidence Based on $BF_{10}^E$, $BF_{10}^U$, $FBF_{10}^{b_1}$, and $FBF_{10}^{b_3}$ for Test A: $H_0: \rho = 0$ Versus $H_1: 0 < \rho < 1$ for Simulated Data Using the Network of the 49 Columbus, OH, Neighborhoods and the Network of the 64 Louisiana Parishes Over 1,000 Replications.

<table>
<thead>
<tr>
<th>Network</th>
<th>$\rho = 0$</th>
<th>$\rho = 0.25$</th>
<th>$\rho = 0.5$</th>
<th>$\rho \sim p^E(\rho)$</th>
<th>$\rho = 0$</th>
<th>$\rho = 0.25$</th>
<th>$\rho = 0.5$</th>
<th>$\rho \sim p^E(\rho)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln(BF_{10}^E)$</td>
<td>$-1.7$</td>
<td>$1.0$</td>
<td>$7.4$</td>
<td>$4.8$</td>
<td>$\ln(BF_{10}^U)$</td>
<td>$-1.7$</td>
<td>$0.7$</td>
<td>$7.0$</td>
</tr>
<tr>
<td>$\ln(BF_{10}^U)$</td>
<td>$-1.7$</td>
<td>$0.7$</td>
<td>$7.0$</td>
<td>$4.5$</td>
<td>$\ln(FBF_{10}^{b_1})$</td>
<td>$-2.8$</td>
<td>$-0.5$</td>
<td>$6.0$</td>
</tr>
</tbody>
</table>

$\ln(BF_{10}^E)/C_0 = 1.7 1.0 7.4 4.8$  
$\ln(BF_{10}^U)/C_0 = 1.7 0.7 7.0 4.5$  
$\ln(FBF_{10}^{b_1})/C_0 = 2.8 0.5 6.0 3.4$  
$\ln(FBF_{10}^{b_3})/C_0 = 2.0 0.2 6.2 3.8$
probability mass still at values close to one for which the likelihood function is vanishing (see Figure 4). As the marginal likelihood under $H_1$ in the fractional Bayes factor approach is essentially an average weighted likelihood over $[0, 1]$, with the updated prior acting as weight function, the fractional Bayes factors based on $b_1$ and $b_2$ tend to favor the null by construction. On the other hand, for large values of $b$, the updated marginal prior for $\rho$ exhibits a local maximum near the maximum likelihood estimate of $\rho$ (see Figure 4). When the data are generated under $H_0$, this results in a relatively larger average weighted likelihood over $(0, 1)$ and considerable support for the alternative which explains the inconsistent behavior of the fractional Bayes factor based on the fraction $b_3$ when the null is true.

Given the results from our simulation study, we recommend using the Bayes factor based on the empirical prior for $\rho$ when testing $H_0 : \rho = 0$ versus $H_1 : 0 < \rho < 1$ or the Bayes factor based on the uniform prior as a second reasonable alternative. We do not recommend any of the fractional Bayes factors for this test, as they are either biased toward the null when the data are generated under the alternative or they show inconsistent behavior when $H_0$ is true. As a result, the fractional Bayes factors provide little improvement in classical null hypothesis significance testing, whereas the Bayes factors based on the empirical and uniform prior for $\rho$ clearly do perform considerably well.

Figure 5 reports the average posterior model probabilities for test B based on the empirical prior, $p^E(H_i | y)$, $i \in \{0, 1, 2, 3, 4\}$ (left column), and the
Figure 5. The average posterior model probabilities of $H_0$ (solid line), $H_1$ (dashed line), $H_2$ (dotted line), $H_3$ (dot-dashed line), and $H_4$ (long-dashed line) for test B: $H_0 : \rho = 0$ versus $H_1 : -0.25 < \rho < 0$ versus $H_2 : 0 < \rho \leq 0.25$ versus $H_3 : 0.25 < \rho \leq 0.5$ versus $H_4 : 0.5 < \rho < 1$, based on the empirical prior for $\rho$, $p^E(H_i|y)$, $i \in \{0, 1, 2, 3, 4\}$, (left column) and based on the uniform prior for $\rho$, $p^U(H_i|y)$, (right column) for simulated data using generated small-world networks over 1,000 replications.
ones based on the uniform prior, $p^U(H_i|y)$ (right column), for different network sizes and data-generating hypotheses.\(^\text{10}\)

