Network Autocorrelation Modelling: Bayes Factor Testing for Beta Coefficients

Afra Kilic

Supervisor: dr. ir. Joris Mulder

Abstract

In network autocorrelation model, the classical hypothesis testing procedures for independent variables' effect on the outcome variable can only be used to falsify a precise null hypothesis of no effect. Classical methods are incompetent for both quantifying evidence for the null and testing multiple hypotheses simultaneously. In order to deal with these limitations, this study presents Bayes factor testing for β coefficients in the network autocorrelation model using Savage-Dickey ratio and BIC approximation methods. We propose Bayes factors for two-sided and multiple hypotheses testing procedures. Simulation results suggest that Bayes factor for the latter shows higher performance and it is the one we recommend. Then, we illustrate the practical use of the proposed Bayes factors with two real data examples and compare the results to those coming from classical tests using p values. Finally, R code used in this study and for computing the proposed Bayes factors is provided.

I. Introduction

The prominent assumption of linear regression models is the independence of data points. However, many situations exist in which data points are not truly independent but instead there is a network structure between observations; in such cases, individual observations (e.g., countries, organizations and persons) depend on the other observations in the data. The network autocorrelation model (Doreian 1981, Ord 1975) is used commonly to quantify and test the strength of network effect on individual behavior with respect to other covariates in a given network (Fujimoto, Chou, and Valente 2011). In this model, individual responses or observations are assumed to be correlated by the amount of network autocorrelation parameter ρ, representing the strength of dependency in the network. The identification and the magnitude of the exogenous variables' effect on the outcome variable controlling for the network effect is often the focus of interest in model applications. For instance, housing prices are similar across adjacent neighborhoods, thus, to investigate the factors affecting housing prices this adjacency has to be taken into consideration. Although this model considers the adjacency between observations, as in the classical linear regression, the identification of exogenous variable's effect on the outcome variable results in an inferential test of $H_0 : \beta = 0$ versus $H_1 : \beta \neq 0$. Eventually, if the null hypothesis is rejected, the researcher then conclude that there is an effect of the exogenous factor on the outcome variable controlling for ρ. This standard approach to test the estimations (p and β) is null hypothesis significance testing. Traditional tests for the null hypothesis' significance, such the Wald test, likelihood ratio test, or Lagrange multiplier test, are based on various test statistics, or summary values derived from the sample which are assumed to have asymptotically known distributions (Leenders, 1995). Despite the fact that this statistical approach have yielded many interesting and theoretically useful insights, more intricate hypothesis tests are more informative. In this sense, Bayesian hypothesis testing procedure can be more insightful for Beta coefficients in network autocorrelation model as it allows researchers to test more than two hypotheses against each other and quantify the relative evidence for each hypothesis. Furthermore, a Bayesian approach to hypothesis testing deals with the problem in the classical approach like asymptotic assumption and inconsistency. Hence, the present paper focuses on Bayesian hypothesis testing procedure for β coefficients in the network autocorrelation model.

In significance test, the decision to reject the null hypothesis depends on p-value which is the probability of obtaining a test result as extreme as the observed test statistics under the assumption of true null hypothesis. After the calculation of this probability, the obtained p-value is compared with the predetermined α value which is set to 0.05 in most cases (Weakliem, 2004). If the p value smaller than α, the null hypothesis is rejected, which is also known as a significant test result. In other words, If the p value is smaller than α level, it is concluded that there is enough evidence in the data to reject the null hypothesis of no effect. Conversely, if the p value is larger than α, data does not contain enough evidence to reject the null.

Even though the classical null hypothesis significance testing is commonly applied in the field due to its convenience, there are several drawbacks of using this method in the network autocorrelation model. First, maximum likelihood-based test statistics such as p values and confidence intervals rely heavily on asymptotic theory. Dittrich et al. (2017b) showed that the probability of making $TypeI$ error is not controlled for small networks which are encountered in social science research such as school classes, care teams, members of a club. Second, in the case of testing multiple competing hypothesis against each other, p values are incompetent (Shaffer, 1995) as they only allow researchers to test each competing hypothesis against the null hypothesis. Thus, this process does not provide the information regarding which hypothesis, out of a set of precise and interval hypothesis, is most supported by the data. In order to make this statement clearer, consider a situation where $\hat{\beta} = 0.32$ (standardized version), with a 95 percent confidence interval for β of $(-.04; .68)$, and a one-sided p value of .08. The null hypothesis $H_0 : \beta = 0$ would not be rejected at the standard significance level of $\alpha = .05$ and it would be conclude that there is no effect of the exogenous variable on the outcome variable. However, if we take a look at the confidence interval there may be a positive effect of the variable on the outcome variable. In that sense, standard approach is very restrictive to asses both existence and the magnitude of the effect.

Furthermore, the method cannot be used to generate evidence to support the null hypothesis; it can only be used to refute it (Wetzels and Wagenmakers, 2012). In other words, the p value larger than α is only a state of ignorance where the null neither be rejected nor supported by the data. Likewise, p smaller than α only shows that there is enough evidence in the data to reject the null, but it does not provide the information regarding how much evidence we have to support the alternative hypothesis. Thus, we actually cannot know whether the alternative is true or not even the null is rejected in significance test. Finally, the classical null hypothesis significance tests are inconsistent. There will always be probability of α (typically .05) of making $TypeI$ error, even if the null is true and the sample size grows to infinity. $TypeI$ error refers to mistaken rejection of a true null hypothesis. Probability of making $TypeI$ error is unrelated to the sample size and the only way to reduce the probability is to set lower α which in turn makes the test very conservative and hard to make inferences with results. This is not desired because a larger sample size should allow for more precise results as it contains more information.

Bayes factor, a Bayesian approach to hypothesis testing, allows us to deal with the problems occurring in the null hypothesis significance testing. First, in this approach Hypothesis can be formulated in a precise way like $H_0 : \beta_1 = 0$, or as interval hypothesis $H_1 : -5 < \beta_1 < 0$. In this example, $B_{01} = 5$ shows that the data is five times more likely to be observed under H_0 compared to H_1. Second, Bayes factors can be easily extended to test more than two hypotheses against each other simultaneously (Raftery, Madigan and Hoeting, 1997). For example, $H_0 : \beta_1 = 0$ versus $H_1 : -3 < \beta_1 < 0$ versus $H_2 : 2.5 > \beta_1 > 0$ versus $H_4 : 3 < \beta_1 < 10$. Hence, they allow researchers to assess and quantify the relative evidence in the data supporting the null or any other hypothesis in comparison to any competing hypothesis. In that sense, such a test is much more insightful than the classical null hypothesis significance testing no effect, $H_0 : \beta = 0$ versus "some" (positive or negative) exogenous effect, $H_1 : \beta \neq 0$. Third, unlike the significance tests, Bayes factor is consistent such that in the case that H_0 is the true model B_{01} converges to infinity as the sample size goes to infinity (Casella et al., 2009). Finally, Bayes factor provides “exact” inference without asymptotic approximation (Berger, 2006; De Oliveira and Song, 2008), which is especially very advantageous when dealing with small network sizes.

In the existent literature, Bayes factor has been used for testing competing connectivity matrices against each other (Hepple, 1995) and for testing different explanatory variables with the Bayesian model averaging procedure (LeSage and

Parent, 2007). Dittrich et al. (2017a), introduced Bayes factor tests (Jeffreys, 1961; Kass and Raftery 1995; Mulder and Wagenmakers, 2016) based on an empirical and a uniform prior for the network effect. Their methodology makes it possible to test multiple competing hypothesis regarding network effect against each other and it works for any combination of precise and/or interval hypothesis. However, to our knowledge, there is no study on Bayes factor for testing β coefficients in network autocorrelation model. As the main interest of the network autocorrelation application is to control the network effect for the identification and magnitude of the exogenous variables' effect on the outcome variable, the present paper proposes Bayes factor testing for beta coefficients as an alternative hypothesis testing to overcome the problems with the classical hypothesis significance testing. To do so, we applied a Bayesian approach to two-sided and multiple hypothesis testing procedures. Two-sided hypothesis testing considers testing no effect no effect versus "some" effect as in the classical testing approach using Bayes factor rather than p-value. In addition to two-sided hypothesis testing, multiple hypothesis makes possible to test multiple competing hypotheses regarding β coefficients and it works for any combination of precise and/or interval hypothesis.

