Revista Colombiana de Estadística - Theorical Statistics
July 2022, volume 45, issue 2, pp. 341–361
http://doi.org/10.15446/rce.v45n2.92390

Spatial Econometric Models: A Bayesian Approach

Modelos econométricos espaciales: una aproximación bayesiana

Edilberto Cepeda-Cuervo[1,a], Jorge Armando Sicacha[2,b]

[1]Department of Statistics, Faculty of Sciences, Universidad Nacional de Colombia, Bogotá D.C., Colomb

[2]Department or School, Faculty or Division, University or Institution, City, Country

Abstract

In this paper we propose Bayesian methods to fit econometric regression models, including those where the variability is assumed to follow a regression structure. We formulate the main functions of the statistical R-package *BSPADATA*, developed according to the proposed methods to obtain posteriori parameter inferences. After that, we include results of simulated studies to illustrate the use of this package and the performance of the proposed methods. Finally, we provide studies to illustrate the applications of the models and compare our results with that obtained by maximum likelihood.

***Key words*:** Bayesian methods; CAR models; Spatial econometric models; SAR models.

Resumen

En este artículo proponemos métodos bayesianos para ajustar modelos de regresión econométrica, incluidos aquellos en los que la variabilidad sigue una estructura de regresión. Formulamos las principales funciones del Rpackage estadístico BSPADATA, desarrollado según los métodos propuestos para obtener inferencias de parámetros a posteriori. Luego, incluimos resultados de estudios de simulación para ilustrar el uso de este paquete y el desempeño de los métodos propuestos. Finalmente, proporcionamos estudios para ilustrar las aplicaciones de los modelos y comparamos nuestros resultados con los obtenidos por máxima verosimilitud.

***Palabras clave*:** Modelos econométricos espaciales; Modelos SAR; Modelos CAR; Métodos bayesianos.

[a]Ph.D. E-mail: ecepedac@unal.edu.co
[b]Ms.C. E-mail: jasicachap@unal.edu.co

Introduction

Spatial econometrics is a branch of the economics whose origin dates back to the Annual Meeting of the Dutch Association of Statistics 1974 (Anselin, 2001), and that since its creation has been a support investigations in economics and ecology. Usual models in the spatial econometrics are the spatial autoregressive models and the spatial error models considered in this paper (Anselin, 1988). Despite the popularity of the maximum likelihood methods to fit spatial econometric models, these lead to specification problems when the sample is very small, when there are atypical data or when there is heterogeneity of variance across space. The application of maximum likelihood methods to fit spatial economic models can be found in Anselin (1988) and LeSage (1997).

Although the possibility applying Bayesian methods to fit econometric models was considered in Hepple (1979), Anselin (1980) and Anselin (1982), it was only in 1997 that James Lesage proposed Bayesian estimation of spatial autoregressive models using MCMC algorithms (LeSage, 1997). He proposed a Bayesian method to fit homoscedastic and heteroscedastic autoregressive models. For the homoscedastic autoregressive models, he assumed diffuse prior distribution for (β, σ^2, ρ). Thus, the conditional distributions of β and σ^2 are the multivariate normal distribution and the Chi-square distribution, respectively. In this model, the conditional distribution of ρ is unknown, so he applied custom generation via "rejection sampling methods". For the heteroscedastic spatial autoregressive models he assumed that the variance-covariance matrix is given by $\sigma^2\mathbf{V}$, where $\mathbf{V} = \text{diag}(\nu_i)$, where the $\nu_i's$ are assumed to be n independent parameters to be sampled using a Chi-square distribution.

In this paper we propose Bayesian methods to fit spatial econometric models defined assuming multivariate normal prior distribution for β, inverse-gamma prior distribution for σ and uniform distributions for the spatial association parameters ρ and λ, all independent. But given that the posterior parameter distribution is unknown, samples of it are obtained from the posterior conditional distributions of parameter blocks, visited in sequence until convergence. Thus, samples of β and σ^{-2} are proposed from a multivariate normal distribution and an inverse gamma distribution, respectively, and accepted with probability 1. Samples of the spatial association parameters ρ and λ, are proposed from a kernel transition function given by a truncated normal distribution, applying the Metropolis Hastings algorithm. If a uniform distribution is assumed as the kernel transition function, Bayesian methods perform poorly.

In the spatial heteroscedastic econometric models, samples of the posterior mean regression parameter distributions, β, and the association parameters, ρ and λ, are obtained as in the homoscedastic regression models, but samples of the variance regression parameters are obtained by using the MCMC algorithm, as proposed in Cepeda-Cuervo (2001) and Cepeda-Cuervo & Gamerman (2005), where samples of the posterior variance regression distribution are proposed from a normal transition kernel defined by the combination of working observational model and a normal prior distribution.

To fit Bayesian homoscedastic and heteroscedastic regression models, we developed the *BSPADATA* R-package, which is available to users at `https://cran.r-project.org/web/packages/BSPADATA/`. To evaluate the performance of the proposed Bayesian method and of the *BSPADATA* R-package, we developed simulated and applied analyses, and we compared our result with that obtained by applying maximum likelihood methods. We used the *BSPADATA* R-package to fit the econometric models to two known datasets: level of crime in 49 districts of Columbus, Ohio, and rate of occurrence of Leukemia in New York. In the first one, the level of crime is explained by the average price of housing in each district and by the average income of its inhabitants (Anselin, 1988; LeSage & Pace, 2009). In a second application, the rate of occurrence of Leukemia in New York is the variable of interest, and is associated with factors such as the potential exposure and the proportion of cases per spatial unit (Waller, 1996).

After this introduction, the rest of the paper is organized into four sections and an appendix. Section 1 includes the Bayesian homoscedastic econometric regression models' definition, the full posterior conditional distributions and the respective Bayesian algorithm formulation. Section 2 includes the Bayesian heteroscedastic econometric regression models' definition, the full posterior conditional distributions and the respective Bayesian algorithm formulation. Section 3 presents results of three simulation studies related to SAR and SARAR/SAC homoscedastic and heteroscedastic models. Section 4 reports results of two application studies. Finally, the appendix contains some homoscedastic and heteroscedastic functions of the *BSPADATA* R-package and some theoretical developments of the proposed Bayesian methods.

1. Bayesian Homoscedastic Econometric Regression Models

Let A_i, $i = 1, 2, 3, \ldots, n$, be a partition of a region S. This partition induces a neighborhood structure $\{Ni : i = 1, \ldots, n\}$, where N_i denotes the set of all subregions that are neighbors of subregion i. The most common neighbor definitions are given by the physical first-order contiguity or by the distance between regions; however, being neighbors does not necessarily connote geographic proximity. If two subregions A_j and A_k, are in the neighborhood of region A_i, it does not mean that the interdependence between A_j and A_i, and between A_k and A_i are the same. This interdependence between A_i and A_j, $j \neq i$, is characterized by nonnegative real numbers w_{ij}, $j \in \{1, 2, \ldots, n\} - \{i\}$ such that $\sum_{j \neq i} w_{ij} = 1$. Thus, spatial interdependence of a spatial variable of interest can be written as a symmetric weight matrix W, with zeros in the main diagonal and non-negative entries outside the diagonal. Thus, assuming two spatial weight matrices W_1 and W_2, the Bayesian econometric regression model is defined as follow: let $\mathbf{Y}$ be an n-dimensional vector of a spatial variable of interest; $\mathbf{X}$ an $N \times K$ a design matrix; $\mathbf{W_1}$ and $\mathbf{W_2}$ be $N \times N$ symmetric spatial weight matrices; $(\boldsymbol{\beta}^t,\ \rho,\ \lambda)^t$ a vector of regression parameters, where $\boldsymbol{\beta}$ is a K-dimensional parameter vector and ρ and λ are related to the spatial association of $\mathbf{Y}$ and $\boldsymbol{\varepsilon}$, respectively. The general homoscedastic econometric model (SARAR/SAC) is given by:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \rho\mathbf{W_1}\mathbf{Y} + \boldsymbol{\varepsilon} \tag{1}$$
$$\boldsymbol{\varepsilon} = \lambda\mathbf{W_2}\boldsymbol{\varepsilon} + \boldsymbol{\nu}, \tag{2}$$

where $\boldsymbol{\nu} \sim N(\mathbf{0}, \boldsymbol{\Omega})$, with $\mathbf{0}$ an n-dimensional vector of zeros and $\boldsymbol{\Omega} = \sigma^2\mathbf{I}_{N\times N}$. For a Bayesian model specification, we assume the following independent prior distribution for the regression parameter: $\boldsymbol{\beta} \sim N(\mathbf{b}, \mathbf{B})$, $\sigma^2 \sim$ Inv-Gamma(ζ, ϑ), $\rho \sim U(1/\omega_{1_{\max}}, 1)$ and $\lambda \sim U(1/\omega_{2_{\max}}, 1)$, where $\omega_{1_{\max}}$ and $\omega_{2_{\max}}$ are the largest negative eigenvalues of the symmetric spatial weight matrices $\mathbf{W_1}$ and $\mathbf{W_2}$, respectively (Anselin, 1988).

