214

DISCUSSION PAPER

SPATIAL MULTIVARIATE REGRESSIONS WITH PANEL DATA

Mário Jorge Mendonça
Luis Alberto Medrano

214
DISCUSSION PAPER
Brasília, September 2016

SPATIAL MULTIVARIATE REGRESSIONS WITH PANEL DATA

Mário Jorge Mendonça[1]
Luis Alberto Medrano[2]

1. Institute for Applied Economic Research <mario.mendonca@ipea.gov.br>.
2. Institute for Applied Economic Research and Federal Rural University of Rio de Janeiro <lmedrado@ufrrj.br>.

Federal Government of Brazil

Ministry of Planning, Development and Management
Minister, Interim Dyogo Henrique de Oliveira

ipea **Institute for Applied Economic Research**

A public foundation affiliated to the Ministry of Planning, Development and Management, Ipea provides technical and institutional support to government actions – enabling the formulation of numerous public policies and programs for Brazilian development – and makes research and studies conducted by its staff available to society.

President
Ernesto Lozardo

Director of Institutional Development
Juliano Cardoso Eleutério

Director of Studies and Policies of the State, Institutions and Democracy
João Alberto De Negri

Director of Macroeconomic Studies and Policies
Claudio Hamilton Matos dos Santos

Director of Regional, Urban and Environmental Studies and Policies
Alexandre Xavier Ywata de Carvalho

Director of Sectoral Studies and Policies, Innovation, Regulation and Infrastructure
Fernanda De Negri

Director of Social Studies and Policies
Lenita Maria Turchi

Director of International Studies, Political and Economic Relations
Alice Pessoa de Abreu

Chief of Staff, Deputy
Márcio Simão

Chief Press and Communications Officer
Regina Alvarez

URL: http://www.ipea.gov.br
Ombudsman: http://www.ipea.gov.br/ouvidoria

DISCUSSION PAPER

A publication to disseminate the findings of research directly or indirectly conducted by the Institute for Applied Economic Research (Ipea). Due to their relevance, they provide information to specialists and encourage contributions.

© Institute for Applied Economic Research – **ipea** 2016

Discussion paper / Institute for Applied Economic Research.- Brasília : Rio de Janeiro : Ipea, 1990-

ISSN 1415-4765

1. Brazil. 2. Economic Aspects. 3. Social Aspects. I. Institute for Applied Economic Research.

CDD 330.908

The opinions expressed in this publication are of exclusive responsibility of the authors, not necessarily expressing the official views of the Institute for Applied Economic Research and the Ministry of Planning, Development and Management.

Reproduction of this text and the data contained within is allowed as long as the source is cited. Reproduction for commercial purposes is prohibited.

Jel Codes: C31, C39.

CONTENTS

ABSTRACT

We develop a new Bayesian estimator that is able to deal with multivariate panel data structure in the presence of spatial correlation. The analysis of panel data introduced here allows us to analyze not only the fixed effect but also the random effect model. This work extends the previous study undertaken by Gamerman and Moreira (2004) which only spatial scale is considered. To estimate the random effect model we use the hierarchical analysis that can be applied to estimate some categories of longitudinal data models. The Monte Carlo simulations demonstrate the ability of this new estimator to replicate quite well simulated data. To show the empirical relevance of this new estimator we apply it to the deforestation data in the Brazilian Amazon.

Keywords: multivariate regressions; spatial correlation; panel data; fixed effect; Markov chain Monte Carlo.

RESUMO

Neste estudo desenvolvemos um estimador Bayesiano capaz de lidar simultaneamente com uma estrutura de regressão multivariada de dados em painel e com correlação espacial. A análise dos dados em painel contempla os casos de *pooling*, efeito fixo e efeito aleatório. Para estimação do modelo com efeito aleatório, usamos a análise hierárquica. As simulações de Monte Carlo via cadeia de Markov demonstraram a capacidade do estimador para replicar os dados muito bem simulados. Usamos ainda dados de desmatamento da Amazônia brasileira para atestar a relevância empírica desse novo estimador.

Palavras-chaves: regressão multivariada; correlação espacial; dados em painel; efeito fixo; Monte Carlo via cadeia de Markov.

1 INTRODUCTION

In this paper we develop a new estimator that is able to deal with multivariate panel data structure in the presence of spatial correlation. We departed from a Bayesian inference practice, which brings important advantages in terms of the ability to obtain "more confident estimates in the presence of small samples with a high-dimensional space of parameters", (Gelman et al., 2003, p. 696). This work extends the previous study undertaken by Gamerman and Moreira (2004) where only spatial scale is considered. Besides that we join cross-section and time series in the same framework based on the panel data structure (Hsiao, 1995; Baltagi, 1995; Arellano, 2003). The analysis of longitudinal data introduced here allows us to analyze not only the fixed[3] effect but also the random effect. The univariate model for panel data was extensively treated in the Bayesian paradigm. Laird and Ware (1982); and Chib and Carling (1999), are the primary references. Good expositions can be found in Koop (2003); Lancaster (2004) and Greenberg (2007).

The introduction of a panel data structure brings some new difficulties, mainly in the random effect model (Laird and Ware, 1982). The specification we follow to model the random effect is denoted by classical econometrics as the error components model (Baltagi, 1995). The method which we apply to estimate the random effect model can be understood as an extension of the one that appear in Greenberg (2007), for univariate regressions. In this case, due to a large number of parameters the simpliest form of Markov chain Monte Carlo (MCMC) method would slow to converge. Hence, we use an algorithm based on hierarchical analysis that can be applied to estimate some categories of longitudinal data models. This methodology frames parameters in groups minimizing time computing.

In a frequentist perspective multivariate regression model was studied by many authors. Baltagi (1981; 1995), Baltagi and Li (1992), and Kinal and Lahiri (1990; 1993) shows that a consistent estimate can be obtained using 3SLS while Cornwell et al. (1992) employ GMM approach. Notwithstanding these models do not take into consideration spatial autocorrelation dependence. An interesting case that can be analysed is the vector autoregressive with panel data (PVAR) as it will be shown in this paper.

Finally, we apply our estimator to deal with the problem associated with deforestation in the Brazilian Amazon. The common wisdom will link cattle ranching to the increase of deforestation in the Brazilian Amazon. We show that since a more complex structure is taken into account cattle ranching is not a important driver of deforestation.

The article is organized as follows. The section 2 presents distint models of multivariate regression with panel data while describe how to frame them into more compact matrix form that will be usefull to undertake our estimation procedure. In section 3 we offers methods to estimate these models based on Bayesian inference. We use artificial data generated by Monte Carlo simulation to validate our methodology. In section 4 we use an empirical perspective to estimate the model of dynamics of the land-use in Brazilian Amazon. Finally, some concluding remarks are presented in section 5.

2 MULTIVARIATE REGRESSIONS WITH PANEL DATA

2.1 General Model (Pooling)

The structural vector regression model with panel data with Q endogenous variables can represented in the following way:

$$y_{it}^{'}A = \alpha_i^{'} + x_{it1}^{'}\Gamma_1 + \ldots + x_{itK}^{'}\Gamma_K + \varepsilon_{it}^{'},\ i = 1,\ldots,N \text{ and } t = 1,\ldots,T. \qquad (1)$$

The indexes i and t are associated, respectively, to each spatial unit and each time period. Let $y_{it} = (y_{it1},\ldots,y_{itQ})^{'}$ is the vector $Q\times 1$ of endogenous variables and $\varepsilon_{it} = (\varepsilon_{it1},\ldots,\varepsilon_{itQ})^{'}$ is the $Q\times 1$ vector of random terms with $\varepsilon_{it} \sim (0,\Lambda_\varepsilon)$ where Λ_ε is the $Q\times Q$ diagonal matrix. The $Q\times 1$ vector of individual effects $\alpha_i = (\alpha_{i1},\alpha_{i2},\ldots,\alpha_{iQ})^{'}$ that not varies in time dimension is introduced to model the spatial heterogeneity in the data. Let $x_{itk} = (x_{itk1},\ldots,x_{itkQ})^{'}$ for $k = 1,\ldots,K$ is the vector of exogenous variables. Each K is associated to the effect of the *k-th* exogenous variable on the system. The contemporaneous relationships among the endogenous variables are represented by the $Q\times Q$ matrix A, and Γ_K is a $Q\times K$ matrix that contains the coefficients related to the effect of x_{itk} on y_{it}. The vector autoregressive with panel data (PVAR) is a special case of (1) in which $x_{it} = y_{it-k}$. The vector of individual effect α_i can be treated as fixed or random component. In the last case, $\alpha_i \sim (0,\Lambda_\alpha)$ with Λ_ε a $Q\times Q$ diagonal matrix. By simplicity from now on we assume that there is just one exogenous variable. In this case (1) can be rewritten such as

$$y_{it}^{'}A = \alpha_i^{'} + x_{it}^{'}\Gamma + \varepsilon_{it}^{'}, \qquad (1')$$

where $x_{it} = (x_{it1},\ldots,x_{itK})^{'}$ is the $K\times 1$ vector and Γ is a matrix the $Q\times K$.