In order to compute Bayes factor, priors have to be specified for the unknown model parameters for each hypothesis prior to the data observation. For testing problem considered in this work, prior specification for the network autocorrelation β can be very time consuming. Besides, given the complexity of the network autocorrelation model, marginal likelihood estimation of the model is also computationally hard. In this article, we developed and explored Bayes factors for two-sided and multiple hypothesis testing by considering these two barriers. In the two-sided hypothesis testing, Bayes factors were calculated using the Savage-Dickey ratio with a unit information prior and the Bayesian Information Criteria (BIC) approximation which implicitly assumes a unit information prior. The Savage-Dickey ratio with a uniform prior was also used to compute Bayes factor in multiple hypothesis testing. Thus, our methodology also aims to propose a general and easily applicable Bayes factors for testing β coefficients: First the proposed Bayes factors are computationally efficient with Savage-Dickey ratio and BIC approximation; second they are "objective" and time efficient with unit information prior. Subsequently, we conducted a simulation study to investigate the numerical properties of and differences between the proposed

Bayes factors. Finally we used the Bayes Factors to reanalyze two data sets from the literature.

The article is organized as follows: In the second section, we discuss the network auto correlation model in more detail and continue with a short introduction to Bayesian hypothesis testing in the third section. In the fourth section, we motivate our prior specification for network autocorrelation β coefficients and Bayes factor calculation methods. The fifth section continues with a simulation to assess the numerical performance of the Bayes factors and then apply them to two real data sets to highlight the practical use of the proposed Bayes factors. The seventh section concludes.

II. Network Autocorrelation Model

The network autocorrelation model has been proposed by geographers (Ord, 1975) to address the issue of structural dependence among the observations. Unlike the conventional linear regression model, the network autocorrelation model does not assume that the observations are independent from each other, but it allows for dependence among them and takes this interdependence into account. The vast majority of social phenomena are interconnected within networks of interdependencies meaning that individual outcomes may not only be influenced by exogenous factors but may also be influenced by the ideas of other network actors. In particular, the network effect, also known as network autocorrelation parameter rho, operates directly on the outcome variable along with the exogenous variables. Formally, the network autocorrelation model is expressed as

$$y = \rho W y + X\beta + \epsilon, \qquad \epsilon \sim \mathsf{N}(0, \sigma^2 I_g) \tag{1}$$

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, β is s $(k \times 1)$ -vector regression coefficients, I_g is s $(g \times g)$-identity matrix and ϵ is a $(g \times 1)$-vector containing independent and identically normally distributed error terms with zero mean and variance of σ^2. Moreover, W is a given $(g \times g)$ - connectivity matrix representing social ties in a network, with $W_{ij}(i, j \in \{1, ..., g\})$ referring the degree of influence of actor j on actor i. Finally, ρ is a scalar termed the network autocorrelation parameter quantifying the social influence for a given y, W, X. The resulting model parameters can be noted as $\theta := (\rho, \sigma^2, \beta)$, where $\rho = 0$ reduces the model to the standard linear regression model. The model's likelihood function given by (e.g., Doreian, 1980)

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

where $A_\rho := I_g -$. To ensure that $|\det(A_\rho|$ is non-zero and model's likelihood function in (2) is well-defined, there are restrictions on the region of support for ρ. Typically, this region chosen as the interval containing $\rho = 0$ for which A_ρ is nonsingular (Hepple, 1995; LeSage and Parent, 2007; Smith, 2009). This interval is given by $(\lambda_g^{-1}, \lambda_1^{-1})$, where $\lambda_1 \geq \lambda_2 \geq ...\lambda_g$ are the ordered eigenvalues of W (Hepple, 1995). For row-standardized connectivity matrices W, meaning that where each row sum equals to one, it holds that $\lambda_1 = 1$ (Anselin, 1982). We will use row-standardized matrices in the remainder and the associated parameter space will becomes $\Omega := \Omega_\rho \times \Omega_{\sigma^2} \times \Omega_\beta = (\lambda_g^{-1}, \lambda_1^{-1}) \times (0, \infty) \times \mathbb{R}^k$.

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 like 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, Lo´pez, and Angulo 2008).

III. Bayesian Hypothesis Testing

In network studies, researchers are interested in both the network effect and exogenous factors effect on the outcome variable. To test the latter one, which is the main interest of the present work, the network effect can be considered an effect to be controlled. After controlling the network effect on outcome variable, Our expectations about whether exogenous factors have an effect on the outcome variable and, if any, the magnitude of this effect can be expressed such as $H_0 : \beta = 0$, or $H_1 : \beta \neq 0$, or $H_2 : -1 < \beta < 0$, or $H_3 : 1 > \beta > 0$. Note that these intervals are based on standardized β coefficients, and can be adjusted based on the expectations of the direction and magnitude of the effect. After setting our expectations on the exogenous factors' effect, the main question we are interested is which of these hypotheses is most likely true; Bayes factor testing aids in answering this question. The Bayes factor of hypothesis H_i against H_j, $i, j \in \{0, 1, ..., T-1\}$, is defined as the ratio of the marginal likelihoods under the two hypotheses.

$$B_{ij} = \frac{p(y|H_i)}{p(y|H_j)} = \frac{\int_\Omega f(y|\theta_i, H_i)p(\theta_i|Hi)d\theta_i}{\int_\Omega f(y|\theta_j, H_j)p(\theta_j|Hj)d\theta_j} \tag{3}$$

The term $p(y|H_i)$ in equation (3) refers to the marginal likelihood of data y under the hypothesis H_i and denotes the probability that data were observed under H_i. It is computed by integrating the product of the model's likelihood function and prior distribution for the model parameters under H_i. Note that θ_i are the model parameters under H_i; $p(\theta_i|Hi)$ refers to their prior density; and Ω corresponds parameter space as $\Omega := \Omega_\rho \times \Omega_{\sigma^2} \times \Omega_\beta = (\lambda_g^{-1}, \lambda_1^{-1}) \times (0, \infty) \times \mathbb{R}^k$. Thus, 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 Thus, the Bayes factor of hypothesis H_i against H_j can be expressed for β parameter in the network autocorrelation as

$$B_{ij} = \frac{\int_{\mathbb{R}_i^k} \int_0^\infty \int_{\lambda_g^{-1}}^{\lambda_1^{-1}} p(\rho)p(\sigma^2)p_i(\beta_i)f(y|\rho, \sigma^2, \beta_i)d_\rho d_{\sigma^2} d_{\beta_i}}{\int_{\mathbb{R}_j^k} \int_0^\infty \int_{\lambda_g^{-1}}^{\lambda_1^{-1}} p(\rho)p(\sigma^2)p_j(\beta_j)f(y|\rho, \sigma^2, \beta_j)d_\rho d_{\sigma^2} d_{\beta_j}} \tag{4}$$

where β_i is the exogenous factor parameter under the hypotheses H_i; $p_i(\beta_i)$ denotes their prior density. We assume common priors for σ^2 and ρ under both hypothesis H_i and H_j are seen as nuisance parameters and their exact form of the priors typically does not affect the magnitude of the Bayes factor (Kass and Raftery, 1995). As we can see from the equation (4), B_{ij}, as the ratio of two marginal likelihoods, quantifies the relative evidence that the data were observed under the hypothesis H_i rather than H_j. For instance, $B_{ij} = 5$ means that it is five times more likely to have observed the data under hypothesis H_i than under hypothesis H_j.

Jeffreys (1961) suggested a classification scheme to divide up the types of Bayes factors (see Table 1). For instance, when the Bayes factor is greater than 3, there is "substantial" evidence for H_i; conversely, when the Bayes factor is less than 1/3, there is "substantial" support for H_j. When referring to relative evidence supporting a hypothesis, these labels offer some illustrative suggestions, but the conclusion should eventually be determined by the context of the research question (Kass and Raftery 1995).