The spatial autoregressive models (SAR) are defined from the homoscedastic regression models defined by (1) and (2), by setting $\lambda = 0$. If $\rho = 0$, the spatial error models (SEM) are defined.

1.1. Full Posterior Conditional Distributions

Assuming the general homoscedastic econometric model defined by equations (1) and (2), the likelihood function is given by:

$$L(\boldsymbol{\beta}, \rho, \lambda, \sigma^2) \propto |\mathbf{I} - \rho\mathbf{W_1}||\mathbf{I} - \lambda\mathbf{W_2}|(\sigma^{2n})^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2\sigma^2}\boldsymbol{\nu}^t\boldsymbol{\nu}\right\}, \tag{3}$$

where $\boldsymbol{\nu} = (\mathbf{I} - \lambda\mathbf{W_2})[(\mathbf{I} - \rho\mathbf{W_1})\mathbf{y} - \mathbf{X}\boldsymbol{\beta}]$. Thus, with the prior parameter distributions defined above, the posterior parameter distribution, obtained by applying Bayes' theorem, is given by:

$$\pi(\boldsymbol{\beta}, \sigma^2, \rho, \lambda) \propto |\sigma^2|^{-n/2}|\mathbf{A}||\mathbf{D}| \exp\left\{ -\frac{1}{2\sigma^2}(\mathbf{A}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^t\mathbf{D}^t\mathbf{D}(\mathbf{A}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\right\}\times$$
$$\exp\left\{ -\frac{1}{2}(\boldsymbol{\beta} - \mathbf{b})^t\mathbf{B^{-1}}(\boldsymbol{\beta} - \mathbf{b})\right\}\frac{\zeta^\vartheta}{\Gamma(\vartheta)}(\sigma^2)^{-\vartheta-1}\exp\left(-\zeta/\sigma^2\right)P(\rho)P(\lambda), \tag{4}$$

where $\mathbf{A} = \mathbf{I} - \rho\mathbf{W_1}$, $\mathbf{D} = \mathbf{I} - \lambda\mathbf{W_2}$ and, $P(\rho)$ and $P(\lambda)$ denote the prior of ρ and λ, respectively (Cepeda-Cuervo, 2001; Cepeda-Cuervo & Gamerman, 2000). Thus, the full posterior conditional distributions of:

1. $\boldsymbol{\beta}$, $\pi(\boldsymbol{\beta}|\sigma^2, \rho, \lambda)$, is a normal distribution $N(\mathbf{b}^*, \mathbf{B}^*)$, where $\mathbf{b}^* = \mathbf{B}^*(\sigma^{-2}\mathbf{X}^t\mathbf{D}^t\mathbf{D}\mathbf{A}\mathbf{y} + \mathbf{B^{-1}}\mathbf{b})$ and $\mathbf{B}^* = (\sigma^{-2}\mathbf{X}^t\mathbf{D}^t\mathbf{D}\mathbf{X} + \mathbf{B}^{-1})^{-1}$.

2. σ^2 is given by:

$$\pi(\sigma^2|\boldsymbol{\beta}, \rho, \lambda) \sim \text{Inv-Gamma}\left(\zeta + \frac{n}{2}, \frac{k}{2} + \vartheta\right), \tag{5}$$

where $k = (\mathbf{A}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^t\mathbf{D}^t\mathbf{D}(\mathbf{A}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$.

3. ρ is given by:

$$\pi(\rho|\boldsymbol{\beta},\sigma^2,\lambda) \propto |\mathbf{A}| \times \exp\left\{-\frac{c}{2}\left(\rho-\frac{d}{c}\right)^2\right\}P(\rho), \tag{6}$$

where $c = \sigma^{-2}\mathbf{y}^t\mathbf{W}_1^t\mathbf{D}^t\mathbf{D}\mathbf{W_1}\mathbf{y}$ and $d = \sigma^{-2}\mathbf{y}^t\mathbf{W}_1^t\mathbf{D}^t\mathbf{D}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})$.

4. λ is given by:

$$\pi(\lambda|\boldsymbol{\beta},\sigma^2,\rho) \propto |\mathbf{D}| \times \exp\left\{-\frac{c'}{2}\left(\lambda-\frac{d'}{c'}\right)^2\right\}P(\lambda), \tag{7}$$

where $c' = \sigma^{-2}(\mathbf{A}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{W}_2^t\mathbf{W_2}(\mathbf{A}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})$ and

$$d' = \sigma^{-2}(\mathbf{A}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{W}_2^t(\mathbf{A}\mathbf{y}-\mathbf{X}\boldsymbol{\beta}).$$

Note that the full conditional distributions of $\boldsymbol{\beta}$ and σ^2 are known, while the full conditional distributions of ρ and λ are analytically intractable. Therefore, to propose a Bayesian algorithm to obtain samples of the posterior distribution, we develop in Section 1.2 normal transition kernels to propose samples of the posterior conditional distributions of ρ and λ.

1.2. Bayesian Algorithm

In order to obtain samples of the posterior distribution $\pi(\boldsymbol{\beta},\sigma^2,\rho,\lambda)$, we propose an iterative process where samples of $\boldsymbol{\beta}$, σ^2, ρ and λ are obtained from their full conditional posterior distributions. Thus, samples $\boldsymbol{\beta}$ are obtained from the posterior conditional distribution $\pi(\boldsymbol{\beta}|\sigma^2,\rho,\lambda)$ and samples of σ^2 are obtained from their posterior conditional distribution $\pi(\sigma^2|\boldsymbol{\beta},\rho,\lambda)$, which are a normal and a gamma distributions, respectively.

To obtain samples of the full posterior conditional distribution of ρ and λ, taking into account that these are analytically intractable, we propose truncated normal kernel transition functions given by the normal distributions $N(c^*,d^*)$, restricted to the interval $(1/\omega_{1_{max}},1)$, to obtain samples of ρ, and the normal distribution $N(c'^*,d'^*)$, restricted to the interval $(1/\omega_{2_{max}},1)$, to obtain samples of λ. In these kernel transition function definitions c^*, d^*,c'^* and d'^* are given by:

$$c^* = \frac{\mathbf{y^t}\mathbf{W}_1^t\mathbf{D}^t\mathbf{D}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}{\mathbf{y}^t\mathbf{W}_1^t\mathbf{D}^t\mathbf{D}\mathbf{W_1}\mathbf{y}} \qquad d^* = \frac{\sigma^2}{\mathbf{y^t}\mathbf{W_1^t}\mathbf{D^t}\mathbf{D}\mathbf{W_1}\mathbf{y}}$$

$$c'^* = \frac{(\mathbf{A}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{W_2^t}(\mathbf{A}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}{(\mathbf{A}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{W_2^t}\mathbf{W_2}(\mathbf{A}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})} \qquad d'^* = \frac{\sigma^2}{(\mathbf{A}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{W_2^t}\mathbf{W_2}(\mathbf{A}\mathbf{y}-\mathbf{X}\boldsymbol{\beta})},$$

where $\mathbf{A}$ and $\mathbf{D}$ are defined as in Section 1.1. Thus, samples of the posterior distribution $\pi(\boldsymbol{\beta},\sigma^2,\rho,\lambda)$ can be obtained, given that samples of the full posterior conditional distributions of ρ and λ can be obtained by applying the Metropolis Hastings algorithm.

2. Bayesian Spatial Heteroscedastic Econometric Models

Let $\mathbf{Y}$ be an n-dimensional vector of a spatial variable of interest. If $\mathbf{Y}$, $\mathbf{X}$, $\boldsymbol{\beta}$, ρ, λ, $\mathbf{W_1}$, $\mathbf{W_2}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\nu}$ are defined as in homoscedastic econometric models, the Bayesian spatial heteroscedastic econometric model is defined by:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \rho\mathbf{W_1}\mathbf{y} + \boldsymbol{\varepsilon} \tag{8}$$

$$\boldsymbol{\varepsilon} = \lambda\mathbf{W_2}\boldsymbol{\varepsilon} + \boldsymbol{\nu}, \tag{9}$$

where $\boldsymbol{\nu} \sim N(\mathbf{0}, \boldsymbol{\Omega})$ with $\boldsymbol{\Omega}$ a diagonal matrix, with diagonal entries given by $\text{diag}(g(\Omega_{ii})) = \mathbf{Z}\boldsymbol{\gamma}$. In the variance regression structure, $\mathbf{Z}$ is an $N \times p$ design matrix of dispersion regression structure, $\boldsymbol{\gamma}$ is a p-dimensional vector variance regression parameter and g is an appropriate real function. For the Bayesian model specification, we assume the following independent prior distribution for the parameter models: $\boldsymbol{\beta} \sim N(\mathbf{b}, \mathbf{B})$, $\boldsymbol{\gamma} \sim N(\mathbf{g}, \mathbf{G})$, $\rho \sim U(1/\omega_{1_{\max}}, 1)$ and $\lambda \sim U(1/\omega_{2_{\max}}, 1)$, where $\omega_{1_{\max}}$ and $\omega_{2_{\max}}$ are defined as in Section 1.