In general, for both $p^E(H_i|y)$ and $p^U(H_i|y)$, the evidence for the true hypothesis is increasing with the sample size in all scenarios. Furthermore, the data-generating hypothesis always receives the most support, except for some very small network sizes and when the data are based on negative network effects in combination with the empirical prior. In the latter case, the prior probability of $H_1, p^E(H_1) = .02$, is very small compared to the one of $H_0, p^E(H_0) = .20$. This means that the data need to support $H_1$ approximately 10 times more than $H_0$ such that the hypotheses receive at least equal posterior probability. However, we believe this behavior of $p^E(H_1|y)$ not to be a real concern, as the empirical literature suggests that it is highly unlikely to observe negative values for $\rho$ in practice (Dittrich et al. 2017). Else, if a researcher deems such negative network autocorrelations to be more plausible than implied by $p^E(H_1) = .02$, $H_1$ should be accordingly given a higher prior probability.

**Empirical Examples**

In the following, we apply the Bayes factor based on the empirical prior for $\rho$ to three data sets from the literature to quantify the relative evidence for different competing hypotheses of interest. We test the five hypotheses $H_0: \rho = 0$, $H_1: -.25 < \rho < 0$, $H_2: 0 < \rho \leq .25$, $H_3: .25 < \rho \leq .5$, and $H_4: .5 < \rho < 1$ against each other, corresponding to the notion of “no network effect,” a “minor negative network effect,” a “minor (positive) network effect,” a “medium network effect,” and a “large network effect,” respectively. We assign prior probabilities to these hypotheses using the unconstrained empirical prior from the fourth section, which yields $p^E(H_0) = .20, p^E(H_1) = .02, p^E(H_2) = .20, p^E(H_3) = .39$, and $p^E(H_4) = .19$. In order to check robustness to the choice of the prior for $\rho$ and to the choice of the prior model probabilities, we also perform our analyses using a uniform prior for the network autocorrelation parameter and for specifying prior model probabilities via the uniform unconstrained prior, that is, $p^U(H_0) = .20, p^U(H_1) = p^U(H_2) = p^U(H_3) = .16$ and $p^U(H_4) = .32$. Finally, we compare the results to those coming from classical tests using $p$ values.

**Crime Data**

In the cross-sectional data set for 49 neighborhoods in Columbus, OH, first analyzed by Anselin (1988), the network autocorrelation model is used to
explain the 1980 neighborhoods’ crime rates, operationalized as the combined total of residential burglaries and vehicle thefts per 1,000 households. This data set is openly accessible as part of the Columbus data from the R package spdep (Bivand and Piras 2015). Figure 6 (left panel) shows the spatial distribution of the crime rates across the 49 neighborhoods.

Anselin (1988) modeled these crime rates as a function of household income and housing value (in 1,000USD$) plus an intercept term. In his study, both explanatory variables had a negative impact on the crime rate, while the maximum likelihood estimate of the network effect was $\hat{\rho}_{ML} = 0.40$, with a 95 percent confidence interval for $\rho$ of $(0.17; 0.64)$, and $p = 0.0008^{11}$.

When using the empirical prior for $\rho$, the posterior probabilities of the hypotheses are $p^E(H_0|y) = 0.01$, $p^E(H_1|y) = 0.00$, $p^E(H_2|y) = 0.11$, $p^E(H_3|y) = 0.74$, and $p^E(H_4|y) = 0.14$. In other words, $H_3$ is by far the most likely hypothesis and approximately 83 ($\approx 0.74/0.01$), 1,744, 7, and 6 times more plausible than $H_0$, $H_1$, $H_2$, and $H_4$, respectively. Furthermore, the Bayes factor of $H_3$ versus $H_0$ is 42.3, which is considered as very strong evidence in the data for $H_3$, see Table 3, and the Bayes factor of $H_3$ versus $H_4$ (the second most supported hypothesis) is 2.6, which implies minor evidence for a medium effect relative to a large effect. R code for computing these posterior model probabilities and corresponding Bayes factors is provided in Online Appendix A. When relying on the uniform prior for $\rho$, the posterior model probabilities are $p^U(H_0|y) = 0.02$, $p^U(H_1|y) = 0.00$, $p^U(H_2|y) = 0.14$, $p^U(H_3|y) = 0.64$, and $p^U(H_4|y) = 0.20$. Again, $H_3$ is by far the most likely hypothesis of the five, followed by $H_4$ which is