In order to accommodate the sample data and introducing notations we firstly consider the pooling model in which the individual component is not included. The treatment of the specific individual component will be undertaken in the next sections. We need some additional notations to allocate the sample data in a more compact way. The information concern each unit i can be put in a matrix Y_i such as

$$Y_i = \begin{bmatrix} (y_{i11},\ldots,y_{i1Q}) \\ \vdots \\ (y_{iT1},\ldots,y_{iTQ}) \end{bmatrix} = \begin{bmatrix} y'_{i1} \\ \vdots \\ y'_{iT} \end{bmatrix}, \text{ where } Y_i \text{ is matrix } T \times Q$$

where T is the size of time dimension. Based on it the information related all the units can be framed defining a $TN \times Q$ matrix Y in the following way.

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix}.$$

The information related to the exogenous variable and the disturbance can be also represented respectively by the matrices X and E elaborated by the same way as Y. Thus the pooling model can be framed in a compact form

$$YA = X\Gamma + E, \tag{2}$$

where Y and E are $NT \times Q$ matrices, and X is $NT \times K$. We assume that E is mactrice Normal such as $E \sim MN(0, \Lambda_\varepsilon \otimes I_{TN})$. Provided that A is invertible, then (2) has a reduced form given by

$$Y = XB^* + U. \tag{3}$$

The relationship between structural form and reduced form[1] is based on the following identities

1. More details about structural and reduced form can be found in Hamilton (2003).

$$B^* = \Gamma A^{-1},\ U = EA^{-1} \text{ and } \Sigma_u = A^{-1}\Lambda_\varepsilon (A^{-1})', \tag{4}$$

Σ_u is also $Q \times Q$ and B^* is the $Q \times K$ matrices. Note that $U \sim MN(0, \Sigma_\varepsilon \otimes I_{TN})$, that is, the reduced form residuals can be interpreted as the result of linear combinations of exogenous shocks that are not contemporaneously (in the same instant of time) correlated. It means that the reduced form representation does not allow to identify the effects of exogenous independent shocks onto the variables (the matrix Σ is not diagonal)[2].

2.2 Fixed Effect (FE) Model

The model with fixed effect can be defined expanding the pooling model. To take into account the fixed effect we need to define a $N \times Q$ matrix of individual effects $\alpha = [\alpha_{ij}]_{NxQ}$ where α_{ij} is the individual effect of unit i in the equation j. Then we can consider it as more one set of parameters. The inclusion of α in the model can be done in such way

$$Y = XB^* + (I_N \otimes i_T)\alpha + U, \tag{5}$$

where i_T is a $T \times 1$ vector with all the elements equal to one. The set of regression in (5) can be accommodate in a more compact framework based in the following procedures

$$Y = \begin{bmatrix} I_N \otimes i_T & X \end{bmatrix} \begin{bmatrix} \alpha \\ B^* \end{bmatrix} + U \text{ or}$$

$$Y = ZB + U, \tag{5'}$$

where $\begin{bmatrix} I_N \otimes i_T & X \end{bmatrix} = Z$ and Z is $TN \times (N+K)$, and $\begin{bmatrix} \alpha \\ B^* \end{bmatrix} = B$ where B is $(N+K) \times Q$.

Because $B = \begin{bmatrix} \alpha \\ B^* \end{bmatrix}$, it means that for a sample data with larger number of units the estimation of matrix α can be time consuming. In the classical grounds the problem can be solved using a convenient orthogonal projection operators on (5) defined as $P = I_N \otimes J_T$ and $Q = I_{NT} - P$ where $J_T = (1/T) i_T i'_T$. In this case we have $QY = QXB^* + QU$ because $Q(I_N \otimes J_T)\alpha = Q\alpha^* = 0$. With data group by unities P transforms a vector of

2. These shocks are primitive and exogenous forces, with no common causes, that affect the variables of the model (Sims, 1980).

observations in a vector of group means[3], i.e. Q produces a vector of deviations from means. In this case the OLS estimator for B

$$\hat{B}^* = \left[X^{*'}X^*\right]^{-1}\left[X^{*'}Y^*\right] = \left[X^{'}QX\right]^{-1}\left[X^{'}QY\right],$$

where $X^* = QX$ and $Y^* = QY$. Since we have estimated B^* we can retrieve the individual effect easily. Unfortunately, due to Q be idempotent, the variance of QU is singular, which limits the practicability to apply the Bayesian methods because we need the inverse of this matrix as we can show in section 3.

2.3 Random Effect (RE) Model

The random effect model that we analyse in this study refers to the one known in the literature on panel data as the error component model. In this case the error regression is assumed to be composed of two independent components: one component associated with cross-sectional units or individuals and a second with each observation. In this case (5) must be interpreted in the following way

$$Y = XB^* + V, \tag{6}$$

where V the matrix of disturbance defined as $V = (I_N \otimes i_T)\alpha + U$. Here $\alpha \sim MN_{N\times Q}(0_{N\times Q}, \Sigma_\alpha \otimes I_N)$ and $U \sim MN_{TN\times Q}(0_{TN\times Q}, \Sigma_u \otimes I_{TN})$. With minor differences this moldel can be considered as the multivariate version of the univariated Gaussian linear mixed model (Laird and Ware, 1982). To find the variance of V it is necessary first to find the variance of $(I_N \otimes i_T)\alpha$. In this sense we have that

$$vec = (I_N \otimes i_T)\alpha = vec((I_N \otimes i_T)\alpha I_Q) = \left(I_Q \otimes \left(I_N \otimes i_T\right)\right)vec(\alpha).$$

Using the following property of Mactrice Normal distribution, that is, if $W \sim MN_{M\times P}\left(\overline{W}, \Sigma \otimes \Omega\right)$ then $vec(W) \sim N_{MP\times 1}\left(vec\left(\overline{W}\right), \Sigma \otimes \Omega\right)$, we have that

$$\begin{aligned}
\Sigma_v &= E\left(vv^{'}\right) = \left(I_Q \otimes \left(I_N \otimes i_T\right)\right)E\left(vec(\alpha)vec(\alpha)^{'}\right)\left(I_Q \otimes \left(I_N \otimes i_T^{'}\right)\right) = \\
&= \left(I_Q \otimes \left(I_N \otimes i_T\right)\right)\left(\Sigma_\alpha \otimes I_N\right)\left(I_Q \otimes \left(I_N \otimes i_T^{'}\right)\right) == \left(I_Q \otimes \left(I_N \otimes i_T\right)\right)\left(\Sigma_\alpha \otimes \left(I_N \otimes i_T^{'}\right)\right) \\
&= \left(\Sigma_\alpha \otimes \left(I_N \otimes i_T i_T^{'}\right)\right) = T\left(\Sigma_\alpha \otimes P\right)
\end{aligned}$$

3. $PY_{it} = Y_{it} - \frac{1}{T}\sum Y_{it} = Y_{it} - \overline{Y}_i.$

Thus $\Sigma_V = E\left(VV'\right) = T(\Sigma_\alpha \otimes P) + \Sigma_U \otimes I_{TN}$ as it appears in Baltagi (1992).

2.4 Spatial FE Model

In this study the spatial dependence or autocorrelation is introduced directly from the data follow the methodology first proposed by Anselin (1988). Other methods could be used. The spatial dependence can be also model using latent component (Gamerman and Moreira, 2004) that incorporate spatial variation of the regression coefficients. The roots of this method appear in Fahmeir and Lang (2001) and Lang et al. (2003).

In order to assess spatial dependence some new elements must be introduced in the analysis.A standard choice is to include in the mean process a spatial autoregressive component that takes the spatial locations into account. This can be done by the use of a contiguity or neighborhood $N \times N$ matrix $W^* = (w_{ij})$ with w_{ij} representing the neighborhood between the sites i and j, such that $w_{ij} \neq 0$ if the sites i and j are neighbors, and $w_{ij} = 0$ otherwise. The standard choice is $w_{ij} = 1/m$ where m_j is the number of neighborhood of the unit j. This spatial weight matrix is a square matrix representing the spatial context. It encodes the neighborhood relationship among the spatial units. The literature on this subject is vast, we shall only retain here that the neighbors relationships chosen by the analyst may change the results[4]. In this sense equation (5') may be rewritten in the following way:

$$Y = WY\Phi + ZB + U, \tag{7}$$

where $W = I_T \otimes W^*$ and W is $NT \times NT$ and Φ is a $Q \times Q$ matrix. The form to accommodate spatial autocorrelation in the data specified in (8) is denoted in spatial econometrics by SAR.[5] The special case is given when Φ is diagonal with entries Φ_1 $\phi_1, \cdots, \phi_N$ consisting of the spatial autoregressive coefficients. When Φ is full the non-zero elements off the diagonal display the effect that one endogenous variable has on the other ones, irrespective of the spatial interactions between regions.

It is commonplace in the literature to consider spatial autocorrelation also present in the disturbance. In this case $U = WU\Psi + V$ which is denoted as the SEM[6] model.

4. See Anselin et al. (1988) for a detailed discussion.
5. Spatial autoregressive model.
6. Spatial error model.

In this study for simplicity we assume that $\Psi = 0$. This simplification does not bring any weakness in our methodology. The methodology that we will apply to estimate the model with spatial autocorrelation can be extended easily to consider both specifications. We decided to include only SAR with the aim to reduce the time of computing.