2.1. Full Posterior Conditional Distributions

Assuming the model defined by equations (8) and (9), the likelihood function is given by:

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho, \lambda) \propto |\mathbf{I} - \rho\mathbf{W_1}||\mathbf{I} - \lambda\mathbf{W_2}||\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\boldsymbol{\nu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\nu}\right], \tag{10}$$

where $\boldsymbol{\nu} = (\mathbf{I} - \lambda\mathbf{W_2})[(\mathbf{I} - \rho\mathbf{W_1})\mathbf{y} - \mathbf{X}\boldsymbol{\beta}]$. Thus, with the prior distribution defined above, the posterior parameter distribution, obtained by applying Bayes' theorem, is given by:

$$\begin{aligned}\pi(\boldsymbol{\beta}, \boldsymbol{\gamma}, \rho, \lambda) \propto |\boldsymbol{\Sigma}|^{-1/2}|\mathbf{A}||\mathbf{D}| \exp\left\{-\frac{1}{2}(\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})^t\mathbf{D^t}\boldsymbol{\Sigma}^{-1}\mathbf{D}(\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})\right\}\times \\ \exp\left\{-\frac{1}{2}(\boldsymbol{\beta} - \mathbf{b})^t\mathbf{B^{-1}}(\boldsymbol{\beta} - \mathbf{b})\right\}\exp\left\{-\frac{1}{2}(\boldsymbol{\gamma} - \mathbf{g})^t\mathbf{G^{-1}}(\boldsymbol{\gamma} - \mathbf{g})\right\}P(\rho)P(\lambda),\end{aligned} \tag{11}$$

where $\mathbf{A}$ and $\mathbf{D}$ is defined as in Section 1.1. Thus, from 11, the full posterior conditional distribution of:

1. $\boldsymbol{\beta}$, $\pi(\boldsymbol{\beta}|\boldsymbol{\gamma}, \rho, \lambda)$, is a normal distribution $N(\mathbf{b}^*, \mathbf{B}^*)$, where $\mathbf{b}^* = \mathbf{B}^*(\mathbf{X}^t\mathbf{D}^t\boldsymbol{\Sigma}^{-1}\mathbf{DAy} + \mathbf{B^{-1}b})$ and $\mathbf{B}^* = (\mathbf{X}^t\mathbf{D}^t\boldsymbol{\Sigma}^{-1}\mathbf{DX} + \mathbf{B^{-1}})^{-1}$ (Cepeda-Cuervo, 2001; Cepeda-Cuervo & Gamerman, 2005).

2. $\boldsymbol{\gamma}$ is given by:

$$\begin{aligned}\pi(\boldsymbol{\gamma}|\boldsymbol{\beta}, \gamma, \lambda) \propto &|\boldsymbol{\Sigma}|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})^t\mathbf{D}^t\boldsymbol{\Sigma}^{-1}\mathbf{D}(\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})\right\}\times \\ &\exp\left\{-\frac{1}{2}(\boldsymbol{\gamma} - \mathbf{g})^t\mathbf{G^{-1}}(\boldsymbol{\gamma} - \mathbf{g})\right\}.\end{aligned}$$

3. ρ is given by:

$$\pi(\rho|\lambda, \boldsymbol{\gamma}, \boldsymbol{\beta}) \propto |\mathbf{A}| \exp\left\{-\frac{c}{2}\left(\rho - \frac{d}{c}\right)^2\right\} P(\rho),$$

   where $c = \mathbf{y^t W_1^t D^t} \boldsymbol{\Sigma}^{-1} \mathbf{D W_1 y}$ and $d = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^t \mathbf{D^t} \boldsymbol{\Sigma}^{-1} \mathbf{D W_1 y}$.

4. λ is given by:

$$\pi(\lambda|\rho, \boldsymbol{\gamma}, \boldsymbol{\beta}) \propto |\mathbf{D}| \exp\left\{-\frac{c'}{2}\left(\lambda - \frac{d'}{c'}\right)^2\right\} P(\lambda),$$

   where $c = (\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})^t \mathbf{W_2^t} \boldsymbol{\Sigma}^{-1} \mathbf{W_2} (\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})$ and

$$d = (\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})^t \mathbf{W_2^t} \boldsymbol{\Sigma}^{-1} (\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta}).$$

The full conditional distribution of $\boldsymbol{\beta}$ is known, while the full conditional distribution of $\boldsymbol{\gamma}$, ρ and λ are analytically intractable. Thus, in order to propose a Bayesian algorithm to obtain samples of the posterior parameter distribution, we develop in Section 2.2 normal transition kernels to propose samples of the posterior conditional distributions of $\boldsymbol{\gamma}$, ρ and λ.

2.2. Proposed Bayesian Algorithm

To obtain samples of the posterior parameter distribution of $\boldsymbol{\theta} = (\boldsymbol{\beta}^t, \boldsymbol{\gamma}, \rho, \lambda)^t$, we propose an interactive process, where samples of $\boldsymbol{\beta}$ are obtained from their posterior conditional distribution $\pi(\boldsymbol{\beta}|\gamma, \rho, \lambda)$, obtained in Section 2.1, while samples of $\boldsymbol{\gamma}$, ρ and λ are obtained by applying the Metropolis-Hastings algorithm, proposing samples from the following kernel transition functions.

1. To obtain samples of the full posterior conditional distribution of $\boldsymbol{\gamma}$, we build a kernel transition function applying the method proposed in Cepeda-Cuervo (2001), and Cepeda-Cuervo & Gamerman (2005) as follows. Given that the random variables $t_i = (y_i - \mu_i)^2 \sim \sigma_i^2 \chi_1^2$, $i = 1, 2, \ldots, n$, have mean and variance given by $E(t_i) = \sigma_i^2$ and $V(t_i) = 2\sigma_i^4$, assuming a differentiable dispersion link function $g(.)$, that takes into account the positivity of the variance, the working variables $\tilde{y}_i$'s are obtained by first-order Taylor approximation of g around $E(t_i) = \sigma_i^2$. That is,

$$g(t_i) \simeq g[E(t_i)] + g'[E(t_i)][t_i - E(t_i)] = \tilde{y}_i, \tag{12}$$

   where g' denotes the first-order derivative of g. For these variables, $E(\tilde{y}) = g(E(t_i))$ and $Var(\tilde{y}) = g'[E(t_i)]^2 Var(t_i)$. When g is the logarithmic function, the working observation variables are given by:

$$\tilde{y}_i = \mathbf{z}_i^t \boldsymbol{\gamma}^{(c)} + \frac{(y_i - \rho^{(c)} W_{1i} y_i - \lambda^{(c)} W_{2i} y_i + \lambda^{(c)} \rho^{(c)} W_{2i} W_{1i} y_i - \mathbf{x}_i^t \boldsymbol{\beta}^{(c)} + \lambda^{(c)} W_{2i} \mathbf{x}_i^t \boldsymbol{\beta}^{(c)})^2}{\exp(\mathbf{z}_i^t \boldsymbol{\gamma}^{(c)})} - 1,$$

for which $E(\tilde{y}_i) = \boldsymbol{z}_i^t\boldsymbol{\gamma}^{(c)}$ and $Var(\tilde{y}_i) = 2$. Thus, the kernel transition function to propose samples of the variance regression parameters $\boldsymbol{\gamma}$ is given by the posterior distribution obtained by the combination of a normal prior distribution and the working observational models given by $\tilde{\mathbf{y}} \sim N(\boldsymbol{z}_i^t\boldsymbol{\gamma}^{(c)}, 2)$:

$$q_\gamma(\boldsymbol{\gamma}^{(c)}, \boldsymbol{\gamma}^{(n)}) = N(\mathbf{g}^*, \mathbf{G}^*),$$

where $\mathbf{g}^* = \mathbf{G}^*(\mathbf{G^{-1}g} + 0.5\mathbf{Z^t\tilde{Y}})$ and $\mathbf{G}^* = (\mathbf{G^{-1}} + 0.5\mathbf{Z^tZ})^{-1}$.

2. To obtain samples of the full posterior conditional distribution of ρ, we propose a normal transition kernel transition function $N(c^*, d^*)$, where:

$$c^* = \frac{\mathbf{y}^t\mathbf{W}_1^t\mathbf{D}^t\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})}{\mathbf{y}^t\mathbf{W}_1^t\mathbf{D}^t\boldsymbol{\Sigma}^{-1}\mathbf{D}\mathbf{W}_1\mathbf{y}} \qquad d^* = \frac{1}{\mathbf{y}^t\mathbf{W}_1^t\mathbf{D}^t\boldsymbol{\Sigma}^{-1}\mathbf{D}\mathbf{W}_1\mathbf{y}}$$

with $\mathbf{D} = \mathbf{I} - \lambda\mathbf{W_2}$, truncated by the interval $(1/\omega_{1_{\max}}, 1)$, where $\omega_{1_{\max}}$ is the largest negative eigenvalue of the neighbor matrix $\mathbf{W_1}$ (Anselin, 1988).