---

**Figure 6.** The number of residential burglaries and vehicle thefts per 1,000 households across 49 neighborhoods in Columbus, OH, in 1980 (left), and the logarithm of voter turnout for the 1980 U.S. presidential election across 3,076 U.S. counties (right). The shading color of an entity indicates to which quintile of the sample it belongs.
approximately three times as unlikely as $H_3$. Thus, in line with the results from classical maximum likelihood–based inference, there is very strong evidence for a positive network effect. Contrary to maximum likelihood–based inference, however, the Bayesian approach also allows one to draw conclusions as to how much support there is in the data for particular values of $\rho$. In these data, evidence is strong that the network effect resides between .25 and .5, a conclusion that is affected neither by the choice of the prior for $\rho$ nor by the choice of the prior model probabilities. Ultimately, among these Columbus neighborhoods, there is the most evidence for medium autocorrelation (between .25 and .5) with respect to crime rates. Table 3 summarizes the findings.

**Threatened Birds Data**

In the following example, McPherson and Nieswiadomy (2005) studied the percentage of threatened birds in 113 countries around the globe in the year 2000 via the network autocorrelation model, considering that “threats to a species in one country may spill over to neighboring countries’ species” (McPherson and Nieswiadomy 2005:401). In this data set, the spatial connectivity matrix was based on the shared border length between two neighboring countries which made several island countries isolates in the network. In addition, the authors included 10, mainly socioeconomic, explanatory variables in the analysis plus an intercept term. The resulting maximum likelihood estimate of the network autocorrelation was $\hat{\rho}_{\text{ML}} = .16$, with a 95 percent confidence interval for $\rho$ of $(-.06; .37)$, and $p = .08$, so “in other words, threats to birds spill over into adjoining countries” (McPherson and Nieswiadomy 2005:405).

Based on the empirical prior for $\rho$, the data yield posterior model probabilities of $p^E(H_0|y) = .26$, $p^E(H_1|y) = .01$, $p^E(H_2|y) = .50$, $p^E(H_3|y) = .23$, and $p^E(H_4|y) = .00$. Hence, the hypothesis that there is a minor spillover effect of threats to birds across adjoining countries, that is, $H_2: 0 < \rho \leq .25$, is most supported by the data and consequently results in the highest Bayes factor (see Table 3). However, the Bayes factor of a minor spillover effect relative to no spillover effect is only 2.0, so the support in the data for $H_2$ is far from decisive. In case of considering a uniform for $\rho$, the resulting posterior probabilities of the hypotheses are similar and given by $p^U(H_0|y) = .32$, $p^U(H_1|y) = .07$, $p^U(H_2|y) = .49$, $p^U(H_3|y) = .12$, and $p^U(H_4|y) = .00$. Again, none of the alternative hypotheses receives convincing evidence to outweigh $H_0$. Overall, the data provide most evidence for $\rho$ being between 0 and .25, but the Bayes factors
Table 3. The Bayes Factors $\text{BF}_E^i$, $\text{BF}_U^i$, $i \in \{0, 1, 2, 3, 4\}$, the Posterior Model Probabilities $p^E(H_i|y)$ and $p^U(H_i|y)$ of $H_0: \rho = 0$, $H_1: -.25 < \rho < 0$, $H_2: 0 < \rho \leq .25$, $H_3: .25 < \rho \leq .5$, and $H_4: .5 < \rho < 1$, and the ML Estimates $\hat{\rho}_{ML}$ of $\rho$ and Corresponding $p$ Values for the Crime Data Set, the Threatened Birds Data Set, and the Voting Data Set.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Crime Data</th>
<th>Threatened Birds Data</th>
<th>Voting Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$H_0$</td>
<td>$H_1$</td>
<td>$H_2$</td>
</tr>
<tr>
<td>$\text{BF}_E^0$</td>
<td>1</td>
<td>.4</td>
<td>12.5</td>
</tr>
<tr>
<td>$\text{BF}_U^0$</td>
<td>1</td>
<td>.2</td>
<td>9.4</td>
</tr>
<tr>
<td>$p^E(H_0</td>
<td>y)$</td>
<td>.01</td>
<td>.00</td>
</tr>
<tr>
<td>$p^U(H_0</td>
<td>y)$</td>
<td>.02</td>
<td>.00</td>
</tr>
<tr>
<td>ML-based</td>
<td>$\hat{\rho}_{ML} = .40$, $p = .0008$</td>
<td>$\hat{\rho}_{ML} = .16$, $p = .08$</td>
<td>$\hat{\rho}_{ML} = .61$, $p &lt; 10^{-6}$</td>
</tr>
</tbody>
</table>

Note: ML = maximum likelihood.
also show that the strength of the evidence is rather small. These findings vividly illustrate the well-known issue that $p$ values tend to overestimate the evidence against the null (Berger and Sellke 1987; Jeffreys 1961; Rouder et al. 2009; Sellke, Bayarri, and Berger 2001; Wagenmakers 2007).