3 BAYESIAN INFERENCE

3.1 The FE model

To estimate the models treated in this study we consider the Bayesian grounds for inference. The Bayes' theorem gives the posterior distribution as

$$P(B,\Sigma_u \mid Y) \propto P(Y \mid B,\Sigma_u)P(B,\Sigma_u),$$

where $P(Y \mid B,\Sigma_u)$ is the likelihood function and $P(B,\Sigma_u)$ is the prior distribution. We assume that likelyhood function is given by a matricvariate normal (Press, 1989) defined as follows

$$P(Y \mid B,\Sigma_u) \propto \mid \Sigma_u \mid^{-Q/2} \exp\left[-\frac{1}{2}tr\left(\Sigma_u^{-1}\left(Y-ZB\right)'\left(Y-ZB\right)\right)\right].$$

To conduct Bayesian inference a prior distribution for (B,Σ_u) is required. We assume the matricvariate normal inverse Wishart (Press, 1989) form for prior distribution, such as, $(B,\Sigma_u) \sim MNIW(B_0,H_0,\nu_0,S_0)$ with density $P(B,\Sigma_u) \propto P(B \mid \Sigma_u)P(\Sigma_u)$ where

$$P(B \mid \Sigma_u) \infty \mid H_0 \mid^{-Q/2} \mid \Sigma_u \mid^{-Q/2} \exp\left[-\frac{1}{2}tr\left(\Sigma_u^{-1}\left(B-B_0\right)' H_0^{-1}\left(B-B_0\right)\right)\right]$$

$$P(\Sigma_u) \propto \mid \nu_0 S_0 \mid^{-Q/2} \mid \Sigma_u \mid^{-(\nu_0+Q+1/2)} \exp\left[-\frac{1}{2}tr\left(\Sigma_u^{-1}\nu_0 S_0\right)\right],$$

where H_0 and S_0 are positive definite matrices and $\nu_0 > 0$ is a real number. In others words conditional to Σ_u, $B \mid \Sigma_u \sim N(B_0,\Sigma_u \otimes H_0^{-1})$, while the marginal distribution for Σ_u^{-1} is an inverse Wishart denoted by $\Sigma \sim IW(\nu_0,S_0)$ with $E[\Sigma] = S_0 / (\nu_0 - (Q+1))$.

Due to the kernel of the likelihood function is also given by a MN distribution it is possible to conjugate. In this case the posterior distribution of (B,Σ_u) (product of likelihood function and prior distribution) is also a *MNIW* such that $B \mid \Sigma_u, Y \sim N(B_T, \Sigma_u \otimes H_T^{-1})$ while $\Sigma_u \mid Y \sim IW(S_T, v_T)$, with

$$H_T = H_0 + Z'Z; \tag{8.1}$$

$$v_T = v_0 + T; \tag{8.2}$$

$$S_T = \frac{v_0}{v_T} S_0 + \frac{T}{v_T}\hat{\Sigma} + \frac{1}{v_T}(\hat{B} - B_0)' H_0 H_T Z' Z(\hat{B} - B_0); \tag{8.3}$$

$$B_T = H_T^{-1}(H_0 B_0 + Z' Z\hat{B}). \tag{8.4}$$

All these conditional distributions are standard therefore the Gibbs sampling algorithm (Gamerman and Lopes, 2005), can be applied. Given the draw of $g - 1$th iteration, the next iteration is done by simulation

Step1 $\quad \Sigma_u^{(g)} \sim IW(S_T, v_T)$.

Step 2 $\quad B^{(g)} \mid \Sigma_u^{(g-1)} \sim N(B_T^{(g-1)}, \Sigma^{(g-1)} \otimes H_T^{-1})$.

This process can be iterated a larger number of times, and after discarding the initial transient all the subsequent draws can be used for inference. Information on convergence can be obtained by monitoring the serial correlation in the draws, the numerical standard errors of estimates based on the output of the sample and through the diagnostic of Gelman and Rubin (1992).

3.2 The RE model

The Bayes' theorem gives the posterior distribution as following

$$P(B, \Sigma_u, \alpha, \Sigma_\alpha \mid Y) \propto P(Y \mid B, \Sigma_u, \alpha, \Sigma_\alpha) P(B, \Sigma_u, \alpha, \Sigma_\alpha).$$

Equation (6) provides the likelihood for RE model as

$$P(Y \mid B, \Sigma_u, \alpha, \Sigma_\alpha) \propto |\Sigma_u|^{-Q/2} \exp\left[-\frac{1}{2} tr\left(\Sigma_u^{-1}\left(Y - ZB^* - \alpha\right)'\left(Y - ZB^* - \alpha\right)\right)\right].$$

In order to compose the posterior distribution we need to complement the likelihood with the prior distribution

$$P(B,\Sigma_u,\alpha,\Sigma_\alpha)=P(B,\Sigma_u)P(\alpha)P(\Sigma_\varepsilon)=P(B|\Sigma_u)P(\Sigma_u)P(\alpha)P(\Sigma_\alpha).$$

One reason the Bayesian methods has gained popularity is associated to the fact that hierarchical priors can surmount some of the problems caused by high dimensional parameters spaces. The random individual effect model has a parameter space. If T is small relative to N, the number of parameters is quite large relative to sample size. This suggests that a hierarchical prior might be appropriate and is indeed the case that such priors are commonly used. A convenient hierarchical prior proposed to dealing with the term associated to random effect α assumes that

$$\alpha \sim MN(0,\Sigma_\alpha \otimes I_N).$$

The hierarchical structure of the prior arises if Σ_α have its own prior. We define the prior distribution for Σ_α through

$$\Sigma_\alpha \sim IW(d_0, D_0).$$

For the remaining parameters, we assume a non-hierarchical prior of the independent Normal-Wishart variety

$$\Sigma_u \sim IW(v_0, S_0) \text{ and } B \sim MN(B_0, H_0).$$

One can apply Gibbs sampling to each block parameters, increasing the sampler with draws for α. Due to the conditional Gaussian structure, $Y|B,\Sigma_u,\Sigma_\alpha \sim MN(XB,\Sigma_V)$ where $\Sigma_V = E(VV') = T(\Sigma_\alpha \otimes P)+\Sigma_U \otimes I_{TN}$. This implies that the posterior distribution of B conditional to Σ_α and Σ_u is $B|\Sigma_\alpha,\Sigma_u,Y \sim MN(B_1,H_1)$ where

$$H_1 = \left[H_0^{-1} + X'\Sigma_V^{-1}X\right]^{-1} \tag{9.1}$$

$$\text{and } B = H_1\left[H_0^{-1}B_0 + X'\Sigma_V^{-1}Y\right]. \tag{9.2}$$

In order to find the conditional posterior distribution for random effect α we first apply the vec operator on (6) such as $y=(I\otimes X)b+(I_Q\otimes J)\alpha^*+u$, where $y = vec(Y)$,

$b = vec(B), u = vec(U)$ and $J = I_N \otimes i_T$. Letting $\tilde{y} = y - (I \otimes X)b$ this posterior distribution is given such as $\alpha^* \mid B, \Sigma_\alpha, \Sigma_u, Y \sim N\left(\alpha_1^*, R\right)$ with

$$R = \left[\left(I_Q \otimes J'\right)\left(\Sigma_u^{-1} \otimes I_{NT}\right)\left(I_Q \otimes J\right) + \left(\Sigma_\alpha^{-1} \otimes I_N\right)\right]^{-1} = \left[\left(\Sigma_u^{-1} \otimes J'J\right) + \Sigma_\alpha^{-1} \otimes I_N\right]^{-1} \quad (9.3)$$

$$\alpha_1^* = R\left[\left(\Sigma_u^{-1} \otimes J'\right)\tilde{y}\right]$$

and $\alpha_1^* = R\left[\left(\Sigma_u^{-1} \otimes J'\right)\tilde{y}\right]$. (9.4)

The posterior precision matrix for random effect Σ_α is simulated from an invert Wishart, $\Sigma_\alpha \mid \alpha, B, \Sigma_u, Y \sim IW\left(d_1, D_1\right)$ with

$$D_1 = \left[D_0 + \alpha'\alpha\right] \quad (9.5)$$

and $d_1 = d_0 + N$. (9.6)

Finally the posterior distribution for precision error Σ_u is also given by an invert Wishart $\Sigma_u \mid B, \alpha, \Sigma_\alpha, Y \sim IW\left(v_1, S_1\right)$ with

$$v_1 S_0 = v_0 S_0 + \left(Y - XB - \left(I_N \otimes i_T\right)\alpha\right)'\left(Y - XB - \left(I_N \otimes i_T\right)\alpha\right) \quad (9.7)$$

and $v_1 = v_0 + NT$. (9.8)

Thus the sampler takes the following form

1. Simulate $B^{(g)}$ and $\alpha^{(g)}$ from $f\left(B^{(g)}, \alpha^{(g)} \mid \Sigma_u^{(g-1)}, \Sigma_\alpha^{(g-1)}, Y\right)$ by sampling

(i) $B^{(g)} \mid \Sigma_u^{(g-1)}, \Sigma_\alpha^{(g-1)}, Y$;

(ii) $\alpha^{(g)} \mid B^{(g)}, \Sigma_u^{(g-1)}, \Sigma_\alpha^{(g-1)}, Y$

2. Simulate $\Sigma_\alpha^{(g)}$ from $f\left(\Sigma_\alpha^{(g)} \mid B^{(g)}, \alpha^{(g)}, \Sigma_u^{(g-1)}, Y\right)$;

3. Simulate $\Sigma_u^{(g)}$ from $f\left(\Sigma_u^{(g)} \mid B^{(g)}, \alpha^{(g)}, \Sigma_\alpha^{(g)}, Y\right)$; and

4. Repeat steps 1-3 for a large number of interactions, after to burn the initial

iterations, to sumarize the posterior density.