3. To propose samples of the posterior conditional distribution of λ, we propose a sample transition kernel given by a normal transition kernel $N(c^*, d^*)$, where

$$c^* = \frac{(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{W}_2^t\boldsymbol{\Sigma}^{-1}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})}{(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{W}_2^t\boldsymbol{\Sigma}^{-1}\mathbf{W_2}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})} \qquad d^* = \frac{1}{(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{W}_2^t\boldsymbol{\Sigma}^{-1}\mathbf{W_2}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})}$$

with $\mathbf{A} = \mathbf{I} - \rho\mathbf{W_1}$. This normal distribution is truncated in the range $(1/\omega_{2_{\max}}, 1)$, where $\omega_{2_{\max}}$ is the larger negative eigenvalue of the neighborhood matrix, $\mathbf{W_2}$ (Anselin, 1988). It is implemented in the function *hetero_general* of the library *BSPADATA* of R, (**?**).

To obtain samples of ρ and λ we also propose uniform kernel transition functions in the intervals $(1/\omega_{1_{\max}}, 1)$ and $(1/\omega_{2_{\max}}, 1)$, respectively.

3. Simulated Studies

To illustrate the performance of the Bayesian methods proposed to fit spatial econometric models, the results of two simulation studies are presented. In what follows, we use a binary neighbor matrix of the Columbus data (**?**), standardized by rows, where if two spatial units i and j have a common border, the (i, j)-input of the binary neighbor matrix is 1 and, if spatial units i and j have no a common border, the (i, j)-input of the binary neighbor matrix is 0.

3.1. A first Simulation Study

In a first study, we assume that the variable of interest follows a normal spatial homoscedastic distribution model given by:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \rho\mathbf{W}\mathbf{Y} + \boldsymbol{\varepsilon}, \text{ where } \boldsymbol{\varepsilon} \sim N(0, \sigma^2 I), \tag{13}$$

as in equations (1) and (2), but with $\lambda = 0$. Thus, to generate a dataset for the analysis, assuming two explanatory variables, samples of size $n = 100$ were generated from $\mathbf{X}_1 \sim U(0, 400)$ and $\mathbf{X}_2 \sim U(10, 23)$, respectively. In all simulations, it is assumed that $\mathbf{X}_0 = \mathbf{I}_{100}$. With these samples and assuming $\boldsymbol{\beta} = (18, 0.478, -1.3)$, the mean vector value was obtained from $\boldsymbol{\mu} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\beta$, where $\mathbf{W}$ is the neighborhood matrix defined in **?** and $0 < \rho < 1$, in order to assume spatial dependence. Assuming $\sigma^2 = 45$, the variance-covariance matrix of $\mathbf{Y}$ was obtained from $V(\mathbf{Y}) = \sigma^2[(\mathbf{I} - \rho\mathbf{W})(\mathbf{I} - \rho\mathbf{W})]^{-1}$. Hence, a sample of $\mathbf{Y}$ was obtained from a multivariate normal distribution $N(\boldsymbol{\mu}, V(\mathbf{Y}))$. In this way, three datasets were obtained, one for each of the three values of $\rho = 0.12$, 0.52 and 0.90.

To complete the Bayesian model specifications, we assume independent normal distributions $\beta_k \sim N(0, 10^5)$, $k = 1, 2, 3$, for the regression parameters, and $\sigma^2 \sim G(a, b)$, with $a = b = 0.0001$ for the variance parameters. With this model specification, the posterior parameter estimates, obtained by applying the proposed Bayesian method and maximum likelihood methods are given in Table 1. The acceptance rate in the proposed Bayesian method is 79%.

Table 1: Homoscedastic SAR models: Parameter estimates

ρ real	Método	Valor	β_0	β_1	β_2	ρ	σ^2	BIC	DIC
0.12	Bay	$\hat{\theta}$	20,595	0,490	−1,429	0,081	54,229	258,551	894,800
		s.d.	6,233	0,008	0,276	0,033	11,997		
	ML	$\hat{\theta}$	21,658	0,489	−1,460	0,076	47,957	258,175	
		s.d.	8,386	0,008	0,261	0,035	9,689		
0,52	Bay	$\hat{\theta}$	20,388	0,483	−1,321	0,510	44,225	251,708	874,179
		s.d.	5,523	0,008	0,231	0,020	9,625		
	ML	$\hat{\theta}$	21,596	0,482	−1,340	0,506	38,971	251,284	
		s.d.	5,613	0,008	0,219	0,022	7,878		
0,90	Bay	$\hat{\theta}$	19,617	0,493	−1,332	0,895	97,175	302,857	1027,753
		s.d.	13,237	0,013	0,374	0,012	21,272		
	ML	$\hat{\theta}$	27,708	0,493	−1,393	0,888	86,669	302,293	
		s.d.	17,383	0,012	0,368	0,016	17,60345		

Table 1 shows that the Bayesian and maximum likelihood parameter estimates are close to the true values, all of with small standard deviations, except for estimates of σ^2 when $\rho = 0.90$, in which there is a high spatial association of the variable of interest.

3.2. A Second Simulated Study

In a second study, we assume that the variable of interest follows a spatial structure given by:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \text{ where } \boldsymbol{\varepsilon} = \lambda\mathbf{W_2}\boldsymbol{\varepsilon} + \boldsymbol{\nu} \text{ and } \boldsymbol{\nu} \sim N(0, \sigma^2\mathbf{I}). \tag{14}$$

Thus, assuming two explanatory variables $\mathbf{X}_1$ and $\mathbf{X}_2$, as in Section 3.1, $\boldsymbol{\beta}^t = (18, 0.026, -0.4)$ and $\sigma^2 = 45$, the mean is given by $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ and the variance given by $V(\mathbf{Y}) = \sigma^2[(\mathbf{I} - \lambda\mathbf{W})(\mathbf{I} - \lambda\mathbf{W})]^{-1}$, where $\mathbf{W}$ is the weighted neighborhood matrix defined in **?** and λ is a real number $0 < \lambda < 1$. Thus, three samples of $\mathbf{Y}$

were obtained from a multivariate normal distribution $N(\boldsymbol{\mu}, V(\mathbf{Y}))$. One for each of the values of $\lambda = 0.25$, 0.55 and 0.85, as in Section 3.1.

To complete the Bayesian model specifications, we assume independent normal distributions $\beta_k \sim N(0, 10^5)$, $k = 1, 2, 3$, for the regression parameters, and $\sigma^2 \sim G(a, b)$, with $a = b = 0.0001$. The posterior parameter estimates, obtained by applying the proposed Bayesian method and maximum likelihood methods are given in Table 2. All are close to the true values. This shows the good performance of the Bayesian estimation method. In this simulation, the acceptance rates of samples of λ are lower when $\lambda = 0.55$ and when $\lambda = 0.85$, since they are smaller than 30%. When $\lambda = 0.25$, the acceptance is bigger than 60%. The BIC and DIC vales are smaller for $\lambda = 0.25$.

Table 2: SAR Homoscedastic model: posterior parameter estimates.

λ	Método	Valor	β_0	β_1	β_2	λ	σ^2	BIC	DIC
0, 25	Bay	$\hat{\theta}$	23, 130	0, 018	−0, 768	0, 383	43, 240	249, 349	869, 568
		s.d.	5, 047	0, 008	0, 274	0, 245	9, 601		
	ML	$\hat{\theta}$	23, 989	0, 017	−0, 800	0, 214	38, 668	248, 084	
		s.d.	4, 787	0, 008	0, 265	0, 188	7, 848		
0, 55	Bay	$\hat{\theta}$	15, 858	0, 031	−0, 438	0, 843	44, 082	260, 609	909, 672
		s.d.	6, 941	0, 007	0, 235	0, 132	9, 607		
	ML	$\hat{\theta}$	16, 838	0, 033	−0, 447	0, 520	41, 691	254, 807	
		s.d.	4, 454	0, 007	0, 232	0, 142	8, 638		
0, 75	Bay	$\hat{\theta}$	13, 313	0, 026	−0, 426	0, 902	48, 816	269, 141	931, 869
		s.d.	9, 546	0, 008	0, 270	0, 078	10, 630		
	ML	$\hat{\theta}$	14, 918	0, 025	−0, 465	0, 713	45, 971	263, 806	
		s.d.	5, 807	0, 008	0, 269	0, 100	9, 687		

3.3. A Third Simulation Study

In this section, we present results of two simulations, one assuming the SAR heteroscedastic model and other assuming the SARAR/SAC heteroscedastic model.