B_{ij}	**log(B_{ij})**	**Interpretation**
>100	>4.61	Decisive evidence for hypothesis H_i
30 to 100	3.40 to 4.61	Very strong evidence for hypothesis H_i
10 to 30	2.30 to 3.40	Strong evidence for hypothesis H_i
3 to 10	1.10 to 2.30	Substantial evidence for hypothesis H_i
1 to 3	0 to 1.10	Not worth more than a bare mention
1/3 to 1	–1.10 to 0	Not worth more than a bare mention
1/10 to 1/3	–2.30 to –1.10	Substantial evidence for hypothesis H_j
1/30 to 1/10	–3.40 to –2.30	Strong evidence for hypothesis H_j
1/100 to 1/30	–4.61 to –3.40	Very strong evidence for hypothesis H_j
<1/100	<–4.61	Decisive evidence for hypothesis H_j

Table 1: Evidence Categories for the Bayes Factor B_{ij} as Given by Jeffreys(1961)

IV. Bayes Factor Computation

Notwithstanding the advantages of Bayesian hypothesis testing using Bayes factor, there are two main challenges limiting popularity of Bayes factor: prior specification can be a very difficult task, especially when prior information is weak or unavailable completely; and the computation can be quite difficult when the statistical model is complex. In order to deal with these barriers given the complexity of network autocorrelation model, in the present paper we calculated Bayes factors via BIC approximation and Savage-Dickey density ratio with unit information prior.

Wagenmakers et al. (2010) defines the first one as a conceptual challenge that researchers have only a vague idea of the vagueness of their prior knowledge, or that they seek to use less informative priors to be more "objective". However, Bayesian hypothesis testing is very sensitive to prior specification (e.g., Bartlett, 1957; Liu Aitkin, 2008) because the marginal likelihood is an average taken with respect to the prior. Hence, in the absence prior definite and specific information about the variable of interest, the major challenge is to determine the amount of information contained in the prior distribution. For instance, assume a researcher without a definite information seek for a prior for the mean μ of a Normal distribution with known variance. In order to refrain from expressing a preference, she may tempt to use a vague prior of a Normal distribution with mean zero and variance of 10.000. However, regardless of the data, more spread priors are conservative in the sense that they to show a preference for the simple model (e.g., one in which $\mu = 0$). A common way to deal with prior specification barrier and to increase the robustness of Bayesian hypothesis testing is to use default priors that are not super vague but also not reflecting subjective prior beliefs.

The second challenge for Bayesian hypothesis testing is that the computation of the marginal likelihood and the Bayes factor are often quite difficult especially when the model is complex. As a solution to this problem, a series of different methods has been developed (Gamerman Lopes, 2006, chap. 7.). However, almost all of these computational methods become less efficient and more difficult to implement as the underlying models become more complex. For instance, it is possible to compute Bayes factor directly without first calculating marginal likelihoods via Markov Chain Monte Carlo (MCMC) sampling method. Even if with increasing computational power it is now easier to apply MCMC sampling method, it is still very time inefficient for complex models (e.g., network autocorrelation models with high density and network size).

I. The BIC Approximation

BIC approximation has computational advantage and it does not require to specify prior distribution because it implicitly assumes unit information prior (Wagenmakers, 2007). Unit information prior was proposed by Kass and Wasserman (1995) based on Fisher information and it is a multivariate normal prior with a mean at the maximum likelihood estimate and variance equal to the expected information matrix for one information (Kass and Wasserman, 1995). The unit information prior is well spread out compared to the likelihood and is relatively flat within the parameter space where the likelihood is relevant , without being much larger outside of that region. Hence, the likelihood dominates the prior and it satisfies the conditions for a *stable estimation situation* where inference about μ is relatively insensitive to the prior (Edwards, Lindman and Savage, 1963). This property ensures that the BIC is "objective" in the sense that different researchers will have the same statistical inference from the same data and set of models. Typically, BIC for H_i is defined as

$$BIC(H_i) = -2logL_i + k_i log(n) \tag{5}$$

where n is the number of observations, k_i is the number of free parameters of model H_i, L_i is the maximum likelihood for model H_i, that is $L_i = P(y|\beta, \sigma^2, \rho, H_i)$. The BIC can be used approximate the prior predictive probability $Pr(y|H_i)$ as $Pr_{BIC}(D|H_i) = exp[-BIC(H_i/2)]$. In case of two models or hypotheses, the Bayes factor is defined as the ratio of the prior predictive probabilities as

$$B_{01} \approx \frac{Pr_{BIC}(D|H_0)}{Pr_{BIC}(D|H_1)} = exp(\triangle BIC_{10}/2) \tag{6}$$

where $\triangle BIC_{10} = BIC(H_1) - BIC(H_0)$. For the derivation of BIC approximation of Bayes factor and its details we refer to Raftery(1995, 1999) and Wagenmakers (2007). We applied BIC approximation only to the two-sided test for $H_0 : \beta = 0$ and $H_1 : \beta \neq 0$. Hence, the model H_1 includes both the intercept term and the covariate of interest β whereas H_0 only includes the intercept term.

II. Savage-Dickey Approach

Bayes factor can also be calculated by taking ratio of the prior and posterior density of the parameter of interest at the point of interest, which is also known as Savage-Dickey method (Wagenmakers, 2010). Mathematically, the Savage-Dickey density ratio for two sided hypothesis testing ($H_0 : \beta = 0$ and $H_1 : \beta \neq 0$) says:

$$B_{ij} = \frac{p(y|H_0)}{p(y|H_1)} = \frac{p(\beta = 0|y, H_1)}{p(\beta = 0|H_1)} \tag{7}$$

Unlike BIC approximation, for Savage-Dickey we need to specify a prior. As explained above, Bayesian hypothesis testing is very sensitive to prior selection, and unit information prior enables the likelihood to dominate the prior, thus, it satisfies the conditions for a *stable estimation situation* (Edwards, Lindman, and Savage, 1963). Hence, to prevent Bayes factor from becoming conservative with a very spread prior, we used unit information prior in Savage-Dickey method for both two-sided test and multiple hypothesis testing. However, to show less preference in multiple hypothesis testing, we used a multivariate normal prior with a mean of zero and variance equal to the expected information matrix for one information, that is, $\mathsf{N}(0, \hat{\sigma_\beta^2}/\sqrt{g})$.

For the multiple hypothesis testing ($H_0 : \beta_1 = 0$ versus $H_1 : \beta_1 < 0$ versus $H_2 : \beta_1 > 0$), we used transitivity feature of Bayes factor (i.e. $B_{12} = B_{10} \times B_{02}$) (Morey and Wagenmakers, 2014). As Savage-Dickey only applies to nested models, we assumed an unconstrained hypothesis $H_u \in \mathbb{R}$ and then computed B_{0u}, B_{1u}, B_{2u} to calculate how much evidence each hypothesis contains relative to the general hypothesis . Hypotheses H_1 and H_2 refer to an interval rather than a point. Hence, when calculating B_{1u}, B_{2u}, we considered the ratio of the area below the posterior distribution to the area below the prior distribution over the *interval* of interest, rather than the ratio of height of posterior to prior at a point. For instance, mathematically this can be shown for $H_1 : \beta_1 < 0$ versus $H_u \in \mathbb{R}$ as

$$B_{1u} = \frac{p(\beta < 0|y, H_u)}{p(\beta < 0|H_u)} \tag{8}$$

Thus, after having Bayes factor for each hypothesis against the complementary hypothesis H_u, Bayes factor can be calculated for the hypotheses of interest with transitive property. For instance, Bayes factor of hypothesis H_1 against H_2 can be computed as $B_{12} = \frac{B_{1u}}{B_{2U}}$.

V. Simulation Study

We performed a two-stage simulation study to examine the performance of Bayes factor for testing β coefficients. First, we conducted a two-sided test for β. Specifically, we tested the following hypotheses against each other: $H_0 : \beta = 0$ and $H_1 : \beta \neq 0$.