3.3 Spatial FE model

In the first step the likelihood function is required. Based on eq. (8) this function can be written as follow

$$P(Y|B,\Sigma_u,\Phi,Z) \propto |\Sigma_u|^{-Q/2} \exp\left[-\frac{1}{2} tr\left(\Sigma_u^{-1}\left(Y-W\Phi Y-ZB\right)'\left(Y-W\Phi Y-ZB\right)\right)\right].$$

This likelihood function can be combined with a prior for (B,Σ_u,Φ), $P(B,\Sigma_u,\Phi)=P(B,\Sigma_u)P(\Phi)=P(B|\Sigma_u)P(\Sigma_u)P(\Phi)$. Here the prior distribution for (B,Σ_u) is the same used in FE model, $(B,\Sigma_u) \sim MNIW(B_0,H_0,\nu_0,S_0)$. When Φ is a diagonal matrix is reasonable to assume the elements ψ_i of the diagonal are i.i.d. with $\psi_i \sim U\,(0,1)$ for all i. Hence,

$$p(\Phi)=\prod_{i=1}^{Q} f(\psi_i). \tag{10}$$

The conditional posterior for (B,Σ_u) is known and it is given by $(B,\Sigma_u \mid \Phi,Y) \sim MNIW(B_1,H_1,\nu_1,S_1)$ with

$$H_1=(Z'Z+H_0^{-1}); \tag{11.1}$$

$$B_1=H_1^{-1}(H_0B_0+Z'Y^*); \tag{11.2}$$

$$\nu_T=\nu_0+T; \tag{11.3}$$

$$\text{and } S_T=\frac{\nu_0}{\nu_T}S_0+\frac{T}{\nu_T}\hat{\Sigma}+\frac{1}{\nu_T}(\hat{B}-B_0)'H_0H_TZ'Z(\hat{B}-B_0). \tag{11.4}$$

Because the posterior conditional distribution for Φ is not analitically tractable, we shall employ \`Metropolis-Hastings sampling' algorithm (Dagpunar, 2007) to generate drawings from it. Thus, the MCMC sampler combines Gibbs and M-H algorithms and must be employed in two stages. In the first stage the posterior distribution for (B,Σ) conditional on Φ is draw by Gibbs sampling. For each g-iteration can be done by simulation in the following way.

Step 1. Draw Φ^g from $P(\Phi \mid B^{(g)}, \Sigma^{(g)}, Y)$ using MH algorithm;

Step 2. Draw $(B^{(g)}, \Sigma^{(g)})$ from $P(B^{(g)}, \Sigma^{(g)} \mid \Phi^{(g)}, Y)$

where $P(B, \Sigma \mid \Phi, Y) \sim MNIW(B_T, H_T, \nu_T, S_T)$.

In the second stage, the M-H algorithm works as follows. Given $\widetilde{\Phi}^{g-1} = \Phi^{g-1}$ the state at the next step is $\widetilde{\Phi}^{g} = \Phi^{g}$ with probability $\alpha(\Phi^{g-1}, \Phi^{g})$, or remains unchanged at Φ^{g-1} with probability $1 - \alpha(\Phi^{g-1}, \Phi^{g})$. Here the candidate the Φ^g is sampling from a proposal density $q(\Phi^g ! \Phi^{g-1})$. The acceptance probability is

$$\alpha(\Phi^{g-1}, \Phi^{g}) = \min\left\{1, \frac{P(\Phi^{g})q(\Phi^{g-1} \mid \Phi^{g})}{P(\Phi^{g-1})q(\Phi^{g} \mid \Phi^{g-1})}\right\}, \tag{12}$$

where $P(\Phi) = P(Y \mid B, \Sigma_u, \Phi, Z)p(\Phi)$. Set $u \sim U(0,1)$, if $(u < \alpha)$ then $\widetilde{\Phi}^{g} = \Phi^{g}$, otherwise $\widetilde{\Phi}^{g} = \Phi^{g-1}$. To implement M-H algorithm it is necessary that a suitable candidate for proposal density be specified. A suitable choice for q might be

$$q(\Phi^{g} \mid \Phi^{g-1}) \propto \exp\left[-\left(\Phi^{g} - \Phi^{g-1}\right)' \Sigma^{-1} \left(\Phi^{g} - \Phi^{g-1}\right)\right],$$

that is, given Φ^{g-1}, $\Phi^{g} \sim N\left(\Phi^{g-1}, \Sigma\right)$. Other specifications can be select from a family of distribution. There is a vast discussion in the literature about how this choice could be made (Dagpunar, 2007; Robert and Casella, 2004; Chib and Greenberg, 1995).

3.4 Spatial RE model

The strategy to implement Bayesian inference in a multivatriate model with random effects is much more complicate. We can write the special RE model as:

$$Y = WY\Phi + XB + V, \tag{13}$$

where V the matrix of disturbance defined as $V = (I_N \otimes i_T)\alpha + U$, $\alpha \sim MN_{N \times Q}(0_{N \times Q}, \Sigma_\alpha \otimes I_N)$ and $U \sim MN_{TN \times Q}(0_{TN \times Q}, \Sigma_u \otimes I_{TN})$.

Following the Bayes theorem we have:

$$P(B,\Sigma_u,\alpha,\Sigma_\alpha,\Phi|Y) \propto P(Y|B,\Sigma_u,\alpha,\Sigma_\alpha,\Phi)P(B,\Sigma_u,\alpha,\Sigma_\alpha,\Phi). \tag{14}$$

The likelihood function of equation (14) is given by

$$P(Y|B,\Sigma_u,\alpha,\Sigma_\alpha,\Phi) \propto$$

$$\propto |\Sigma_u|^{-Q/2} \exp\left[-\frac{1}{2} tr\left(\Sigma_u^{-1}\left(Y - W\Phi Y - ZB - (I_N \otimes i_T)\alpha\right)'\left(Y - W\Phi Y - ZB - (I_N \otimes i_T)\alpha\right)\right)\right].$$

To build the posteriori function we should combine de likelihood function above with the a priori function for the parameters $(B,\Phi,\Sigma_u,\Sigma_\alpha,\Phi)$, that is,

$$P(B,\Sigma_u,\alpha,\Sigma_\alpha,\Phi) = P(B)P(\Sigma_u)P(\alpha)P(\Sigma_\alpha)P(\Phi).$$

In relation to $(B,\alpha,\Sigma_u,\Sigma_\alpha)$ we can use the same set of priori distributions that we use before in section 3.2 where $\alpha \sim MN(0,\Sigma_\alpha \otimes I_N)$, $\Sigma_\alpha \sim IW(d_0,D_0)$, $\Sigma_u \sim IW(v_0,S_0)$ and $B \sim MN(B_0,H_0)$, for Φ we use the specification described in (10).

Since we have defined the a priori distribution, and the posteriori distribution we have that the Gibbs sample assume the following format $B|\Sigma_\alpha,\Sigma_u,\alpha,\Phi,Y \sim MN(B_1,H_1)$, $\alpha|B,\Sigma_\alpha,\Sigma_u,\Phi \sim N(\alpha_1^*,R)$, $\Sigma_\alpha|\alpha,B,\Sigma_u,\Phi \sim IW(d_1,D_1)$ and $\Sigma_u|B,\alpha,\Sigma_\alpha,\Phi \sim IW(v_1,S_1)$ with the hyperparameters stated in equations 9.1 – 9.8. For the estimation of the parameter Φ, we follow the M-H algorithm.