1. **SAR heteroscedastic model** . In this case, we assume that the variable of interest follows a spatial structure given by:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \rho\mathbf{W_2}\mathbf{Y} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{\Omega}) \tag{15}$$

where $\mathbf{\Omega} = \text{diag}(\Omega_{ii})$ and $\log(\Omega_{ii}) = \gamma_0 + \gamma_1 z_{1i} + \gamma_2 z_{2i}$. To generate a dataset for the analysis, a sample of size $n = 100$ was generated from each of the following uniform distributions: $\mathbf{X}_1 \sim U(0, 400)$, $\mathbf{X}_2 \sim U(10, 23)$ and $\mathbf{X}_3 \sim U(0, 10)$. Thus, assuming that $\mathbf{X} = (\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2)$, where $\mathbf{X}_0 = \mathbf{I}_{100}$, a 100-dimensional unit vector, the mean vector value was obtained from $\boldsymbol{\mu} = (\mathbf{I}_{100} - \rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}$, where $\mathbf{W}$ is the neighborhood matrix defined in **?**, $\boldsymbol{\beta} = (-35, 0.35, -1.7)^t$, and ρ is a real number $(0 < \rho < 1)$. To obtain the variance of the variable of interest we assume that variance explanatory variables are given by $\mathbf{Z} = (\mathbf{Z}_0, \mathbf{Z}_1, \mathbf{Z}_2)$, where $\mathbf{Z}_0 = \mathbf{X}_0$, $\mathbf{Z}_1 = \mathbf{X}_1$ and

$\mathbf{Z}_2 = \mathbf{X}_3$. Thus, assuming that $\mathbf{\Omega}$ is a diagonal covariance matrix, such that $\log(\mathbf{\Omega}_{ii}) = -8 + 0.026x_1 - 0.4x_3$, the variance-covariance matrix of $\boldsymbol{Y}$ was obtained from $V(\mathbf{Y}) = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{\Omega}(\mathbf{I} - \rho\mathbf{W})^{-1}$. Finally, with the mean and the variance of $\boldsymbol{Y}$ thus obtained, three samples of the variable of interest $\mathbf{Y}$ were obtained, one for each of values of ρ: 0.12, 0.52, 0.90.

Assuming independent normal distributions $\boldsymbol{\beta} \sim N(\mathbf{0}, \mathbf{I}10^5)$ and $\boldsymbol{\gamma} \sim N(\mathbf{0}, \mathbf{I}10^5)$, for the regression parameters, the SAR heteroscedastic model 15 was fitted to each of the three simulated datasets. The posterior parameter estimates and standard deviations are reported in Table 3.

TABLE 3: Heteroscedastic SAR models: Parameter estimates

ρ	$\hat{\theta}$/s.d.	β_0	β_1	β_2	γ_0	γ_1	γ_2	ρ	BIC	DIC
0,12	$\hat{\theta}$	−34,975	0,350	−1,700	−9,201	0,026	−0,226	0,120	−155,610	−369,127
	s.d.	0,020	0,002	0,001	0,616	0,002	0,073	0,000		
0,52	$\hat{\theta}$	−35,021	0,350	−1,696	−5,786	0,016	−0,094	0,516	−73,282	−90,000
	s.d.	0,069	0,003	0,005	0,597	0,000	0,106	0,001		
0,75	$\hat{\theta}$	−34,806	0,350	−1,710	−4,541	0,012	0,031	0,749	−16,782	49,117
	s.d.	0,175	0,000	0,010	0,680	0,002	0,094	0,001		

For all values of ρ, the mean regression parameter estimates are close to the true values and all have small standard deviations. The estimates of variance regression parameters, γ_i, $i = 0, 1, 2$, change with values of ρ: the differences between true and estimated parameters increase with ρ. In all cases, the estimates of ρ are close to the true values with smalls standard deviations.

2. **General (SARAR/SAC) heteroscedastic model**. In this section we consider the model where the autoregressive effect is included in the error term. That is, we assume that:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \rho\mathbf{W}_1\mathbf{Y} + \boldsymbol{\varepsilon},$$

where $\boldsymbol{\varepsilon} = \lambda\mathbf{W}\boldsymbol{\varepsilon} + \boldsymbol{\nu}$ and $\boldsymbol{\nu} \sim N(\mathbf{0}, \mathbf{\Omega})$. In this model, the explanatory variables and the parameter vector are defined as in 15. The mean vector value is obtained from $\boldsymbol{\mu} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}$, where $\mathbf{W}_1$ is the neighborhood matrix defined in **?**, $\boldsymbol{\beta} = (-35, 0.35, -1.7)$, and ρ is a real number $(0 < \rho < 1)$. The variance-covariance matrix of $\mathbf{Y}$ is given by:

$$Var(\mathbf{Y}) = (\mathbf{I} - \rho\mathbf{W})^{-1}(\mathbf{I} - \lambda\mathbf{W}_2)^{-1}\mathbf{\Omega}(\mathbf{I} - \rho\mathbf{W})^{-1}(\mathbf{I} - \lambda\mathbf{W}_2)^{-1},$$

where $\mathbf{\Omega}$ is a diagonal matrix, with $\log(\mathbf{\Omega}_{ii}) = -8 + 0.026x_1 - 0.4x_3$. With the mean and the variance thus obtained, three samples of the variable of interest $\mathbf{Y}$ were obtained, one for each of the vector values (ρ, λ)=$(0.12, 0.48)$, $(0.45, 0.65)$ and $(0.70, 0.20)$, from a multivariate normal distribution $N(\boldsymbol{\mu}, Var(\mathbf{Y}))$. The parameter estimates obtained by applying maximum likelihood and Bayesian methods are given in Table 4.

As in the SAR heteroscedastic models, for all values of ρ and λ, the mean regression parameter estimates are close to the true values and all have small standard deviations. The estimates of variance regression parameters, γ_i,

Table 4: General (SARAR/SAC) heterocedastic models: parameter estimates

ρ	λ	$\hat{\theta}$/s.d.	β_0	β_1	β_2	γ_1	γ_2	γ_3	ρ	λ	BIC	DIC
$0,12$	$0,48$	$\hat{\theta}$	$-34,736$	$0,349$	$-1,712$	$-7,252$	$0,023$	$-0,301$	$0,120$	$0,392$	$-92,669$	$-187,788$
		s.d.	$0,095$	$0,000$	$0,004$	$0,602$	$0,002$	$0,070$	$0,001$	$0,030$		
$0,45$	$0,65$	$\hat{\theta}$	$-33,517$	$0,349$	$-1,706$	$-0,813$	$-0,001$	$-0,176$	$0,445$	$0,953$	$2,616$	$102,289$
		s.d.	$9,400$	$0,000$	$0,016$	$0,819$	$0,002$	$0,092$	$0,005$	$0,057$		
$0,70$	$0,20$	$\hat{\theta}$	$-35,382$	$0,350$	$-1,675$	$-2,629$	$0,004$	$-0,373$	$0,699$	$0,231$	$-86,014$	$-166,748$
		s.d.	$0,115$	$0,000$	$0,005$	$0,594$	$0,002$	$0,083$	$0,001$	$0,090$		

$i = 0, 1, 2$, change with values of ρ. Differences between true and estimated parameters are smaller for $\rho = 0.12$ and the differences between $\gamma_i - \hat{\gamma}_i$, $i = 1, 2, \ldots, n$ increase when ρ increases. The estimates of ρ and λ, are close to the true values and all have small standard deviation, except when $\rho = 0.45$ and $\lambda = 0.65$.

4. Applications

4.1. Columbus Crime Data

In order to compare the results obtained by applying the proposed Bayesian method with that obtained by applying maximum likelihood, the Columbus crime data presented in Anselin (1988) were analyzed. These data are related to the crime level in 49 districts of Columbus (Ohio), associated with income measures and housing values. In this analysis, it is assumed that the variable of interest is $CRIME$, denoting residential burglaries and car thefts for every 1000 homes, and that the explanatory variables are $HOUSE$, average value of the residence in USS 1000, and INC, average household income in USS 1000. The neighborhoods matrix, available in the dataset *columbus*, is given by the standardized binary contiguity by rows.

In the analysis of this dataset, the following homoscedastic models were fitted: SAR model (16), SEM model (17) and the general (SARAR/SAC) homoscedastic model (18), where $\varepsilon \sim N(0, \sigma^2)$ and $\nu_i \sim N(0, \sigma^2)$.

$$CRIME = \beta_0 + \beta_1 HOUSE + \beta_2 INC + \rho \mathbf{W_1} CRIME + \boldsymbol{\varepsilon} \tag{16}$$

$$CRIME = \beta_0 + \beta_1 HOUSE + \beta_2 INC + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} = \lambda \mathbf{W_2} \boldsymbol{\varepsilon} + \boldsymbol{\nu} \tag{17}$$

$$CRIME = \beta_0 + \beta_1 HOUSE + \beta_2 INC + \rho \mathbf{W_1} CRIME + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} = \lambda \mathbf{W_2} \boldsymbol{\varepsilon} + \boldsymbol{\nu} \tag{18}$$

All cases assumed the following prior distributions: $\boldsymbol{\beta} \sim N(0\mathbf{I}_3, 10^4 \mathbf{I}_{3\times3})$, $\rho \sim N(0, 10^4)$ and $\sigma^2 \sim GammaInv(0.01, 0.01)$, where $I_{3\times3}$ is the 3×3 identity matrix.

For SAR, SEM and General homoscedastic regression models, the parameter estimates and respective standard deviation, obtained by applying Bayesian methods (B) and obtained by applying maximum likelihood (ML) are given in Table 5. The acceptance rates are: 62.88% for ρ in the SAR models, 40.73% for λ in the SARAR/SAC models and, 73.77% for ρ, 71.11% for λ in the SARAR/SAC model.