**Voting Data**


$$y = \rho Wy + \alpha 1 + X \beta + WX \gamma + \epsilon$$

(where $\alpha$ denotes the model’s intercept term, $1$ a vector of ones, and $\gamma$ another vector of regression coefficients), to analyze the logarithm of the voter turnout ($\text{TURNOUT}$) across these U.S. counties. The spatial distribution of the logarithm of the voter turnout is shown in Figure 6 (right panel). As explanatory variables, the authors used the logarithms of, first, the population of 18 years of age or older eligible to vote in each county ($\text{POP}$); second, the population of 25 years of age or older with a twelfth grade or higher education in each county ($\text{EDUCATION}$); third, the number of owner-occupied housing units in each county ($\text{HOUSES}$); fourth, the aggregate income of each county ($\text{INCOME}$), and the four corresponding spatially lagged variables. Furthermore, the adjacency matrix $W$ in this example was constructed on the basis of the four nearest neighbors of each county. Except for $\log(\text{POP})$, and the spatially lagged $\log(\text{HOUSES})$ and $\log(\text{INCOME})$ variables, all other predictors were found to have a positive impact on $\log(\text{TURNOUT})$, with a maximum likelihood estimate of $\hat{\rho}_{\text{ML}} = .62$ (Pace and Barry 1997:242), an associated 95 percent confidence interval for $\rho$ of (.58; .65), and $p < 10^{-6}$.

Computing the corresponding posterior model probabilities underpins the decisive evidence for a large network effect, as

$$p^E(H_0|y) = .00, \quad p^E(H_1|y) = .00, \quad p^E(H_2|y) = .00, \quad p^E(H_3|y) = .00, \quad \text{and} \quad p^E(H_4|y) = 1.$$ 

Accordingly, the concomitant Bayes factor of $H_4$ versus any other considered hypothesis exceeds $10^6$ which provides “decisive” evidence in favor of $H_4$ compared to $H_0$, $H_1$, $H_2$, and $H_3$, respectively. When using the uniform prior, the posterior model probabilities remain essentially unchanged, as the evidence in the data for a large network effect is conclusive in this example. Although the implications of the Bayesian approach seem in line with the traditional approach, this is not the case completely: The only thing we can deduce from classical null hypothesis significance testing here is that we can reject the null hypothesis that $\rho = 0$. On the other hand, the Bayesian
approach gives us much more detail about which values of $\rho$ are most supported by the data and quantifies to what extent.

Conclusions

In this article, we developed three Bayes factors for testing precise and interval hypotheses for the network effect in the network autocorrelation model. The Bayesian approach to these tests comes with several practical advantages compared to classical null hypothesis significance testing. For instance, the Bayes factors and the resulting posterior model probabilities allow us to quantify the amount of evidence for a precise null hypothesis, or any other hypothesis, and they allow us to test multiple precise and interval hypotheses simultaneously without any of the drawbacks of classical null hypothesis significance testing. We ran an extensive simulation study to evaluate the numerical behavior of the presented Bayes factors for a wide range of different network configurations. We found that the Bayes factor based on an empirical prior for the network effect—relying on a summary of published network autocorrelations from many different sources—is always consistent, displays superior performance, and is the Bayes factor we recommend. At the same time, using a uniform prior for $\rho$ instead yields properties that are almost as good as those based on the empirical prior. Finally, we do not recommend employing improper priors for $\rho$ in combination with the fractional Bayes factor methodology. We illustrated the practical use of the recommended Bayes factors with three examples and provided computer code in R, making the proposed Bayes factors easily available to researchers interested in testing for the existence and magnitude of a network effect in the network autocorrelation model.