3.5 Simulation Exercise

To evaluate the possibilities for conducting inference, it is necessary to first consider the performance of our Bayesian methodology with artificial data. For the sake of brevity, we present results here for only one simulated data set. We generate an artificial data with $N = 49$ individual units, $T = 10$ time periods, $k = 2$ regressors and $Q = 4$ endogenous variables (equations). The explanatory variables in the 490×2 ($NT \times k$) matrix were generated from *N(0,1)* distribution. The spacial dependence matrix W was obtained in accordance to section 2.4. Finally, the set of parameters $(B,\Phi,\Sigma_u,\Sigma_\alpha)$ is taken to be equal to

$$B' = \left[\beta_{ij}\right] = \begin{pmatrix} -0.5 & 0.1 & -0.3 & -0.8 \\ 0.2 & 0.8 & 0.6 & 0.3 \end{pmatrix}, \ \Sigma_u = \left[\sigma_{ij}\right] = \begin{pmatrix} 0.05 & -0.01 & -0.01 & -0.01 \\ -0.01 & 0.05 & 0.02 & 0.02 \\ -0.01 & 0.02 & 0.05 & -0.03 \\ -0.01 & 0.02 & -0.03 & 0.05 \end{pmatrix}$$

$$\Phi = \left[\psi_{ij}\right] = \begin{pmatrix} 0.3 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & -0.5 & 0 \\ 0 & 0 & 0 & 0.6 \end{pmatrix} \text{ and } \Sigma_\alpha = \left[\sigma_{\alpha ij}\right] = \begin{pmatrix} 1 & 0.05 & 0.05 & 0.05 \\ 0.05 & 1 & 0.05 & 0.05 \\ 0.05 & 0.05 & 1 & 0.05 \\ 0.05 & 0.05 & 0.05 & 1 \end{pmatrix}.$$

We show in the following that the algorithm successfully recovers the true parameters. The Table 1 presents the median, the standard deviation of the posterior distribution for all parameters obtained from MCMC algorithm. We note the posterior means are all near the true values and the true values are always contained in the central interval covering 95% of the posterior mass. We also apply the convergence diagnostics for the parameter chains generated to estimate the above parameters. The Gelman-Rubin statistics R is applied to multiple chains to verify convergence that occurs when R is close to 1, below a critical level. The independent chains were run 20,000 times. In all cases the chains converged to its invariant measure around the true parameter. Finally, simulated data allows us to verify that all parameters are identified and precise inference can be conduced on the basis of a sample that is of magnitude likely to occur in practice.

TABLE 1
Spatial RE Model – results of MCMC for 20000 interactions

θ	True	E(θ)	S.D(θ)	R	θ	True	E(θ)	S.D(θ)	R
σ^2_{11}	0.0500	0.0468	0.0049	1.0004	$\sigma^2_{\alpha 11}$	1.0000	1.2339	0.2627	1.0004
σ^2_{21}	-0.0100	-0.0052	0.0036	0.9999	$\sigma^2_{\alpha 21}$	0.5000	0.7397	0.2161	1.0003
σ^2_{31}	-0.0100	-0.0074	0.0033	1.0000	$\sigma^2_{\alpha 31}$	0.5000	0.6662	0.2013	1.0001
σ^2_{41}	-0.0100	-0.0094	0.0035	0.9999	$\sigma^2_{\alpha 41}$	0.5000	0.5098	0.1712	1.0006
σ^2_{12}	-0.0100	-0.0052	0.0036	0.9999	$\sigma^2_{\alpha 12}$	0.5000	0.7397	0.2161	1.0003
σ^2_{22}	0.0500	0.0509	0.0052	1.0001	$\sigma^2_{\alpha 22}$	1.0000	1.2862	0.2723	1.0003
σ^2_{32}	0.0200	0.0184	0.0037	1.0001	$\sigma^2_{\alpha 32}$	0.5000	0.6373	0.2044	1.0001
σ^2_{42}	0.0200	0.0223	0.0039	1.0002	$\sigma^2_{\alpha 42}$	0.5000	0.5527	0.1750	1.0000
σ^2_{13}	-0.0100	-0.0074	0.0033	1.0000	$\sigma^2_{\alpha 13}$	0.5000	0.6662	0.2013	1.0001

(Continues)

(Continued)

θ	True	E(θ)	S.D(θ)	R	θ	True	E(θ)	S.D(θ)	R
σ^2_{23}	0.0200	0.0184	0.0037	1.0001	$\sigma^2_{\alpha 23}$	0.5000	0.6373	0.2044	1.0001
σ^2_{33}	0.0500	0.0432	0.0045	1.0006	$\sigma^2_{\alpha 33}$	1.0000	1.1476	0.2416	1.0001
σ^2_{43}	-0.0300	-0.0245	0.0037	1.0005	$\sigma^2_{\alpha 43}$	0.5000	0.6412	0.1746	1.0000
σ^2_{14}	-0.0100	-0.0094	0.0035	0.9999	$\sigma^2_{\alpha 14}$	0.5000	0.5098	0.1712	1.0006
σ^2_{24}	0.0200	0.0223	0.0039	1.0002	$\sigma^2_{\alpha 24}$	0.5000	0.5527	0.1750	1.0000
σ^2_{34}	-0.0300	-0.0245	0.0037	1.0005	$\sigma^2_{\alpha 34}$	0.5000	0.6412	0.1746	1.0000
σ^2_{44}	0.0500	0.0473	0.0048	1.0004	$\sigma^2_{\alpha 44}$	1.0000	0.8550	0.1805	1.0002
β_{11}	-0.5000	-0.5111	0.0145	1.0007	φ_{11}	0.3000	0.3257	0.0218	1.0078
β_{21}	0.2000	0.2200	0.0154	1.0002	φ_{22}	0.4000	0.3938	0.0091	1.0127
β_{12}	0.1000	0.0987	0.0155	1.0000	φ_{33}	-0.5000	-0.5029	0.0132	1.0046
β_{22}	0.8000	0.7963	0.0158	1.0004	φ_{44}	0.6000	0.5998	0.0069	1.0014
β_{13}	-0.3000	-0.3170	0.0149	1.0000					
β_{23}	0.6000	0.5885	0.0154	1.0000					
β_{14}	-0.8000	-0.7788	0.0151	1.0001					
β_{24}	0.3000	0.3097	0.0157	1.0001					

4 COMPARISON BETWEEN MODELS

The estimation of the random effect model in the Bayesian multivariate spatial model is time consuming, and the implementation of the MCMC algorithm is not trivial. Particularly, the calibration procedure of the prior parameters related to the spatial correlation is a delicate process. This problem is cumbersome in high dimensionalities samples. In this way, a relevant question is related to simplify the problem to a tractable dimension. For the empirical researcher the fundamental question is related to the values of both the parameters associated with the explicative variables (b´s) and the spatial correlation coefficients (φ's). That is, can alternative models like MS models (Gamerman and Moreira, 2004), that are less time consuming, generate coefficients similar to our estimator?

Table 2 describes the alternative methodologies that will be compared. We are going to generate data from an artificial model elaborated from four endogenous variables and two regressors. In all those models the spatial autocorrelation is represented by a diagonal matrix.

TABLE 2
Alternative methodologies

Models	Econometric specification
Model 1 - MSRE	Multivariate Spatial Regression with Random Effect (Section 3.4)
Model 2 to 5 - US	Univariate Spatial Regression without Random Effect
Model 6 - MS	Multivariate Spatial Regression without Random Effect (Gamerman and Moreira, 2004)
Model 7 to 10 - USRE	Univariate Spatial Regression with Random Effect

Tables 3 to 5 compare the accuracy and efficiency from different models and specifications. The analysis of the parameters of interest [[(β's), (φ's)] shows the gains associated to the use of our methodology.

TABLE 3
MSRE x Univariate Spatial Regression without Random Effect

	Model 1			Model 2		Model 3		Model 4		Model 5	
θ	True	E(θ)	S.D(θ)	E(θ)	S.D(θ)	E(θ)	S.D(θ)	E(θ)	S.D(θ)	E(θ)	S.D(θ)
σ^2_{11}	0,0500	0,0468	0,0049	0.7149	0.0684						
σ^2_{21}	-0,0100	-0,0052	0,0036								
σ^2_{31}	-0,0100	-0,0074	0,0033								
σ^2_{41}	-0,0100	-0,0094	0,0035								
σ^2_{12}	-0,0100	-0,0052	0,0036								
σ^2_{22}	0,0500	0,0509	0,0052			0.8289	0.0802				
σ^2_{32}	0,0200	0,0184	0,0037								
σ^2_{42}	0,0200	0,0223	0,0039								
σ^2_{13}	-0,0100	-0,0074	0,0033								
σ^2_{23}	0,0200	0,0184	0,0037								
σ^2_{33}	0,0500	0,0432	0,0045					0.7803	0.0716		
σ^2_{43}	-0,0300	-0,0245	0,0037								
σ^2_{14}	-0,0100	-0,0094	0,0035								
σ^2_{24}	0,0200	0,0223	0,0039								
σ^2_{34}	-0,0300	-0,0245	0,0037								
σ^2_{44}	0,0500	0,0473	0,0048							0.4554	0.0446
β_{11}	-0,5000	-0,5111	0,0145	-0.493	0.0497						
β_{21}	0,2000	0,2200	0,0154	0.311	0.0539	0.0575	0.0553				
β_{12}	0,1000	0,0987	0,0155			0.8044	0.0575				
β_{22}	0,8000	0,7963	0,0158					-0.3138	0.0580		
β_{13}	-0,3000	-0,3170	0,0149					0.5932	0.0525		
β_{23}	0,6000	0,5885	0,0154							-	0.0401
β_{14}	-0,8000	-0,7788	0,0151							0.7009	
β_{24}	0,3000	0,3097	0,0157							0.2906	0.0434

(Continues)

(Continued)