Table 5: SAR, SEM and General (SARAR/SAC) homoscedastic regression models: Parameter estimates

	Meth.	β_0	β_1	β_2	σ^2	ρ	λ	BIC	DIC
SAR homoscedastic regression models									
$\hat{\theta}$	B	47.441	-0.308	-1.079	120.672	0.441	$--$	300.203	1010.982
s.d.		7.087	0.099	0.343	26.660	0.125	$--$	327.51	$--$
$\hat{\theta}$	ML	54.201	-0.305	-1.254	107.460	0.315	$--$	$--$	$--$
s.d.		7.295	0.094	0.317	10.366	0.134	$--$	$--$	$--$
SEM homoscedastic regression models									
$\hat{\theta}$	B	60.484	-0.336	-0.936	123.714	$--$	0.685	306.146	1043.28
s.d.		8.068	0.099	0.392	27.440		0.220	$--$	$--$
$\hat{\theta}$	ML	59.893	-0.302	-0.941	98.575	$--$	0.562	304.262	$--$
s.d.		5.366	0.090	0.331	19.8737		0.134	$--$	$--$
General (SARAR/SAC) homoscedastic regression models									
$\hat{\theta}$	B	28,676	$-0,125$	$-0,963$	93,504	0,726	$-1,408$	321,595	1069,836
s.d.		6,640	0,076	0,232	20,753	0,087	0,335	$--$	$--$
$\hat{\theta}$	ML	47,784	$-0,282$	$-1,026$	95,604	0,091	0,167	342,849	
s.d.		9,903	0,090	0,326	19,475	0,197	0,297	$--$	$--$

In general, the parameter estimates of the models obtained by applying Bayesian methods and maximum likelihood methods agree (Table 5). The maximum likelihood estimates presented in the table are those reported by (Neath & Cavanaugh, 2012). Although estimates related to the SAR model 16 show differences between parameter estimates of σ^2 and λ, the BIC values associated with the posterior parameter estimates are smaller than the BIC value associated with estimates obtained by applying maximum likelihood. For the General homoscedastic model 18, the ML and Bayesian parameter estimates of $\boldsymbol{\beta}$ are very similar, despite the differences of standard errors. In this model, although the ML and Bayesian parameter estimates of σ^2 and λ also present larger differences, in both parameter estimates and standard errors the BIC indicates that the best fit is obtained when the Bayesian method is applied. The best model (that with the lowest DIC and BIC values) is the homoscedastic SAR model.

4.2. Leukemia in New York Data

In this section we analyze the New York dataset presented in Waller & Gotway (2004), associated with the risk factors of leukemia in New York between 1978 and 1982. This dataset is related to the following variables:

- Z: Proportion of patients with leukemia in the population at risk. This ratio was transformed in Waller & Gotway (2004) for use in models where normal errors are assumed.
- $PCTAGE65P$: Percentage of people in each space unit 65 years of age.
- $PCTOWNHOME$: Percentage of people in each space unit with their own home.

- $PEXPOSURE$: Potential exposure to risk factors defined as the inverse distance between each spatial unit and the nearest site where trichloroethylene is found.

Given that this dataset has been used in various epidemiological studies for different purposes (Waller, 1996; Gangnon & Clayton, 1998; Waller & Gotway, 2004), we only try to understand the Bayesian method proposed and the use of the $BSPADATA$ r-package, not to understand the behavior of leukemia rates. Thus, we fit the model presented in Waller & Gotway (2004) given by:

$$Z = \beta_0 + \beta_1 PCTAGE65P + \beta_2 PCTOWNHOME + \rho \mathbf{W_1 Z} + \boldsymbol{\varepsilon}$$
$$\boldsymbol{\varepsilon} = \lambda \mathbf{W_2} \boldsymbol{\varepsilon} + \boldsymbol{\nu},$$

where $\boldsymbol{\nu}$ is a normal error term of mean $\mathbf{0}$ and diagonal variance matrix $\mathbf{\Omega}$, which have diagonal values given by $\Omega_{ii} = \exp(\gamma_0 + \gamma_1 PEXPOSURE_i)$, $i = 1, \ldots n$. In this case we assume $\mathbf{W_1} = \mathbf{W_2}$, where $\mathbf{W_1}$ is a matrix obtained from the *listw_NY* object of the dataset NY_data of R (**?**). It is a matrix of binary contiguity, standardized by rows, where if a pair of spatial units, i and j, are neighbors, then the input $[i, j]$ of the matrix is different from 0, guaranteeing that the sum of each row of the matrix is 1. The vector pair $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ have normal prior distributions, $N(0.10000\mathbf{I})$.

For SAR, SEM, and General (SARAR/SAC) heteroscedastic regression models, the parameter estimates and respective standard deviations, obtained by applying the proposed Bayesian methods, are given in Table 6. The acceptance rates are: 50.92 % for λ and 60.65 % for ρ in the SAR model, 50.90 % for λ and 41.15 % for ρ in the SEM model and; 50.22 % for λ , 0.06 % for ρ and 22.31 % for λ in the general (SARAR/SAC) heteroscedastic model.

TABLE 6: Leukemia application: Parameter estimates.

SAR heteroscedastic regression models									
Method	β_0	β_1	β_2	γ_0	γ_1	ρ	λ	BIC	DIC
$\hat{\theta}$	$-0,397$	$3,622$	$-0,418$	$-0,594$	$-0,164$	$0,298$	$--$	$64,903$	$1059,569$
s.d.	$0,143$	$0,617$	$0,171$	$0,164$	$0,078$	$0,099$		$--$	$--$
SEM heterocedastic regression models									
$\hat{\theta}$	$-0,458$	$3,887$	$-0,464$	$-0,570$	$-0,175$	$0,372$	$--$	$69,909$	$1087,139$
s.d.	$0,163$	$0,634$	$0,200$	$0,163$	0.077	$0,128$		$--$	$--$
General (SARAR/SAC) heteroscedaastic models									
$\hat{\theta}$	-0.345	2.493	-0.071	-1.044	-0.176	-1.816	0.996	$209,042$	$1527,007$
s.d.	4.872	0.525	0.197	0.162	0.077	0.134	0.005		

The results show agreement between parameter estimates of the SARAR/SEC and SAR models, but the BIC (DIC) value of the SAR model are the smallest. However, the parameter estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ do not agree with the parameter estimates of the homoscedastic general model, which have a bigger BIC (DIC) value and smaller acceptance rates.

As in the first aplication of the Columbus crime data, from the BIC (DIC) values we conclude that the homoscedastic and heteroscedastic general

(SARAR/SAC) models do not fit are not fit the respective datasets well, possible due to misidentification of problems, taking into acount that in these aplications we assume that $W_1 = W_2$.

5. Conclusions

In this paper we propose Bayesian methods for econometric regression models, including those where the variability is assumed to follow a regression structure. We also introduce the BSPADATA R-package, which can be used to fit Bayesian econometric regression models by applying the Bayesian method proposed by Cepeda-Cuervo (2001) and Cepeda-Cuervo & Gamerman (2005), and summarize the results of the posterior inferences obtained in studies of simulations and applications. In these simulations, we consider spatial homoscedastic regression models with $\rho = 0.12, 0.52, 0.90$ and $\lambda = 0$ and with $\rho = 0$ and $\lambda = 0.25, 0.52, 0.90$, and SEM heteroscedastic regression models with $\rho = 0.12, 0.52, 0.75$. In all of them, the posterior inferences show good performance of the proposed Bayesian methods: all the parameter estimates are close to the true parameter values and the respective standard deviations are small. However, researches and students can develop new simulations using the BSPADATA R-package to verify the performance of the Bayesian methods proposed here, assuming different values of the variance in the homoscedastic econometric models and diferent variancee regression structures in heteroscedastic models. Finally, we provide results of the applications of the models and compare our results with that obtained by maximum likelihood.

In the Bayesian method proposed to fit spatial econometric models, the normal transition kernel shows good performance for β, γ or σ, and ρ in fitting of spatial econometric models. However, if the kernel transition function is a Gaussian one, the results are better than if the kernel is a uniform distribution, so the review of the simulations allows us to identify at least a possibility of extending this Bayesian method, proposing a new kernel transition function for λ.

Appendix A.

The BSPADATA R-package is the computational implementation of the spatial econometric models proposed in sections 1 and 2, under an innovative Bayesian approach. This package is composed of six functions. Three of them (*hom_sar*, *hom_sem* and *hom_general*) fit the Bayesian homoscedastic econometric regression models presented in section 1, and the other three (*hetero_sar*, *hetero_sem* and textithetero_general) to fit the Bayesian heteroscedastic econometric regression models presented in Section 3.

These functions enable obtaining posterior parameter estimates as their standard errors. Users can also retrieve the Markov chains associated with each parameter and their graphical representation. Besides this, they can obtain the Bayesian information criteria (BIC), the deviance information criterion (DIC) and the acceptance rates for those parameters, whose sampling involves a Metropolis-Hastings step.