In order to do that we have applied two different methodology for calculating the Bayes factor; Savage-dickey density ratio and Bayesian Information Criteria (BIC). In particular, we explored which Bayes factor converges fastest to the true hypothesis and assesed the sensitivity of Bayes factors calculated with two different approach. Second, we conducted a multiple hypothesis for β coefficients with a precise hypothesis and two interval hypothesis: $H_0 : \beta = 0$ versus $H_1 : \beta > 0$ versus $H_2 : \beta < 0$. For the two-sided hypothesis testing, we only used Savage-dickey Bayes factor approach. We particularly considered the following network autocorrelation model with one-covariate and an intercept value;

$$y = \beta_0 + X\beta_1 + \rho W y + \epsilon, \qquad \epsilon \sim \mathsf{N}(0, \sigma^2 I_g) \tag{9}$$

I. Study Design

We designed the simulation based on the previous simulation studies of the network autocorrelation model (Mizruchi and Neuman, 2008; Neuman and Mizruchi, 2010: Wang et al., 2014). By considering the model as $y = (I_g - pW)^{-1}(X\beta + \epsilon)$, the outcome variable y was generated via random networks by altering the network density (d), the network size (g), the magnitude of network effect (ρ) and the magnitude of the effect of covariates (β_1). We considered six levels of network densities $\{d \in .1, .2, .3, .4, .5, .6\}$, three different network sizes $\{g \in 20, 50, 100\}$, three fixed levels of network sizes $\{\rho \in 0, .2, .5\}$. For β_1 values, we generated a sequence of 20 between -2 and 2. We first generated binary connectivity matrices via "rgraph" function from the sna package in R (Butts, 2008), that is, if there is a tie between actor i and j, $W_{ij} = 1(i, j \in \{1, ..., g\}$ and zero otherwise. Binary connectivity matrices also imply zero diagonal entries.We then row-normalized the generated symmetric raw connectivity matrices. Second, we drew independently values from a standard normal distribution for the X elements (excluding the first column of X, which is a vector of ones) and ϵ, which implies $\sigma^2 = 1$. In total we considered 240 scenarios. Further information on data generation is given in the study details below. R code used for the simulation study can be found under the Appendix as HTML file.

II. Bayes Factor for Two-Sided Test

In this part of the simulation, we conducted a two-sided test to examine Bayes factor performance for testing β_1 coefficients in the network autocorrelation model (9) based on two different approaches: Savage-Dickey density ratio and Bayesian Information Criteria (BIC). The data used in this test generated based on the parameters: $\rho = 0.2$, $d = 0.2$ $\sigma^2 = 1$, $\beta 1 \in (-2, ..., 2)$ and $g \in \{20, 50, 100\}$. Hence, we considered 60 different scenarios and ran 1000 replications for each data set to see how two B_{01} behave across different g and β_1 values. After generating the data, first, we calculated B_{01} to quantify the relative evidence that the data were observed under the hypothesis $H_0 : \beta = 0$ rather than $H_1 : \beta \neq 0$. Furthermore, we have examined how posterior probabilities $p(H_0|y)$ and $log(B_{01})$ behaves to see which B_{01} is more sensitive or responsive to changes in the real value of β_1.

First, posterior probabilities quantify how probable each hypothesis after observing the data, thus it is the quantity that researchers are typically interested in. The higher posterior probability refers to a stronger evidence in the data in favor of that hypothesis. As we interested in B_{01} to test two hypotheses (H_0 and H_1) against each other and each has the same prior, we calculated the posterior probability as $p(H_0|y) = B_{01}/(B_{01} + 1)$.

Second, as β_1 values move away from zero, B_{01} approaches to zero however it is hard to observe how much it is close to zero which in turns makes hard to interpret how strongly B_{01} responds to a change in β_1. Hence, we also evaluated the $log(B_{01})$ to compare the sensitivity of Savage-Dickey and BIC B_{01}s.

As g increases, the data will contain more information which allows us to estimate β_1 more precisely. Figure 1 shows that in both Savage-Dickey and BIC as sample size increases from 20 to 100, the amount of information to estimate β_1 increases such that in all graphs standard error decreases. Hence, precision level of β_1 estimates increases and B_{01} becomes more responsive to changes in β_1. B_{01} and posterior probability graphs show that BIC approximation and Savage-Dickey ratio do not differ from each other. However, as g increases B_{01} based on the Savage-Dickey ratio becomes more sensitive to changes in β_1 such that $log(B_{01})$ for the Savage-Dickey ratio goes

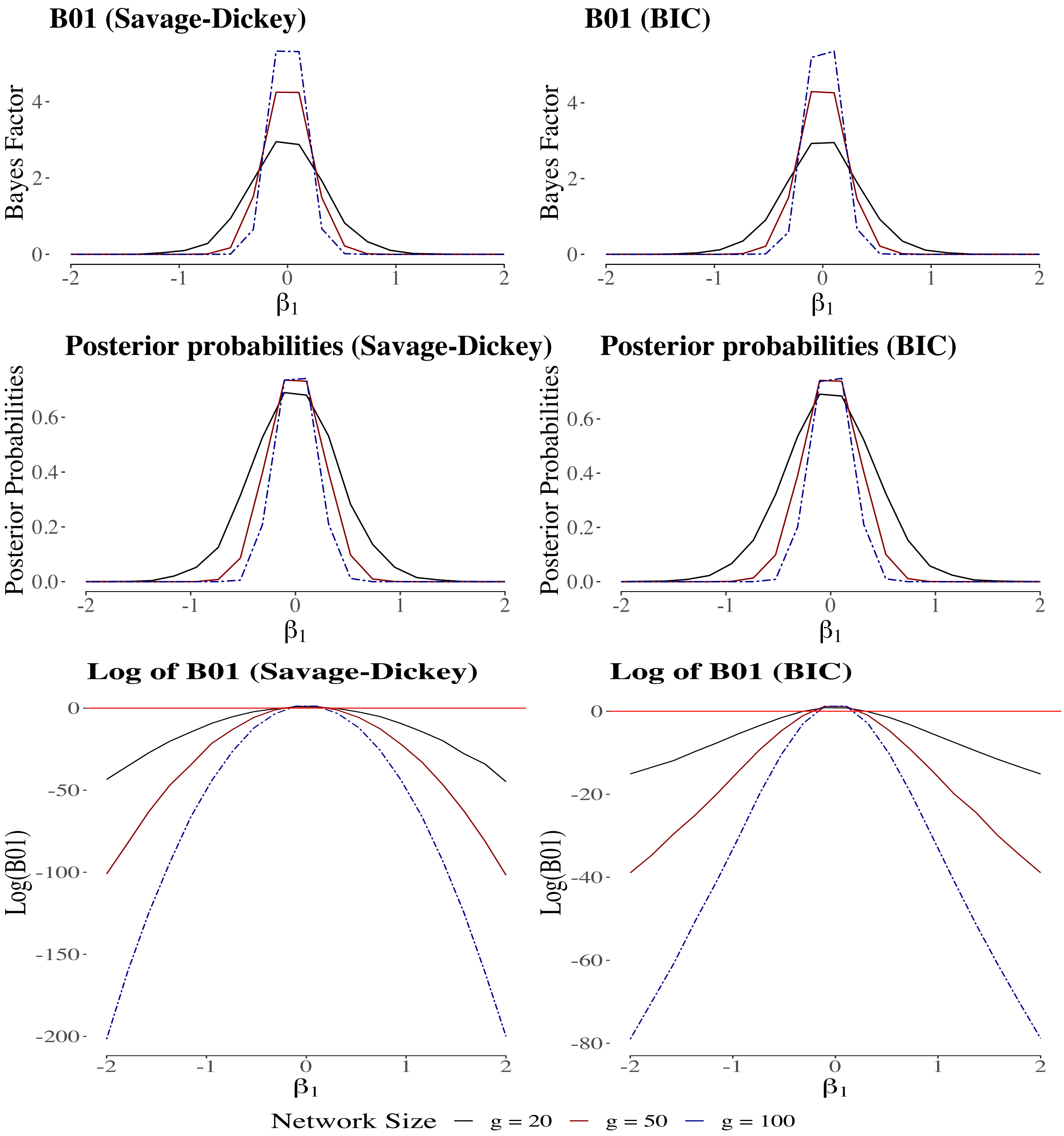

Figure 1: B_{01}, posterior probability ratios and $log(B_{01})$ for simulated data.The right side displays the results for Savage-Dickey estimation and the left side for BIC estimation.

until -200, whereas $log(B_{01})$ for the BIC goes until -80 (Figure 1. Two graphs at the bottom). Another notable difference observed in $log(B_{01})$ graphs, the part above the zero line is slightly larger for the BIC approximation, meaning that BIC is better at identifying the cases where β equals to zero regardless of the network size.

Two-sided test only informs us about the existence of an effect of the covariate of interest. However, it is very unlikely to observe β values exactly equal to zero in the real world. In $log(B_{01})$ graphs, even though BIC approximation is doing better, observing a value above the zero line is still extremely unlikely. Furthermore, as $\beta 1$ goes either to 2 or -2, B_{01} approaches to zero and $log(B_{01})$ approaches to $-\infty$, meaning that this test also does not inform us about the direction of the effect. In this sense, two-sided test is not very informative about β values.