θ	Model 1 True	Model 1 E(θ)	Model 1 S.D(θ)	Model 2 E(θ)	Model 2 S.D(θ)	Model 3 E(θ)	Model 3 S.D(θ)	Model 4 E(θ)	Model 4 S.D(θ)	Model 5 E(θ)	Model 5 S.D(θ)
φ_{11}	0,3000	0,3257	0,0218	0.728	0.0357						
φ_{22}	0,4000	0,3938	0,0091			0.7307	0.0341				
φ_{33}	-0,5000	-0,5029	0,0132					0.2529	0.0668		
φ_{44}	0,6000	0,5998	0,0069							0.8568	0.0209
σ^2_{11}	1,0000	1,2339									
σ^2_{21}	0,5000	0,7397									
σ^2_{31}	0,5000	0,6662									
σ^2_{41}	0,5000	0,5098									
σ^2_{12}	0,5000	0,7397									
σ^2_{22}	1,0000	1,2862									
σ^2_{32}	0,5000	0,6373									
σ^2_{42}	0,5000	0,5527									
σ^2_{13}	0,5000	0,6662									
σ^2_{23}	0,5000	0,6373									
σ^2_{33}	1,0000	1,1476									
σ^2_{43}	0,5000	0,6412									
σ^2_{14}	0,5000	0,5098									
σ^2_{24}	0,5000	0,5527									
σ^2_{34}	0,5000	0,6412									
σ^2_{44}	1,0000	0,8550									

TABLE 4
MSRE x Multivariate Spatial Regression without Random Effect

θ	Model 1 True	Model 1 E(θ)	Model 1 S.D(θ)	Model 6 E(θ)	Model 6 S.D(θ)
σ^2_{11}	0,0500	0,0468	0,0049	0.6788	0.0623
σ^2_{21}	-0,0100	-0,0052	0,0036	0.3993	0.0548
σ^2_{31}	-0,0100	-0,0074	0,0033	0.2929	0.0508
σ^2_{41}	-0,0100	-0,0094	0,0035	0.1969	0.0379
σ^2_{12}	-0,0100	-0,0052	0,0036	0.3993	0.0548
σ^2_{22}	0,0500	0,0509	0,0052	0.8044	0.0739
σ^2_{32}	0,0200	0,0184	0,0037	0.3989	0.0576
σ^2_{42}	0,0200	0,0223	0,0039	0.2318	0.0418
σ^2_{13}	-0,0100	-0,0074	0,0033	0.2929	0.0508
σ^2_{23}	0,0200	0,0184	0,0037	0.3989	0.0576
σ^2_{33}	0,0500	0,0432	0,0045	0.7838	0.0721
σ^2_{43}	-0,0300	-0,0245	0,0037	0.2279	0.0412
σ^2_{14}	-0,0100	-0,0094	0,0035	0.1969	0.0379
σ^2_{24}	0,0200	0,0223	0,0039	0.2318	0.0418
σ^2_{34}	-0,0300	-0,0245	0,0037	0.2279	0.0412
σ^2_{44}	0,0500	0,0473	0,0048	0.4378	0.0405

(Continues)

(Continued)

θ	Model 1			Model 6	
	True	E(θ)	S.D(θ)	E(θ)	S.D(θ)
β_{11}	-0,5000	-0,5111	0,0145	-0.4803	0.0493
β_{21}	0,2000	0,2200	0,0154	0.2751	0.0517
β_{12}	0,1000	0,0987	0,0155	0.0537	0.0537
β_{22}	0,8000	0,7963	0,0158	0.7762	0.0568
β_{13}	-0,3000	-0,3170	0,0149	-0.3099	0.0527
β_{23}	0,6000	0,5885	0,0154	0.5895	0.0549
β_{14}	-0,8000	-0,7788	0,0151	-0.6806	0.0399
β_{24}	0,3000	0,3097	0,0157	0.2707	0.0418
φ_{11}	0,3000	0,3257	0,0218	0.9081	0.0351
φ_{22}	0,4000	0,3938	0,0091	0.8244	0.0311
φ_{33}	-0,5000	-0,5029	0,0132	0.4016	0.0687
φ_{44}	0,6000	0,5998	0,0069	0.9301	0.0223
$\sigma^2_{\alpha 11}$	1,0000	1,2339			
$\sigma^2_{\alpha 21}$	0,5000	0,7397			
$\sigma^2_{\alpha 31}$	0,5000	0,6662			
$\sigma^2_{\alpha 41}$	0,5000	0,5098			
$\sigma^2_{\alpha 12}$	0,5000	0,7397			
$\sigma^2_{\alpha 22}$	1,0000	1,2862			
$\sigma^2_{\alpha 32}$	0,5000	0,6373			
$\sigma^2_{\alpha 42}$	0,5000	0,5527			
$\sigma^2_{\alpha 13}$	0,5000	0,6662			
$\sigma^2_{\alpha 23}$	0,5000	0,6373			
$\sigma^2_{\alpha 33}$	1,0000	1,1476			
$\sigma^2_{\alpha 43}$	0,5000	0,6412			
$\sigma^2_{\alpha 14}$	0,5000	0,5098			
$\sigma^2_{\alpha 24}$	0,5000	0,5527			
$\sigma^2_{\alpha 34}$	0,5000	0,6412			
$\sigma^2_{\alpha 44}$	1,0000	0,8550			

TABLE 5
MSRE x Univariate Spatial Regression with Random Effect

θ	Model 1			Model 7		Model 8		Model 9		Model 10	
	True	E(θ)	S.D(θ)	E(θ)	S.D(θ)	E(θ)	S.D(θ)	E(θ)	S.D(θ)	E(θ)	S.D(θ)
σ^2_{11}	0,0500	0,0468	0,0049	0.0462	0.0047						
σ^2_{21}	-0,0100	-0,0052	0,0036								
σ^2_{31}	-0,0100	-0,0074	0,0033								
σ^2_{41}	-0,0100	-0,0094	0,0035								
σ^2_{12}	-0,0100	-0,0052	0,0036								
σ^2_{22}	0,0500	0,0509	0,0052			0.0507	0.0052				
σ^2_{32}	0,0200	0,0184	0,0037								
σ^2_{42}	0,0200	0,0223	0,0039								
σ^2_{13}	-0,0100	-0,0074	0,0033								
σ^2_{23}	0,0200	0,0184	0,0037								
σ^2_{33}	0,0500	0,0432	0,0045					0.0483	0.0049		
σ^2_{43}	-0,0300	-0,0245	0,0037								

(Continues)

(Continued)

	Model 1			Model 7		Model 8		Model 9		Model 10	
θ	True	E(θ)	S.D(θ)	E(θ)	S.D(θ)	E(θ)	S.D(θ)	E(θ)	S.D(θ)	E(θ)	S.D(θ)
σ^2_{14}	-0,0100	-0,0094	0,0035								
σ^2_{24}	0,0200	0,0223	0,0039								
σ^2_{34}	-0,0300	-0,0245	0,0037								
σ^2_{44}	0,0500	0,0473	0,0048							0.0517	0.0054
β_{11}	-0,5000	-0,5111	0,0145	-0.5106	0.0144						
β_{21}	0,2000	0,2200	0,0154	0.2196	0.0149						
β_{12}	0,1000	0,0987	0,0155			0.1004	0.0151				
β_{22}	0,8000	0,7963	0,0158			0.7916	0.0159				
β_{13}	-0,3000	-0,3170	0,0149					-0.3163	0.0149		
β_{23}	0,6000	0,5885	0,0154					0.5907	0.0152		
β_{14}	-0,8000	-0,7788	0,0151							-0.7775	0.0154
β_{24}	0,3000	0,3097	0,0157							0.3057	0.0162
φ_{11}	0,3000	0,3257	0,0218	0.3169	0.0279						
φ_{22}	0,4000	0,3938	0,0091			0.4366	0.0185				
φ_{33}	-0,5000	-0,5029	0,0132					-0.4809	0.0304		
φ_{44}	0,6000	0,5998	0,0069							0.6268	0.0179
$\sigma^2_{\alpha 11}$	1,0000	1,2339		1.2556	0.2712						
$\sigma^2_{\alpha 21}$	0,5000	0,7397									
$\sigma^2_{\alpha 31}$	0,5000	0,6662									
$\sigma^2_{\alpha 41}$	0,5000	0,5098									
$\sigma^2_{\alpha 12}$	0,5000	0,7397									
$\sigma^2_{\alpha 22}$	1,0000	1,2862				1.1959	0.2517				
$\sigma^2_{\alpha 32}$	0,5000	0,6373									
$\sigma^2_{\alpha 42}$	0,5000	0,5527									
$\sigma^2_{\alpha 13}$	0,5000	0,6662									
$\sigma^2_{\alpha 23}$	0,5000	0,6373									
$\sigma^2_{\alpha 33}$	1,0000	1,1476						1.1219	0.2362		
$\sigma^2_{\alpha 43}$	0,5000	0,6412									
$\sigma^2_{\alpha 14}$	0,5000	0,5098									
$\sigma^2_{\alpha 24}$	0,5000	0,5527									
$\sigma^2_{\alpha 34}$	0,5000	0,6412									
$\sigma^2_{\alpha 44}$	1,0000	0,8550								0.7888	0.1696

5 APPLICATION: IS THE CATTLE RANCHING RESPONSIBLE FOR DEFORESTATION IN THE BRAZILIAN AMAZON? NEW EVIDENCES

5.1 The Land-Use Model for the Brazilian Amazon

Over the past two decades the international community has become aware of the global and regional environmental risks associated to possible massive forest losses in the Brazilian Amazon. The impacts on global carbon cycle, regional climate and the loss of biodiversity are among the main consequences of extensive land-use change processes in this region. Therefore a good understanding of land-use dynamics in the Brazilian Amazon is a fundamental step before action whether controlling the problem or even changes its course towards a more sustainable path.