Given the importance of the network autocorrelation model in a wide variety of fields, we believe there is much value to having available an approach that makes it possible for researchers to test hypotheses that go beyond the standard significance test $H_0 : \rho = 0$ versus $H_1 : \rho \neq 0$ which is only useful for falsifying the null. The new Bayesian tests provide means for quantifying evidence in favor of any hypothesis and enable researchers to test multiple hypotheses against one another in a single analysis including, but not restricted to, any combination of precise and interval hypotheses. Overall, we hope that these tools will enrich the tool kit of researchers studying network effects through the network autocorrelation model.

Acknowledgment

We thank Michael A. McPherson and Michael L. Nieswiadomy for sharing their data with us.
Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The third author was supported by a Veni Grant (016.145.205) provided by The Netherlands Organization for Scientific Research.

Supplementary Material

Supplementary material for this article is available online.

Notes

1. By convention, we exclude loops, that is, relationships from an actor to himself, so \( W_{ii} = 0 \) for all \( i \in \{1, ..., g\} \).
2. Improper priors for nuisance parameters, in our case \( \sigma^2 \) and \( \beta \), appearing in both the null and the alternative model(s) are routinely used in Bayesian hypothesis testing (Hepple 1995; Jeffreys 1961).
3. In R, these probabilities are calculated as
   \[
   \frac{\text{pnorm}(0.5, \text{mean} = 0.36, \text{sd} = 0.19) - \text{pnorm}(0.25, \text{mean} = 0.36, \text{sd} = 0.19)}{\text{pnorm}(1, \text{mean} = 0.36, \text{sd} = 0.19) - \text{pnorm}(0.5, \text{mean} = 0.36, \text{sd} = 0.19)} = 0.23.
   \]
   We rounded all probabilities to two decimal places.
4. Only Mulder (2014a) discussed a similar method for assigning prior probabilities to hypotheses with order constraints on the tested parameters.
5. We use the terms hypothesis and model interchangeably throughout the article.
6. Note that the prior odds of any two interval hypotheses do not depend on the choice of the prior probability of the precise hypothesis.
7. The prior in equation (10) has the same asymptotic behavior as the model’s conditional independence Jeffreys prior for \( p \) for \( p \to 1 \), see Online Appendix C for a proof.
8. According to Watts and Strogatz (1998), small-world networks are usually large and sparse with \( g \gg d \gg \ln(g) \). Obviously, this relationship does not hold for all of our generated networks as, for example, \( \ln(100) = 4.61 \). However, we aim at constructing networks with a partially realistic, nonrandom configuration and simulate networks that at least resemble small-world structures.
9. In Table 2, we do not present the results for the fractional Bayes factor based on \( b_2 \) as \( b_1 = b_2 \) for both spatial networks when \( k = 4 \).
10. Simulation results for \( d = 8 \) are available from the authors on request. We do not present them here, as they do not provide any additional, that is, different, insights.
11. All reported \( p \) values are for the one-sided test \( H_0 : \rho = 0 \) versus \( H_1 : 0 < \rho < 1 \) and based on the Wald test statistic using the expected Fisher information matrix for computing standard errors.

12. All raw connectivity matrices in the examples have been subsequently row-normalized by the authors.

13. Note that a spatial Durbin model can be represented as a network autocorrelation (equation 1) by rewriting \( X\beta + WX\gamma \) as \( \bar{X}\hat{\beta} \), where \( \bar{X} = (1, X, WX) \) and \( \hat{\beta}^T = (\alpha, \beta^T, \gamma^T) \).

14. We created the U.S. county map by using the function “map” from the maps package in R (code by Becker et al. 2016). The depicted county map does not include some very small counties in Virginia which is why there are data for only 3,076, instead of the full 3,107, counties displayed.


References


**Author Biographies**

**Dino Dittrich** is a PhD candidate at the Department of Methodology and Statistics at Tilburg University, the Netherlands. His project is about Bayesian estimation and hypothesis testing techniques for social network models.

**Roger Th. A. J. Leenders** is a professor at the Department of Organization Studies at Tilburg University, the Netherlands. He holds a PhD in Sociology from the University of Groningen. He has published broadly on creativity, product development, and knowledge transfer in organizations, in leading journals such as *Organization Science*, *Journal of Applied Psychology*, *Journal of Product Innovation Management*, *Social Networks*, and the *Academy of Management Journal*.

**Joris Mulder** is an assistant professor at the Department of Methodology and Statistics at Tilburg University, the Netherlands. He holds a PhD in Statistics from Utrech University. His research focuses on Bayesian statistical methods for social science research.