Lastly, we examined how B_{01} behaves across different ρ and density values when $g = 50$ for both Savage-Dickey and BIC methods. The data used in this test generated based on the paramaters: $\rho \in \{0, .2, .5\}$, $d \in \{.1, .2, .3, .4, .5, .6\}$ $\sigma^2 = 1$, $\beta_1 \in (-2, ..., 2)$ and $g = 50$. In particular, other parameters are the same, in ρ testing we hold d constant at .2 and in d testing ρ at .2. Hence, we considered 180 different scenarios (60 for ρ testing 120 for d testing) and ran 1000 replications for each data set. Figure 2 shows that B_{01} behaves similar across different ρ and d values either in Savage-Dickey or in BIC. $log(B_{01})$ graphs, again, clearly show that B_{01} based on Savage-Dickey method is much more sensitive to changes in β_1.

III. Bayes Factor for Multiple Hypothesis Testing

In order to deal with the limitations of the two-sided test, in this part we conducted a multiple hypothesis test for $\beta 1$ which is more informative about both magnitude and the strength of the effect. We considered the following three hypotheses: $H_0 : \beta_1 = 0$ versus $H_1 : \beta_1 < 0$ versus $H_2 : \beta_1 > 0$. For two-sided hypothesis testing, we only used Savage-Dickey method to calculate Bayes factor. With this method, it is also possible to increase the number of hypothesis and adjust the intervals to test more specific hypothesis such as $H_0 : \beta_1 = 0$ versus $H_1 : -3 < \beta_1 < 0$ versus $H_2 : 2.5 > \beta_1 > 0$ versus $H_4 : 3 < \beta_1 < 10$.

As in the two-sided test simulation, we generated the data based on the paramaters:

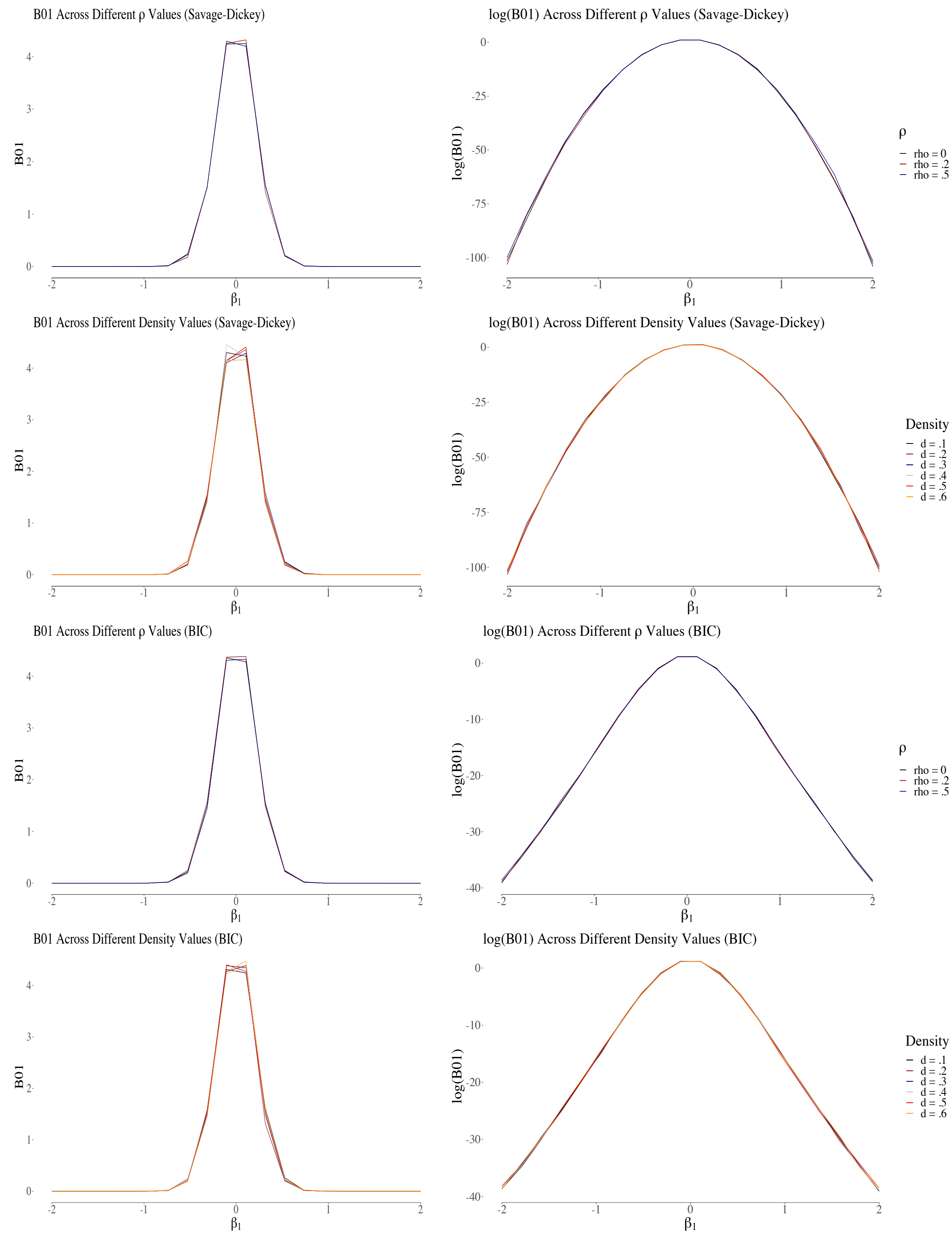

Figure 2: B_{01} and $log(B_{01})$ across different ρ and density (d) values when $g = 50$. The top four graphs display the results for Savage-Dickey method and the bottom fours graphs display the results for BIC method.

$\rho = 0.2$, $d = 0.2$, $\sigma^2 = 1$, $\beta_1 \in (-2, ..., 2)$ and $g \in \{20, 50, 100\}$. Hence, we considered 60 different scenarios and ran 1000 replications for each data set to see how $log(B_{12})$, $log(B_{01})$,and $log(B_{02})$ behave across different g and β_1 values.

Figure 3 shows how $log(B_{12})$, $log(B_{01})$ and $log(B_{02})$ behave as sample size increases from 20 to 100, and where β_1 values vary in the interval (-2,2). As g increases, the strength of evidence for the true hypothesis increases. For instance, B_{12}, showing the relative evidence that the data were observed under $H1 : \beta_1 < 0$ rather than $H_2 : \beta_1 > 0$, approaches to infinity as β_1 goes to -2 and it approaches to zero line as β_1 goes to 2. Notice that as sample size increases, the curves are bending towards zero line such that $log(B_{01})$ in the positive side and $log(B_{02})$ in the negative side are bending towards zero with higher g . Likewise, $log(B_{12})$ shows a more "S" trend by bending towards zero line as β_1 goes to zero.

Finally, to see how $log(B_{12})$, $log(B_{01})$ and $log(B_{02})$ behave across different ρ and d values, we first generated data for ρ testing based on the parameters: $\rho \in \{0, .2, .5\}$, $d \in \{.1, .2, .3, .4, .5, .6\}$, $\sigma^2 = 1$, $\beta_1 \in (-2, ..., 2)$ and $g = 50$. In particular, other parameters are the same, in ρ testing, we hold d constant at .2 and in d testing, ρ at .2. As a result, in total 180 scenarios (120 for d, 60 for ρ) and ran 1000 replications for each data set. Simulation results show that Bayes factors also in multiple hypothesis test behave similarly across different ρ and d values (Fig. 4).

VI. Empirical Examples

In the following, we applied the two-sided and multiple Bayes factor hypothesis testing to two data sets from the literature to analyze the performance of Bayes factor test for β coefficients. First, we tested the hypotheses $H_0 : \beta = 0$ versus $H_1 : \beta \neq 0$ for each variable with Savage-Dickey ratio method and BIC approximation. Second, we conducted multiple hypothesis testing for the same variables and test the hypotheses $H_0 : \beta_1 = 0$ versus $H_1 : \beta_1 < 0$ versus $H_2 : \beta_1 > 0$. Finally, we compared the results to those coming from classical tests using p values. R code used for the testing the empirical examples can be found under the Appendix as HTML file.