Econometric and statistical methods have been used with the objective to determine the main dynamical features between succeeding anthropogenic land-use sequence (Soares-Filho et al., 2006; Reis and Blanco, 1997; PFAFF, 1999; Reis and Guzmán, 1994, *inter allia*). In order to seek for deforestation drivers, it is important to know up to what extent, in past deforestation trends, cropland or pastures uses are more or less keen to immediately follow deforestation. Several studies have already contributed to a better understanding of the underlying processes that drive land-use change in the Amazon.

Andersen et al. (2002; 1997) estimated a model to evaluated the dynamic of land use of Brazilian Amazon over the period 1970 through 1985. But some caveats can be posed about this model. A first difficulty lies in the fact that spatial heterogeneity among cross-section units (municipalities) has not to be taken into account in order to contemplate the geographic, environmental and economic diversity among those different units and the effect of their mutual interactions. In this case, the heterogeneity among units can be contemplated using an appropriated panel data[7] structure. Second, land-use is clearly a spatial process that results from the complex interplay of many phenomena occurring in a much extended spatial domain, it is obviously a spatial phenomenon and we may expect spatial autocorrelation to be present in the data. These two points were not appropriated treated by Andersen et al. (2002; 1997).

The objective of this section is to apply the methodology developed in this paper to estimate the dynamic of land-use for Brazilian Amazon taking into account the issues pointed out in last paragraph. For that we follow the study undertaken by Andersen *et al*. (2002; 1997) where these authors model the dynamic land-use in Amazon using a vector autoregressive. Notwithstanding our study differently considers that all the issues pointed out above deserve to be correctly considered in the same model. In this case the approach used by Andersen et al. (2002; 1997) must be reformulated in order to take into account the panel data structure of the data base and spatial dependence.

7. Important sources of variation may be left out if the data is only pooled in a single (temporal or spatial) dimension, and more precise parameter estimates can be obtained in panel approaches that explore the variability present in the data both across counties and within counties over time.

Hence, we propose to use a panel data vector autoregressive PVAR[8] with spatial autocorrelation to estimate the land-use model of Brazilian Amazon. This model can be seen as a particular case of the model studied in sections 3.4 and 3.5 that assumes the following specification

$$Y = WY\Phi + XB + V \qquad (15')$$

where $X = Y(t-1)$ and V, the matrix of disturbance, is defined such as $V = (I_N \otimes i_T)\alpha + U$. Here $\alpha \sim MN_{N\times Q}(0_{N\times Q}, \Sigma_\alpha \otimes I_N)$ and $U \sim MN_{TN\times Q}(0_{TN\times Q}, \Sigma_u \otimes I_{TN})$.

This is a one order lag spatial PVAR with random effect. The vector of endogenous variables is given by $y_{it} = (forest_{it}, pastureit_{it}, fallow_{it}, crop_{it})'$. In this vector each variable represents one category of land-use that is explained in the next section. To summarize, our model tries to explain the evolution of land-use in Brazilian Amazon over time by the following factors: the heterogeneity among cross-section units (municipalities) due to the geographic, environmental, the sources of spatial interactions among municipalities and economic diversity among those different units and the state of the land-use occupation in the last period.

5.2 Database

The database available for this study comes from Brazilian National Agriculture Census elaborated by the Brazilian Institute for Geography and Statistics (IBGE) which is usually conducted every five years. Others original data sources used are the Industrial and Commercial Census that were also elaborated by the IBGE for the same periods. The data were collected for the following for years 1970, 1975, 1980, 1985 and 1995 at the municipality level. The data were cleaned, harmonized and merged with data of other sources by the Ipea (Institute for Applied Economic Research) managed by the team of Ipeadata.[9] The original database includes data on economic, demographic, ecological and agriculture variables.

In the Brazilian Amazon Basin a county can be subjected to ongoing change in its size mainly during the expansion of the agricultural frontier in Amazon. This fact obstructs the comparison between periods at county level. That is why the concept of

8. Panel vector autoregressive.

9. Available at: <http://www.ipeadata.gov.br>.

Minimum Comparable Area (MCA)[10] was introduced, which is the smallest stable spatial unity during these five censuses that accommodates the changing county boundaries over the panel. The aggregation of counties in the later census years, in order to match the county area in 1970, is greatest in the more recently populated and sub-divided regions found in the legal Amazon.

The agricultural censuses group all land into private land and public land. Private land is stratified into eight categories according to agricultural use. These are *i*) annual crops; *ii*) perennial crops; *iii*) planted forest; *iv*) planted pasture; *v*) short fallow; and *vi*) long fallow are classified in cleared land; while *vii*) natural forest and natural pasture are considered non-cleared land. A small category of private non-usable land (rivers, mountains, etc.) is also considered non-cleared land. Finally, all land that is not claimed by anyone is considered public land and by definition non-cleared.

Based on these definitions the dependent variables used in our land-use model fall into one of the following four categories: cropland (*crop*), pasture (*pasture*), fallow (*fallow*) and natural land (*forest*). Cropland covers annual crops, perennial crops and planted forest. Pasture is planted pasture only. Fallow land includes short fallow, long fallow and non-usable land like roads, dams, etc. Finally, natural land considers natural forest and natural pasture.

5.3 Results and Analysis

We now apply the Bayesian methods to the data set described in the last section. This land-use data set contains N=X individual MCA, T=5 time periods and Q=4 endogenous variables (forest, pasture, fallow and crops). In Table 6 we present the estimated results of the model (15'). The sequence of the endogenous variable is the same that appears in vector y_{it} in section 4.1. In this table the estimated coefficients can be read in the following way: $\boldsymbol{\Sigma}_u=[\sigma^2_{ij}]$, $\boldsymbol{\Phi}=[\phi_{ij}]$, and $B=[\beta_{ij}]$. For instance, the coefficient β_{21} means the effect of one unit change of *forest* area in the last period on the occupation of the current area of *pasture*. The logit transformation was applied to the variables to become the hypothesis of error normality acceptable.[11] Table 6 presents the median, standard deviation and the 0.025 and 0.975 quantiles of the posterior distribution for all parameters. Evidence on convergence in accordance to statistics R is shown.

10. For further details about the concept of MCA see Reis et al. (2007).

11. If x is the proportion of an area in a certain region and time, this amount is converted to y = h (x/1–x).

It can be seen from table 6 that matrix Φ has not non-zero elements. It implies that the spatial interaction plays important role in the process of land-use in the Amazon. Results reported in this table also show that wood extraction does not seem to affect agricultural activities since the zero belongs to the confidence interval for the posterior distribution of β_{41}, i.e., this coefficient is concentrated around zero. Looking the coefficient β_{44} in *crops* equation also reveals that the current expansion of agricultural activities strongly induces to an increasing in the area of *crop* in the future. A weak positive relation is observed between agriculture and cattle in the sense that expansion of cattle weakly affect the future increasing in the area of crops. But, the reverse is not true. The coefficient β_{24} is 0.16 while the coefficient β_{42} is 0.02. It can indicate some form of complementarity between these activities in the sense that more agriculture induces to the future need of more area of cattle.

Finally, differently from the common wisdom we do not any find evidence that cattle ranching are an important driver of deforestation. The analysis of forest equation shows that deforestation is guide by itself. In other words, the deforestation is connected to any other economic activity. It is important to note that agriculture is indeed a factor that does not reduces deforestation in the view that the coefficient B_{14} is negative.