Appendix A.1. Homoscedastic Models

The BSPADATA R-package functions *hom_sar*, *hom_sem* and *hom_general* can be used to fit the Bayesian homoscedastic SAR, SEM and General spatial econometric models, respectively. The syntaxes of these functions are:

- *hom_sar*

```
hom_sar(y,X,W,nsim,burn,step,b_pri,B_pri,r_pri,lambda_pri,
beta_0,sigma2_0,rho_0,kernel,plot,chains)
```

- *hom_sem*

```
hom_sem(y,X,W,nsim,burn,step,b_pri,B_pri,r_pri,lambda_pri,
beta_0,sigma2_0,lambda_0,kernel,plot,chains)
```

- *hom_general*

```
hom_general(y,X,W1,W2,nsim,burn,step,b_pri,B_pri,r_pri,
lambda_pri,beta_0,sigma2_0,rho_0,lambda_0,kernel,mateq,
plot,chains)
```

For these functions, `y`, `X` ,`W1` and `W2` represent the response variable, the explanatory variables and the spatial contiguity matrices, respectively. `W1` and `W2` are assumed equal by default in the *hom_general* function. On the other hand, `nsim`, `burn` and `step` are the number of simulations, the burn-in period of the chain and a parameter that indicates how often a sample does not have to be discarded. `b_pri`, `B_pri`, `r_pri` and `lambda_pri` are the parameters for the prior distributions of $\boldsymbol{\beta}$ and σ^2. Finally, `beta_0`,`sigma2_0` , `rho_0` and `lambda_0` are the initial values of each chain. For ρ and λ, the transition kernel is chosen through `kernel`. It can be either `normal` or `uniform` . The `mateq` parameter in the *hom_general* function indicates whether or not `W1`=`W2` is assumed. The `plot` parameter is a logical input that indicates whether or not the chains of each parameter are to be shown, and the logical parameter `chains` indicate if the Markov chains are to be saved and returned.

Appendix A.2. Heteroscedastic Models

The *hetero_sar*, *hetero_sem* and *hetero_general* functions can be used to fit the SAR, SEM and General heteroscedastic models. The syntax of these functions is:

- *hetero_sar*

```
hetero_sar(y,X,Z,W,nsim,burn,step,b_pri,B_pri,g_pri,G_pri,
beta_0,gamma_0,rho_0,kernel,plot)
```

- *hetero_sem*

```
hetero_sem(y,X,Z,W,nsim,burn,step,b_pri,B_pri,g_pri,G_pri,
beta_0,gamma_0,lambda_0,kernel,plot)
```

- *hetero_general*

```
hetero_general(y,X,Z,W1,W2,nsim,burn,step,b_pri,B_pri,g_pri,
G_pri,beta_0,gamma_0,rho_0,lambda_0,kernel,mateq,plot)
```

For these three models, `y`,`X`, `Z`,`W1` and `W2` represent the response variable, the explanatory variables of the mean model, the explanatory variables of the dispersion model and the spatial contiguity matrices, respectively. `W1` and `W2` are equal by default in the `hetero_general` function. On the other hand, `nsim`, `burn` and `step` are the number of simulations, the burn-in period of the chain and how often a sample does not have to be discarded. `b_pri`, `B_pri`, `g_pri` and `G_pri` are the parameters of the prior distributions of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. Finally, `beta_0`, `gamma_0`, `rho_0` and `lambda_0` are the initial values for the Markov chains. The parameter `kernel` stands for the distribution of the transition kernels of ρ y λ. It can be either `normal`(default) or `uniform`. For the *hetero_general* function, the `mateq` parameter indicates whether or not `W1`=`W2` is assumed. The `plot` parameter is a logical input that indicates if the chains of each parameter are to be shown and the logical parameter `chains` indicate if the Markov chains are to be saved and returned.

Table 7 sumarize the six `BSPADATA` R-package functions.

Table 7: *BSPADATA* R-package functions

Function	Description
hom_sar	Bayesian fit of the Bayesian homoscedastic Spatial Autorregresive (SAR) Model with normal error term
hom_sem	Bayesian fit of the Bayesian homoscedastic Spatial Error Model (SEM) with normal error term
hom_general	Bayesian fit of the Bayesian homoscedastic general model with normal error term
hetero_sar	Bayesian fit of the Bayesian heteroscedastic Spatial Autorregresive (SAR) Model with normal error term
hetero_sem	Bayesian fit of the Bayesian heteroscedastic Spatial Error Model (SEM) with normal error term
hetero_general	Bayesian fit of the Bayesian heteroscedastic general model with normal error term

Appendix B. Proofs

Appendix B.1. General Homoscedastic Regression Models

From the posterior distribution 4, the conditional posterior distributions of ρ and λ are given by:

1. Posterior conditional distribution of ρ:

$$\pi(\rho|\boldsymbol{\beta},\sigma^2,\lambda) \propto |\mathbf{A}| \exp\left\{ -\frac{1}{2\sigma^2}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{D^t}\mathbf{D}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})\right\}\mathbf{P}(\rho)$$

where where $\mathbf{A}$ and $\mathbf{D}$ are defined as in Section 1.1. Thus,

$$\begin{aligned}\pi(\rho|\boldsymbol{\beta},\sigma^2,\lambda) &\propto |\mathbf{I}-\rho\mathbf{W_1}| \exp\left\{ -\frac{1}{2\sigma^2}\left[\mathbf{y^tA^tD^tDAy}-\mathbf{y^tA^tD^tDX}\boldsymbol{\beta}-\boldsymbol{\beta}^t\mathbf{X^tD^tDAy}\right]\right\}\mathbf{P}(\rho)\\ &\propto |\mathbf{I}-\rho\mathbf{W_1}| \exp\left\{ -\frac{1}{2\sigma^2}\left[\rho^2(\mathbf{y^tW_1^tD^tDW_1y})-2\rho((\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\mathbf{D}'\mathbf{DW_1y})\right]\right\}\mathbf{P}(\rho)\\ &\propto |\mathbf{I}-\rho\mathbf{W_1}| \exp\left\{ -\frac{c}{2\sigma^2}\left(\rho-\frac{d}{c}\right)^2\right\}\mathbf{P}(\rho)\end{aligned}$$

where $c = \mathbf{y^tW_1^tD^tDW_1y}$ y $d = (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{D^tDW_1y}$.

2. Posterior conditional distribution of λ

$$\pi(\lambda|\boldsymbol{\beta},\sigma^2,\rho) \propto |\mathbf{D}| \times \exp\left\{ -\frac{1}{2\sigma^2}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{D^t}\mathbf{D}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})\right\}\mathbf{P}(\lambda),$$

Thus,

$$\pi(\lambda|\boldsymbol{\beta},\sigma^2,\rho) \propto |\mathbf{I}-\lambda\mathbf{W_2}| \exp\left\{-\frac{1}{2\sigma^2}\left[(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})^t(\mathbf{I}-\lambda\mathbf{W_2}-\lambda\mathbf{W_2^t}+\lambda^2\mathbf{W_2^tW_2})(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})\right]\right\}\mathbf{P}(\lambda)$$

Defining $\mathbf{e} = \mathbf{Ay}-\mathbf{X}\boldsymbol{\beta}$, it can be written as:

$$\pi(\lambda|\boldsymbol{\beta},\sigma^2,\rho) \propto |\mathbf{I}-\lambda\mathbf{W_2}| \exp\left\{ -\frac{1}{2\sigma^2}\left[\lambda^2(\mathbf{e^tW_2^tW_2e})-2\lambda(\mathbf{e^tW_2e})\right]\right\}\mathbf{P}(\lambda).$$

If $c = \mathbf{e^tW_2^tW_2e}$ and $d = \mathbf{e^tW_2e}$, then:

$$\begin{aligned}\pi(\lambda|\boldsymbol{\beta},\sigma^2,\rho) &\propto |\mathbf{I}-\rho\mathbf{W_2}| \exp\left\{ -\frac{c'}{2\sigma^2}\left[\lambda^2-2\lambda\frac{d'}{c'}+\frac{b^2}{c^2}-\frac{d^2}{c^2}\right]\right\}\mathbf{P}(\lambda)\\ &\propto |\mathbf{I}-\lambda\mathbf{W_2}| \exp\left\{ -\frac{c'}{2\sigma^2}\left(\lambda-\frac{d'}{c'}\right)^2\right\}\mathbf{P}(\lambda)\end{aligned}$$

where $c = (\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})'\mathbf{W_2'W_2}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})$ y $d = (\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})'\mathbf{W_2}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})$.

Appendix B.2. General Heteroscedastic Regression Models

From the posterior conditional distribution 11, the posterior conditional distributions of $\boldsymbol{\gamma}$, ρ and λ are given by:

1. Posterior conditional distribution of $\boldsymbol{\gamma}$:

$$\pi(\boldsymbol{\gamma}|\boldsymbol{\beta},\rho,\lambda) \propto |\boldsymbol{\Sigma}|^{-1/2}\exp\left\{-\frac{1}{2}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{D}^t\boldsymbol{\Sigma}^{-1}\mathbf{D}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})\right\}\exp\left\{-\frac{1}{2}(\boldsymbol{\gamma}-\mathbf{g})^t\mathbf{G}^{-1}(\boldsymbol{\gamma}-\mathbf{g})\right\}.$$

This posteriori conditional distribution does not have a known functional form. Thus we propose a Bayesian method based on working variables, in order to obtain samples of the posterior distribution of $\boldsymbol{\gamma}$, based on Cepeda-Cuervo (2001) and Cepeda-Cuervo & Gamerman (2005).