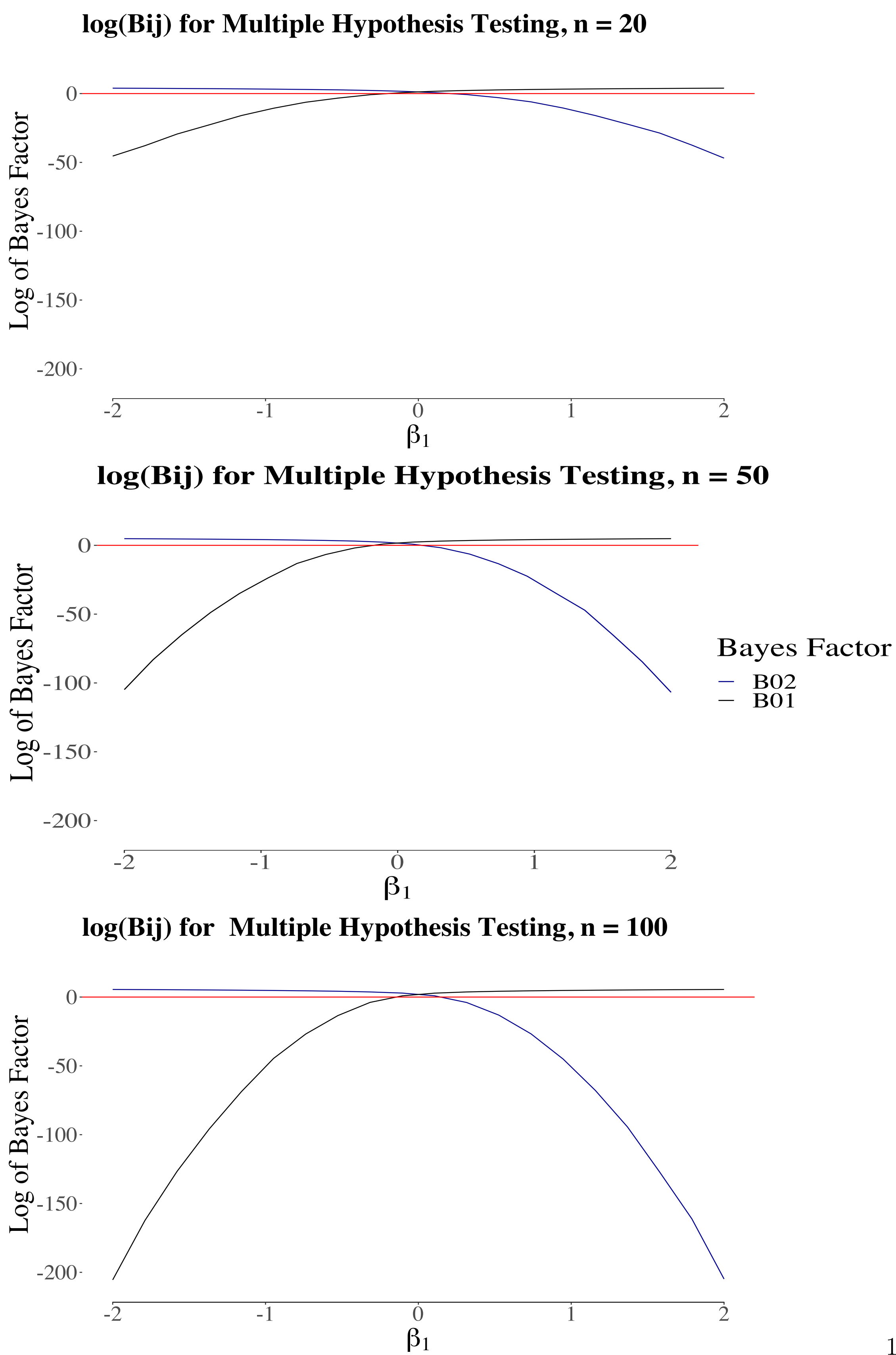

Figure 3: Logarithms of B_{01}, B_{02}, B_{12} when sample size equals to 20, 50 and 100.

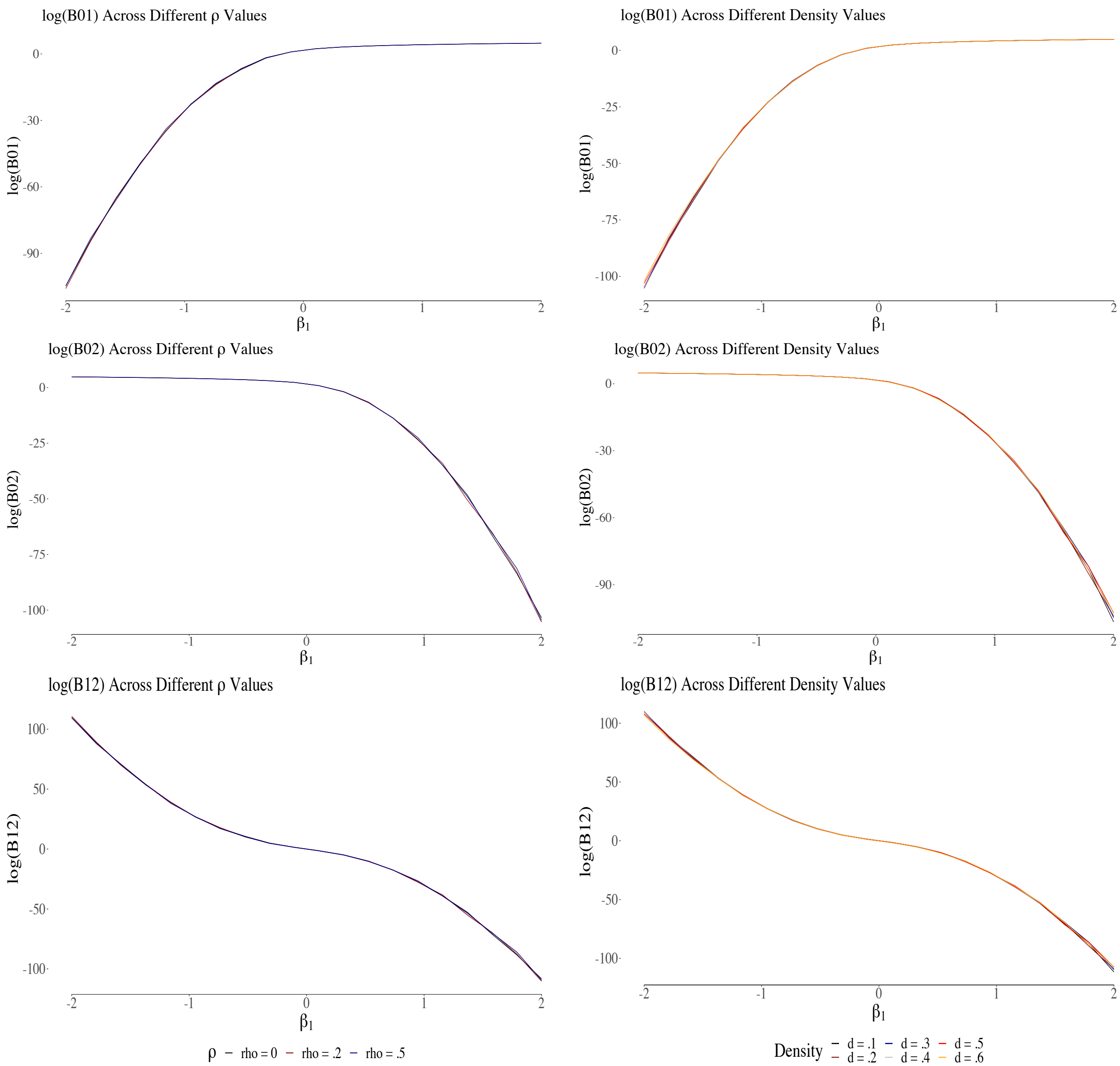

Figure 4: $log(B_{01})$, $log(B_{02})$, and $log(B_{12})$ across different ρ and density (d) values where $g = 50$.