TABLE 6
Spatial RE Model - Results of MCMC for 20000 interactions

			Quantil						Quantil		
θ	E(θ)	S.D(θ)	2,5%	97,5%	R	θ	E(θ)	S.D(θ)	2,5%	97,5%	R
σ^2_{11}	2.00	0.10	1.84	2.16	1.00	σ^2_{11}	0.19	0.06	0.11	0.29	1.00
σ^2_{21}	-2.13	0.16	-2.39	-1.87	1.00	σ^2_{21}	0.01	0.07	-0.12	0.11	1.00
σ^2_{31}	-0.61	0.15	-0.86	-0.36	1.00	σ^2_{31}	0.10	0.10	-0.07	0.27	1.00
σ^2_{41}	-0.04	0.03	-0.09	0.01	1.00	σ^2_{41}	-0.02	0.03	-0.06	0.03	1.00
σ^2_{12}	-2.13	0.16	-2.39	-1.87	1.00	σ^2_{12}	0.01	0.07	-0.12	0.11	1.00
σ^2_{22}	8.41	0.40	7.76	9.08	1.00	σ^2_{22}	0.57	0.19	0.29	0.90	1.00
σ^2_{32}	0.59	0.31	0.09	1.10	1.00	σ^2_{32}	0.56	0.21	0.23	0.91	1.00
σ^2_{42}	0.13	0.06	0.03	0.23	1.00	σ^2_{42}	-0.18	0.06	-0.27	-0.09	1.00
σ^2_{13}	-0.61	0.15	-0.86	-0.36	1.00	σ^2_{13}	0.10	0.10	-0.07	0.27	1.00
σ^2_{23}	0.59	0.31	0.09	1.10	1.00	σ^2_{23}	0.56	0.21	0.23	0.91	1.00
σ^2_{33}	9.06	0.48	8.30	9.86	1.00	σ^2_{33}	1.50	0.38	0.90	2.16	1.00
σ^2_{43}	0.30	0.06	0.19	0.41	1.00	σ^2_{43}	-0.28	0.07	-0.40	-0.16	1.00
σ^2_{14}	-0.04	0.03	-0.09	0.01	1.00	σ^2_{14}	-0.02	0.03	-0.06	0.03	1.00
σ^2_{24}	0.13	0.06	0.03	0.23	1.00	σ^2_{24}	-0.18	0.06	-0.27	-0.09	1.00
σ^2_{34}	0.30	0.06	0.19	0.41	1.00	σ^2_{34}	-0.28	0.07	-0.40	-0.16	1.00
σ^2_{44}	0.35	0.02	0.32	0.38	1.00	σ^2_{44}	0.22	0.03	0.18	0.27	1.00

(Continues)

(Continued)

			Quantil						Quantil		
θ	E(θ)	S.D(θ)	2,5%	97,5%	R	θ	E(θ)	S.D(θ)	2,5%	97,5%	R
β_{11}	0.23	0.03	0.18	0.27	1.00	φ_{11}	0.07	0.01	0.05	0.08	1.03
β_{21}	0.09	0.01	0.07	0.11	1.00	φ_{22}	0.06	0.01	0.04	0.07	1.01
β_{12}	0.02	0.01	0.00	0.04	1.00	φ_{33}	0.06	0.01	0.05	0.07	1.01
β_{22}	-0.20	0.03	-0.24	-0.15	1.00	φ_{44}	0.03	0.00	0.02	0.03	1.00
β_{13}	0.35	0.06	0.26	0.44	1.00						
β_{23}	0.23	0.03	0.18	0.28	1.00						
β_{14}	0.12	0.02	0.08	0.16	1.00						
β_{24}	0.16	0.06	0.06	0.25	1.00						
β_{11}	0.42	0.06	0.32	0.52	1.00						
β_{21}	0.21	0.03	0.16	0.27	1.00						
β_{12}	0.11	0.03	0.07	0.15	1.00						
β_{22}	0.83	0.07	0.73	0.94	1.00						
β_{13}	0.01	0.01	-0.01	0.03	1.00						
β_{23}	0.02	0.01	0.01	0.03	1.00						
β_{14}	0.02	0.01	0.01	0.03	1.00						
β_{24}	0.76	0.02	0.73	0.79	1.00						

6 FINAL COMMENTS

We develop a new Bayesian estimator that is able to deal with multivariate panel data structure in the presence of spatial correlation. The Monte Carlo simulations demonstrate the ability of this new estimator to replicate quite well simulated data.

To show the empirical relevance of this new estimator we apply it to deforestation data in the Brazilian Amazon. The empirical results suggest that differently from the common wisdom we do not find any evidence that cattle ranching are an important driver of deforestation.

REFERENCES

ANDERSEN, L. E. *et al.* **The dynamic of deforestation and economic growth in the brazilian amazon**. Cambridge: Cambridge University Press, 2002.

ANDERSEN, L. E.; GRANGER, C.; REIS, E. J. A randon coefficient var transition model of the changes in land-use in brazilian amazon. **Revista Brasileira de Econometria**, v. 17, n. 1, p. 1-13, 1997.

ANSELIN, L. Spatial econometrics methods and models. **Kluwer Academic Publishers**, v. 4, 1988.

ARELLANO, M. **Panel data econometrics**. Oxford: Oxford University Press, 2003.

BALTAGI, B. H. Simultaneous equations with error components. **Journal of Econometrics**, v. 17, p. 189-200, 1981.

______. Econometric analysis of panel data**. John Wiley and Sons**, v. 4, 1995.

BALTAGI, B. H.; QI, L. A note on the estimation of simultaneous equations with error components. **Econometric Theory**, v. 8, n. 1, p. 113-119, 1992.

CHIB, S.; CARLIN, B. On MCMC sampling in hierarchical longitudinal models. **Statistics and Computing**, v. 9, p. 17-26, 1999.

CHIB, S.; GREENBERG, E. Understanding metropolis-hasting algorithm. **The American Statistician**, v. 49, n. 4, p. 327-335, 1995.

CORNWELL, C.; SCHMIDT, P.; WYHOWSKI, D. Simultaneous equations and panel data. **Journal of Econometrics**, v. 51, p. 151-181, 1992.

DAGPUNAR, J. S. **Simulation and Monte Carlo:** with applications in finance and MCMC. Wiley: England, 2007.

FAHRMEIR, L.; LANG, S. Bayesian inference for generalized additive mixed models based on Markov random field priors. **Applied Statist**, v. 50, p. 201-220, 2001.

GAMERMAN, D.; LOPES, H. F. **Markov chain Monte Carlo**: stochastic simulation for Bayesian inference. London: Chapman & Hall, 2006.

GAMERMAN, D.; MOREIRA, A. R. B. Multivariate spatial regression model. **Journal of Multivariate Analysis**, v. 91, p. 262-281, 2004.

GELMAN, A.; RUBIN, D. R. Inference from iterative simulation using multiple sequences, **Statist. Sci**, v. 7, p. 457-511, 1992. (With discussion).

GELMAN, A. *et al.* **Bayesian data analysis**. Chapman & Hall/CRC, 2003.

GREENBERG, E. **Introduction to bayesian econometrics**. Cambridge: Cambridge University Press, 2007.

HAMILTON, J**. Time series analysis**. Princeton: Princeton University Press, 1993.

KINAL, T.; LAHIRI, K. A computational algorithm for multiple equation models with panel data. **Economics Letters**, v. 34, p. 143-146**,** 1990.

______. On the estimation of simultaneous-equations error-components with an application to a model of developping country foreign trade. **Journal of Applied Econometrics**, v. 8, n. 1, p. 81-92, 1993.

KOOP, G. **Bayesian econometrics**. New Jersey: John Wiley-Interscience, 2003.

LAIRD, N. M.; WARE, J. H. Random effects models for longitudinal data. **Biometrics**, v. 38, p. 963-974, 1982.

LANCASTER, T. **Introduction to modern bayesian econometrics**. New Jersey: Wiley-Blackweel, 2004.

LANG, S. *et al.* Bayesian geoadditive seemingly unrelated regression. **Statistics**, v. 18, p. 263-292, 2003.

PRESS, S. J. **Bayesian statistics**: principle, models and applications. New Jersey: Wiley, 1989.

REIS, E. J.; BLANCO, F. A. **The causes of brazilian amazon deforestation**. IPEA, 1997. (Working Paper)

REIS, E.; GUZMÁN, R. An econometric model of amazon deforestation. *In:* BROWN, K.; PEARCE, D. (Eds.) **The causes of tropical deforestation, the economic and statistical analysis of factors giving rise to the loss of tropical forests**, p. 172-91. London: University College London Press, 1994.

REIS, E.; PIMENTEL, J.; ALVARENGA, A. I. Áreas mínimas comparáveis entre os anos de 1920 a 2000. WP NEMESIS, 2007.

ROBERT, C.; CASELLA, G. **Monte Carlo statistical methods**. Springer, 2005.

SIMS, C. (1980). Macroeconomics and reality. **Econometrica**, v. 48, n. 1, p. 1-48, 1980.

SOARES-FILHO, B. S.; *et. al.* Modeling conservation in the amazon basin. **Nature**, v. 440, p. 520-523, 2006.

Ipea – Institute for Applied Economic Research

PUBLISHING DEPARTMENT

Coordination
Cláudio Passos de Oliveira

Supervision
Everson da Silva Moura
Reginaldo da Silva Domingos

Typesetting
Bernar José Vieira
Cristiano Ferreira de Araújo
Daniella Silva Nogueira
Danilo Leite de Macedo Tavares
Jeovah Herculano Szervinsk Junior
Leonardo Hideki Higa

Cover design
Luís Cláudio Cardoso da Silva

Graphic design
Renato Rodrigues Buenos

The manuscripts in languages other than Portuguese published herein have not been proofread.

Ipea Bookstore

SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo
70076-900 – Brasília – DF
Brazil
Tel.: + 55 (61) 2026 5336
E-mail: livraria@ipea.gov.br

Composed in Adobe Garamond 11/13.2 (text)
Frutiger 47 (headings, graphs and tables)
Brasília – DF – Brazil

Ipea's mission

Enhance public policies that are essential to Brazilian development by producing and disseminating knowledge and by advising the state in its strategic decisions.

ISSN 1415-4765

ipea Institute for Applied Economic Research

MINISTRY OF PLANNING, DEVELOPMENT AND MANAGEMENT

BRASIL BRAZILIAN GOVERNMENT