2. Posterior conditional distribution of ρ:

$$\pi(\rho|\boldsymbol{\beta},\boldsymbol{\gamma},\lambda) \propto |\mathbf{I}-\rho\mathbf{W_1}|\exp\left\{-\frac{1}{2}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{D}^t\boldsymbol{\Sigma}^{-1}\mathbf{D}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})\right\}\mathbf{P}(\rho).$$

Thus,

$$\begin{aligned}
&\pi(\rho|\boldsymbol{\beta},\sigma^2,\lambda)\\
&\propto |\mathbf{I}-\rho\mathbf{W_1}|\exp\left\{-\frac{1}{2}\left[\mathbf{y}^t\mathbf{A}^t\mathbf{D}^t\boldsymbol{\Sigma}^{-1}\mathbf{DAy}-\mathbf{y}^t\mathbf{A}^t\mathbf{D}^t\boldsymbol{\Sigma}^{-1}\mathbf{DX}\boldsymbol{\beta}-\boldsymbol{\beta}^t\mathbf{X^tD^t}\boldsymbol{\Sigma}^{-1}\mathbf{DAy}\right]\right\}\mathbf{P}(\rho)\\
&\propto |\mathbf{I}-\rho\mathbf{W_1}|\exp\left\{-\frac{1}{2}\left[\mathbf{y^t}(\mathbf{I}-\rho\mathbf{W_1})\mathbf{D^t}\boldsymbol{\Sigma}^{-1}\mathbf{D}(\mathbf{I}-\rho\mathbf{W_1})\mathbf{y}-\mathbf{y^t}(\mathbf{I}-\rho\mathbf{W_1})^t\mathbf{D^t}\boldsymbol{\Sigma}^{-1}\mathbf{DX}\boldsymbol{\beta}\right.\right.\\
&\left.\left.-\boldsymbol{\beta}^t\mathbf{X^tD^t}\boldsymbol{\Sigma}^{-1}\mathbf{D}(\mathbf{I}-\rho\mathbf{W_1})\mathbf{y}\right]\right\}\mathbf{P}(\rho)\\
&\propto |\mathbf{I}-\rho\mathbf{W_1}|\exp\left\{-\frac{1}{2}\left[-\rho\mathbf{y^tW_1^tD^t}\boldsymbol{\Sigma}^{-1}\mathbf{Dy}-\rho\mathbf{y^tD^t}\boldsymbol{\Sigma}^{-1}\mathbf{DW_1y}+\rho^2\mathbf{y^tW_1^tD^t}\boldsymbol{\Sigma}^{-1}\mathbf{DW_1y}\right.\right.\\
&\left.\left.+\rho\mathbf{y^tW_1^tD^t}\boldsymbol{\Sigma}^{-1}\mathbf{DX}\boldsymbol{\beta}+\rho\boldsymbol{\beta}^t\mathbf{X^tD^t}\boldsymbol{\Sigma}^{-1}\mathbf{DW_1y}\right]\right\}\mathbf{P}(\rho)\\
&\propto |\mathbf{I}-\rho\mathbf{W_1}|\exp\left\{-\frac{1}{2}\left[\rho^2(\mathbf{y^tW_1^tD^t}\boldsymbol{\Sigma}^{-1}\mathbf{DW_1y})-2\rho((\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{D^t}\boldsymbol{\Sigma}^{-1}\mathbf{DW_1y})\right]\right\}\mathbf{P}(\rho)\\
&\propto |\mathbf{I}-\rho\mathbf{W_1}|\exp\left\{-\frac{c}{2}\left(\rho-\frac{d}{c}\right)^2\right\}\mathbf{P}(\rho)
\end{aligned}$$

where $c=\mathbf{y^tW_1^tD^t}\boldsymbol{\Sigma}^{-1}\mathbf{DW_1y}$ y $d=(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{D^t}\boldsymbol{\Sigma}^{-1}\mathbf{DW_1y}$.

3. Posterior conditional distribution of λ.

$$\pi(\lambda|\boldsymbol{\beta},\boldsymbol{\gamma},\rho) \propto |\mathbf{I}-\lambda\mathbf{W_2}|\exp\left\{-\frac{1}{2}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})^t\mathbf{D^t}\boldsymbol{\Sigma}^{-1}\mathbf{D}(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})\right\}\mathbf{P}(\lambda).$$

Thus,

$$\begin{aligned}
\pi(\lambda|\boldsymbol{\beta},\boldsymbol{\gamma},\rho) \propto |\mathbf{I}-\lambda\mathbf{W_2}|\exp&\left\{-\frac{1}{2}\left[(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})^t\right.\right.\\
&\left.\left.\times(\boldsymbol{\Sigma}^{-1}-\lambda\boldsymbol{\Sigma}^{-1}\mathbf{W_2}-\lambda\mathbf{W_2^t}\boldsymbol{\Sigma}^{-1}+\lambda^2\mathbf{W_2^t}\boldsymbol{\Sigma}^{-1}\mathbf{W_2})(\mathbf{Ay}-\mathbf{X}\boldsymbol{\beta})\right]\right\}\mathbf{P}(\lambda)
\end{aligned} \tag{19}$$

Defining $\mathbf{e} = \mathbf{Ay} - \mathbf{X}\boldsymbol{\beta}$, we get:

$$\pi(\lambda|\boldsymbol{\beta},\boldsymbol{\gamma},\rho) \propto |\mathbf{I} - \lambda\mathbf{W_2}| \exp\left\{\frac{1}{2}\left[\lambda^2(\mathbf{e^t W_2^t \Sigma^{-1} W_2 e}) - 2\lambda(\mathbf{e^t \Sigma^{-1} W_2 e})\right]\right\}\mathbf{P}(\lambda)$$

If $c = \mathbf{e^t W_2^t \Sigma^{-1} W_2 e}$ and $d = \mathbf{e^t \Sigma^{-1} W_2 e}$, then:

$$\begin{aligned}\pi(\lambda|\boldsymbol{\beta},\boldsymbol{\gamma},\rho) &\propto |\mathbf{I} - \lambda\mathbf{W_2}| \exp\left\{-\frac{c'}{2}\left[\lambda^2 - 2\lambda\frac{d'}{c'} + \frac{b'^2}{c'^2} - \frac{d'^2}{c'^2}\right]\right\}\mathbf{P}(\lambda)\\ &\propto |\mathbf{I} - \lambda\mathbf{W_2}| \exp\left\{-\frac{c'}{2}\left(\lambda - \frac{d'}{c'}\right)^2\right\}\mathbf{P}(\lambda),\end{aligned}$$

where $c = (\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})^t \mathbf{W_2^t \Sigma^{-1} W_2}(\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})$
and $d = (\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})^t \mathbf{\Sigma^{-1} W_2}(\mathbf{Ay} - \mathbf{X}\boldsymbol{\beta})$.

With these results for the heteroscedastic general model, the formulas of the heteroscedastic and SAR models can be obtained.

[Received: December 2020 — Accepted: January 2022]

References

Anselin, L. (1980), 'Estimation methods for spatial autoregressive structures', *Regional Science Dissertation and Monograph* **8**.

Anselin, L. (1982), 'A note on small sample properties of estimators in a first-order spatial autoregressive model', *Environment and Planning A* **14**.

Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Kluwer Academic, Boston.

Anselin, L. (2001), 'Spatial econometrics', *A companion to theoretical econometrics* pp. 310–330.

Cepeda-Cuervo, E. (2001), 'Modelagem da Variabilidade em Modelos Lineales Generalizados', *(Unpublished PhD thesis. Univerisidade Federal de Rio de Janeiro,Brazil)* .

Cepeda-Cuervo, E. & Gamerman, D. (2000), 'Bayesian modeling of variance heterogeneity in normal regression models', *Brazilian Journal of Probability and Statistics* pp. 207–221.

Cepeda-Cuervo, E. & Gamerman, D. (2005), 'Bayesian methodology for modeling parameters in the two parameters exponential family', *Revista Estadística* **57**, 93–105.

Gangnon, R. & Clayton, M. (1998), Bayesian Spatial Disease Clustering: An Application, Technical report, Department of Biostatistics, University of Wisconsin, Madison.

Hepple, L. W. (1979), Bayesian analysis of the linear model with spatial dependence, *in* 'Exploratory and explanatory statistical analysis of spatial data', Springer, pp. 179–199.

LeSage, J. (1997), 'Bayesian Estimation of Spatial Autoregressive Models', *Journal of Composite Materials* **16**(4), 928–940.

LeSage, J. & Pace, R. K. (2009), *Introduction to Spatial Econometrics*, Chapman & Hall/CRC, Boca Raton.

Neath, A. A. & Cavanaugh, J. E. (2012), 'The bayesian information criterion: background, derivation, and applications', *Wiley Interdisciplinary Reviews: Computational Statistics* **4**.

Waller, L. (1996), 'Statistical power and design of focused clustering studies', *Statistics in Medicine* (15), 765–782.

Waller, L. A. & Gotway, C. A. (2004), *Applied Spatial Statistics for Public Health Data*, Wiley series in probability and statistics, 1 edn, John Wiley & Sons.