I. Crime Data

We considered a cross-sectional data set for 49 neighbourhoods in Columbus, OH, in 1980 which is a part of Columbus data and openly accessible in the R package "spdep" (Bivand and Piras, 2015). Anselin (1988) analyzed the data to explain the neighbourhoods' crime rates, operationalized as the total residential burglaries and vehicle thefts per thousand households in the neighborhood, by using network autocorrelation model. Crime rates were modeled as a function of intercept term, household income (INC), housing value (HOVAL) and spatial structure of the data. Adjacency is defined in the R object "col.gal.nb". Results showed that both household income and housing value have a negative impact on crime rates. In particular, maximum likelihood estimates of the coefficients were $\beta_{INC} = -1.074$ with $p = 0.001$ and $\beta_{HOVAL} = -0.27$ with $p = 0.003$ respectively.

	Two-Sided Test		**Multiple Testing**			**ML**		
Variable	$log(B_{01})_{sd}$	$log(B_{01})_{bic}$	$log(B_{01})_m$	$log(B_{02})_m$	$log(B_{12})_m$	β	z	p
INC	-3.089	-3.001	-3.887	3.417	7.304	-1.0735	-3.206	0.001
HOVAL	-2.474	-2.193	-3.258	3.362	6.62	-0.27	-3.004	0.003

Table 2: Log Bayes factors for two-sided and multiple hypothesis testing for the Crime data. The last three columns show the maximum likelihood estimates of the coefficients, z statistics and the resulted p-values.

Table 2 shows log Bayes factor for each variable resulted from the two-sided test using Savage-Dickey ratio ($log(B_{01})_{sd}$) and BIC approximation ($log(B_{01})_{bic}$); and for the multiple hypothesis test ($log(B_{01})_m$, $log(B_{02})_m$ and $log(B_{12})_m$). Based on the classification scheme (Table 1) suggested by Jeffreys (1961), Bayes factor tests result in the same conclusion, as well as when using p-values (Table 2). If we compare the first two columns, $log(B_{01})_{sd}$ and $log(B_{01})_{bic}$, Savage-Dickey ratio method seems more sensitive and shows greater evidence against $H_0 : \beta = 0$. For instance, for the housing value variable, $log(B_{01})_{bic}$ refers to a *substantial evidence* for the hypothesis $H_1 : \beta \neq 0$ whereas $log(B_{01})_{sd}$ refers to a *strong evidence* for H_1.

II. Boston Housing Data

In this example, we re-analyzed the Boston housing data described by Harrsion and Rubinfeld (1978) and openly accessible in the R package "spdep". Bivand et al. (2014) analyzed the data fitting spatial lag regression model with maximum likelihood estimation. In particular, they estimated the median value of owner-occupied houses (which has been censored $50K) using 13 relevant covariates (Table 3) and spatial structure of the data (Pace and Gilley, 1997). Adjacency is defined in the R object "boston.soi" and we used the row standardized matrix as in the original study (Bivand et al., 2014). The resulted maximum likelihood estimates of the coefficients and p-values are diplayed in the last two columns of Table 3.

	Two-Sided Test		**Multiple Testing**			**ML**		
Variable	$log(B_{01})_{sd}$	$log(B_{01})_{bic}$	$log(B_{01})_m$	$log(B_{02})_m$	$log(B_{12})_m$	β	z	p
CRIM	$-\infty$	-22.056	$-\infty$	5.333	28.90	-0.0071	-7.209	0.000
ZN	2.628	2.628	3.754	2.11	-1.644	0.0004	0.986	0.324
INDUS	2.87	2.869	3.593	2.454	-1.14	0.0013	0.699	0.485
CHAS	3.072	3.072	3.328	2.869	-0.458	0.0074	0.286	0.775
I(NOX^2)	-1.47	-1.467	-2.171	4.536	6.707	-0.269	-3.031	0.002
I(RM^2)	$-\infty$	-18.534	5.267	$-\infty$	$-\infty$	0.0067	6.731	0.000
AGE	2.875	2.875	2.463	3.588	1.126	-0.0003	-0.690	0.49
log(DIS)	$-\infty$	-15.514	$-\infty$	5.189	22.016	-0.1583	-6.204	0.000
log(RAD)	-8.568	-8.373	4.955	-9.316	$-\infty$	0.0707	4.841	0.000
TAX	-4.536	-4.443	-5.244	4.761	10.005	-0.0004	-3.915	0.000
PTRATIO	-1.50	-1.50	-2.201	4.539	6.74	-0.012	-3.041	0.002
B	-3.417	-3.357	4.69	-4.123	-8.805	0.0003	3.618	0.000
log(LSTAT)	$-\infty$	-55.029	$-\infty$	5.761	65.807	-0.2321	-11.177	0.000

Table 3: Log Bayes factors for two-sided and multiple hypothesis testing for the Boston Housing data. The last three columns show the maximum likelihood estimates of the coefficients, z statistics and the resulted p-values.

Table 3 shows log Bayes factor for each variable resulted from the two-sided test using Savage-Dickey ratio ($log(B_{01})_{sd}$) and BIC approximation ($log(B_{01})_{bic}$); and for the multiple hypothesis test ($log(B_{01})_m$, $log(B_{02})_m$ and $log(B_{12})_m$). Based on the classification scheme (Table 1), Bayes factors and significance test mostly agree especially for the variables with p-values very close to zero such as CRIM, TAX and log(RAD). Likewise, when the z statistics are very close to zero, which implies higher p-values, both two-sided and multiple hypothesis tests show strong evidence for $H_0 : \beta = 0$ as in the variables CHAS, INDUS, and AGE. However, for the estimates I(NOX^ 2) and PTRATIO with the same p-values of .002, both two sided tests show *substantial evidence* for H_1, such that $log(B_{01})_{bic}$ and $log(B_{01})_{sd}$ for both variables are closer to -1.10 which is the bare minimum for the substantial evidence for H_1.

If we take I(NOXˆ 2) variable as an example, posterior probability equals 0.19 (i.e. $B_{01sd} = exp(log(B_{01})_{sd})$ and $p(H_0|y) = B_{01sd}/(B_{01sd} + 1)$), which means that the β is equal to zero with the probability of 19 percent. In this sense, multiple testing results are more informative about the effect. $log(B_{02})_m$ and $log(B_{12})_m$ show strong evidence for non-positive β. Even though, $log(B_{01})_m$, -2.171, does not show a strong evidence against H_0, multiple hypothesis testing allows us quantify the relative evidence for negative β against zero β. Based on the significance test H_0 is rejected and it is concluded that the I(NOXˆ 2) has an effect of on the outcome variable. However, this test does not inform us about how much different the *beta* from zero. Hence, p-value does not say much about the magnitude of the effects, thus, Bayes factor test is more informative in terms of magnitude of the effect.

As we observed in the simulation study, compared to BIC approximation, Savage-Dickey ratio is more sensitive for identifying the *beta* coefficients different than zero in this example. In particular, in both two-sided and multiple test log of Bayes factor for greater effects resulted in $-\infty$ whereas the greatest evidence displayed by BIC approximation resulted in -55.029 for the variable log(LSTAT).

VII. Conclusion

In this article, we developed Bayes factors for two-sided and multiple hypothesis testing procedures for β coefficients in the network autocorrelation model. The Bayesian approach to these tests provides several practical benefits to researchers compared to the classical null hypothesis significance testing. The Bayes factors allow us to quantify the relative evidence for a precise null hypothesis, or any other hypothesis and to test multiple price and interval hypotheses without any drawbacks of the classical null hypothesis significance testing. We ran a thorough simulation study to evaluate the numerical behaviors of the proposed Bayes factors for a large variety of different network setups. Simulation results showed that Bayes factor based on BIC approach is more conservative to show evidence against the null hypothesis and it is better at identifying the cases where β equals to zero regardless of the network size. On the other hand, Bayes factor based on the Savage-Dickey ratio with uniform prior is more sensitive to the changes in β and it converges to infinity as sample size increases. This may imply that consistency is more prominent in the

Savage-Dickey ratio approach.

Then, we illustrated the practical use of the proposed Bayes factors with two real data examples. The comparison between Bayes factors and p-value results shows that Bayesian answers and p-value conclusions typically agree especially as the test statistics diverges from zero point. However, significance test is not very insightful about the magnitude of the effect. Thus, significance test becomes parsimonious when the p-value is closer to the cut point, α, by either straightforwardly rejecting or accepting the null based on seemingly (non)significant p-value. In this sense, Bayes factor quantifies the statistical evidence for each hypothesis without over or underestimating the effect. Furthermore, as in the line with simulation results, we observed that Bayes factor for multiple hypothesis testing are more informative in terms of quantifying the evidence for each hypothesis compared to the ones for two-sided hypothesis testing. Hence, it is the Bayes factor we recommend as a result of this study.

The proposed Bayes factors contribute to the existing literature on testing exogenous factors' effect in network autocorrelation model which is the standard significance test which is only useful for falsifying the null. Given the importance of the network autocorrelation model across different fields, we hope that the present work will contribute to the tool kit of researchers studying with network data through the network autocorrelation model.

References

[1] Anselin, L. 1988. Spatial Econometrics: Methods and Models. Dordrecht, the Netherlands: Kluwer Academic.

[2] Anselin, L. 2002. "Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models." Agricultural Economics 27:247-67.

[3] Baller, R. D., L. Anselin, S. F. Messner, G. Deane, and D. F. Hawkins. 2001. "Structural Covariates of U.S. County Homicide Rates: Incorporating Spatial Effects." Criminology 39:561-88.

[4] Beck, N., K. S. Gleditsch, and K. Beardsley. 2006. "Space Is More than Geography: Using Spatial Econometrics in the Study of Political Economy." International Studies Quarterly 50:27-44.

[5] Bartlett, M. S. 1957. A comment on D.V. Lindley's statistical paradox. Biometrika, 44, 533–534.

[6] Bivand, R. and G. Piras. 2015. "Comparing Implementations of Estimation Methods for Spatial Econometrics." Journal of Statistical Software 63:1-36.

[7] Bivand, R. S., Gómez-Rubio, V., Rue, H. 2014. Approximate Bayesian inference for spatial econometrics models. Spatial Statistics, 9, 146–165. https://doi.org/10.1016/j.spasta.2014.01.002

[8] Butts, C.T., 2008. Social network analysis with sna. J. Stat. Softw. 24 (6), 1–51.

[9] Casella, G., F. J. Giro ́n, M. L. Martı ́nez, and E. Moreno. 2009. "Consistency of Bayesian Procedures for Variable Selection." The Annals of Statistics 37: 1207-28.

[10] Dittrich, D., Leenders, R. T. A. J., Mulder, J. 2017a. Network Autocorrelation Modeling: A Bayes Factor Approach for Testing (Multiple) Precise and Interval Hypotheses. Sociological Methods Research, 48(3), 642–676. https://doi.org/10.1177/0049124117729712

[11] Dittrich, D., Leenders, R. T., Mulder, J. 2017b. Bayesian estimation of the network autocorrelation model. Social Networks, 48, 213–236. https://doi.org/10.1016/j.socnet.2016.09.002

[12] Doreian, P. 1980. "Linear Models with Spatially Distributed Data: Spatial Disturbances or Spatial Effects?" Sociological Methods Research 9:29-60.

[13] Edwards, W., Lindman, H., Savage, L. J. 1963. Bayesian statistical inference for psychological research. Psychological Review, 70(3), 193–242. https://doi.org/10.1037/h0044139

[14] Doreian, Patrick. 1981. "Estimating Linear Models with Spatially Distributed Data." Sociological Methodology 12:359–88.

[15] Duke, J. B. 1993. "Estimation of the Network Effects Model in a Large Data Set." Sociological Methods Research 21:465-81.

[16] Fujimoto, Kayo, Chih-Ping Chou, and Thomas W. Valente. 2011. "The Network Autocorrelation Model Using Two-Mode Data: Affiliation Exposure and Potential Bias in the Autocorrelation Parameter." Social Networks 33(3):231–43.

[17] Gamerman, D., Lopes, H. F. 2006. Markov chain Monte Carlo: Stochastic simulation for Bayesian inference. Boca Raton, FL: Chapman Hall.

[18] Harrison, D., Rubinfeld, D. L. 1978. Hedonic housing prices and the demand for clean air. Journal of Environmental Economics and Management, 5(1), 81–102. https://doi.org/10.1016/0095-0696(78)90006-2

[19] Hepple, L. W. 1995. "Bayesian Techniques in Spatial and Network Econometrics: 1. Model Comparison and Posterior Odds." Environment and Planning A 27:447-69.

[20] Hoeting, J. A., Madigan, D., Raftery, A. E., Volinsky, C. T. 1999. Bayesian model averaging: A tutorial. Statistical Science, 14, 382–417.

[21] Jeffreys, H. 1961. Theory of Probability. 3rd ed. Oxford: Oxford University Press.

[22] Kass, R. E. and A. E. Raftery. 1995. "Bayes Factors." Journal of the American Statistical Association 90:773-95.

[23] Kass, R. E., Wasserman, L. 1995. A Reference Bayesian Test for Nested Hypotheses and its Relationship to the Schwarz Criterion. Journal of the American Statistical Association, 90(431), 928–934. https://doi.org/10.1080/01621459.1995.10476592

[24] Kirk, D. S. and A. V. Papachristos. 2011. "Cultural Mechanisms and the Persistence of Neighborhood Violence." American Journal of Sociology 116:1190-233.

[25] Leenders, R.T., 1995. Structure and Influence: Statistical Models for the Dynamics of Actor Attributes, Network Structure and Their Interdependence. PhD thesis. Thesis Publishers, Amsterdam.

[26] LeSage, J. P. and O. Parent. 2007. "Bayesian Model Averaging for Spatial Econometric Models." Geographical Analysis 39:241-67.

[27] Liu, C. C., Aitkin, M. 2008. Bayes factors: Prior sensitivity and model generalizability. Journal of Mathematical Psychology, 52, 362–375.

[28] Marsden, P. V. and N. E. Friedkin. 1993. "Network Studies of Social Influence." Sociological Methods Research 22: 127-51.

[29] McMillen, D. P. 2010. "Issues in Spatial Data Analysis." Journal of Regional Science 50:119-41.

[30] Mizruchi, M. S., L. B. Stearns, and C. Marquis. 2006. "The Conditional Nature of Embeddedness: A Study of Borrowing by Large US Firms, 1973-1994." American Sociological Review 71:310-33.

[31] Morey, R. D., Wagenmakers, E. J. 2014. Simple relation between Bayesian order-restricted and point-null hypothesis tests. Statistics Probability Letters, 92, 121–124. https://doi.org/10.1016/j.spl.2014.05.010

[32] Mulder, J. and E.-J. Wagenmakers. 2016. "Editors Introduction to the Special Issue "Bayes Factors for Testing Hypotheses in Psychological Research: Practical Relevance and New Developments."" Journal of Mathematical Psychology 72:1-5.

[33] Mur, J., F. Lo´pez, and A. Angulo. 2008. "Symptoms of Instability in Models of Spatial Dependence." Geographical Analysis 40:189-211.

[34] Ord, Keith. 1975. "Estimation Methods for Models of Spatial Interaction." Journal of the American Statistical Association 70(349):120–26.

[35] Pace, R. K., Gilley, O. W. 1997. Using the spatial configuration of the data to improve estimation. The Journal of Real Estate Finance and Economics, 14(3), 333–340. https://doi.org/10.1023/a:1007762613901

[36] Raftery, Adrian E., David Madigan, and Jennifer A. Hoeting. 1997. "Bayesian Model Averaging for Linear Regression Models." Journal of the American Statistical Association 92(437):179–91.

[37] Shaffer, J. P. 1995. “Multiple Hypothesis Testing.” Annual Review of Psychology 46: 561-84.

[38] Tita, G. E. and S. M. Radil. 2011. “Spatializing the Social Networks of Gangs to Explore Patterns of Violence.” Journal of Quantitative Criminology 27:521-45.

[39] Wagenmakers, E. J., Lodewyckx, T., Kuriyal, H., Grasman, R. 2010. Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. Cognitive Psychology, 60(3), 158–189. https://doi.org/10.1016/j.cogpsych.2009.12.001

[40] Wagenmakers, E.-J. 2007. A practical solution to the pervasive problems of p values. Psychonomic Bulletin Review, 14, 779–804.

[41] Weakliem, D. L. 2004. “Introduction to the Special Issue on Model Selection.” Sociological Methods Research 33:167-87.

[42] Wetzels, R. and E.-J. Wagenmakers. 2012. “A Default Bayesian Hypothesis Test for Correlations and Partial Correlations.” Psychonomic Bulletin Review 19: 1057-64.

VIII. Appendix

Please find the R-code including simulation study and application of the method to the real data examples in the following HTML link:

R Code for the Bayes Factor Testing for β coefficients in the Network Autocorrelation Model