The Interplay Between the Bayesian and Frequentist Approaches: A General Nesting Spatial Panel Data Model

Andrés Ramírez Hassan*

June 14, 2016

Abstract

We propose an econometric framework mixing the Frequentist and Bayesian approaches to estimate a general nesting spatial model. First, it avoids specific dependency structures between unobserved heterogeneity and regressors, which improves mixing properties of MCMC procedures in the presence of unobserved heterogeneity. Second, it allows model selection based on a strong statistical framework, characteristics that are not easily introduced using a Frequentist approach. We perform some simulation exercises finding good performance of the properties of our approach, and apply our methodology to analyse the relation between productivity and public investment in the USA.

JEL Classification: C11, C23.
Keywords: Bayesian Econometrics, General Nesting Spatial Model, Model Selection, Spatial Econometrics.

*Department of Economics, School of Economics and Finance, Universidad EAFIT, Carrera 49 # 7 Sur-50, Medellín, Colombia; email: aramir21@eafit.edu.co. Phone: (57)(4) 2619500 Ext. 9402

1 INTRODUCTION

Spatial econometrics is an area in econometrics that incorporates spatial effects into conventional econometric methods (Anselin, 1988). These effects emerge due to spatial heterogeneity or spatial dependence between cross sectional units. At first, these methods originated from spatial statistics (Cressie, 1993), and later were extended to regional science and geographical economics. In the past twenty five years, spatial econometrics has made significant advances, and nowadays it is common to find applications to different fields of economics, as well as theoretical developments (Anselin, 2001; Elhorst, 2001, 2003; Baltagi and Li, 2006; Baltagi et al., 2007).

Although spatial panel data models are part of this prominent development (see Anselin et al. (2008) for an excellent review), to our knowledge, few authors have proposed to estimate these models from a Bayesian perspective, and none mix the Frequentist and Bayesian approaches (see Lee and Yu (2010b) for the different approaches that have been used, and Chib (2008) for an excellent review of the estimation of Bayesian panel models: unfortunately, it does not include spatial panel models). There are some methodological, pedagogical and philosophical advantages to each approach, although at the methodological level the debate has become considerably muted, with the recognition that each approach has a great deal to contribute to statistical practice (Good, 1992; Bayarri and Berger, 2004; Kass, 2011).

In spatial econometrics, it is well known that some Frequentist methods, like Instrumental Variable estimation and the Generalized Method of Moments, may generate estimates that are outside of the stationary region, while other methods, like Maximum Likelihood, do not suffer from this shortcoming (Elhorst, 2014). Bayesian methods do not suffer from such a shortcoming. The latter framework also allows calculating the Posterior Model Probability, which helps to select the contiguity matrix that best supports the data, or introduce model uncertainty through Bayesian Model Averaging. On the other hand, Frequentist methods, such as Likelihood Ratio tests, require nested models, and Information Criteria do not permit consistently introducing model uncertainty. Moreover, the Bayesian approach leads to computational simplification in spatial panel models (Parent and LeSage, 2010). Parent and LeSage (2011, 2012) argue that Maximum Likelihood estimation can be stuck in a local optimum, which may generate erroneous inferences, whereas Bayesian Markov chain Monte Carlo methods avoid this. However, using a Bayesian framework can generate poor mixing properties of the Markov chain Monte Carlo procedure in the presence of unobserved heterogeneity (Best et al., 1999). In addition, Bayesian panel data models with fixed effects explicitly introduce specific dependency structures between unobserved heterogeneity and the regressors (Greenberg, 2008), whereas fixed effects are treated as nuisance parameters in a Frequentist approach, so that some simple transformations permit controlling for them.

Most of the spatial panel literature is focused on a Frequentist approach (Elhorst, 2014), although a Bayesian approach offers good advantages (Ohtsuka et al., 2010; Parent and LeSage, 2010, 2011, 2012). Therefore, the main objective of the present paper is to propose a methodological approach to estimate a general nesting spatial model (GNS, Elhorst (2014)) in a panel data setting. We will take into consideration the advantages of both statistical schools using a simple Frequentist transformation avoiding specific dependency structures between unobserved heterogeneity and the regressors, and improving the mixing properties of the Markov chain Monte Carlo procedure in the presence of unobserved heterogeneity, but also taking advantage of a Bayesian approach such as model selection or model uncertainty.

Introducing fixed effects allows controlling for dependence between the unobserved heterogeneity of the cross sectional units and the regressors, a characteristic that is not present in panel data models with random effects, but that is very noticeable in econometric applications. Moreover, we introduce spatial Autoregressive effects that allow estimating the global direct and indirect (spillover) effects due to neighbours as well as spatial Autoregressive disturbances that allow capturing uncontrollable transitory effects, and also spatial lags of the independent variables that generate local spatial effects. In addition, we calculate the marginal likelihood of our model to calculate the Posterior Model Probability (PMP). As a consequence, we can select the contiguity matrix that best fits the database on the PMPs or apply Bayesian Model Averaging to take into account uncertainty regarding this issue. Choosing a contiguity matrix is crucial in interaction models like spatial models (Stetzer, 1982; Griffith and Lagona, 1998; Florax and De Graaff, 2004; Paez et al., 2008; Stakhovych and Bijmolt, 2009; Seya et al., 2013).

Unfortunately there is little guidance for the choice of this matrix (Anselin, 2002). Although simulation exercises have shown that Information Criteria (Kakamu and Wago, 2008) and the spatial J test (Piras and Lozano-Garcia, 2012) help to select the contiguity matrix, these mechanisms do not allow model averaging omitting model uncertainty, and can be computationally expensive if many candidates exist (Seya et al., 2013).

We perform some simulation exercises to check the properties of our approach regarding the use of the Posterior Model Probabilities to select the contiguity matrix that best supports the data. Finally, we apply our methodology to introduce spatial effects and select the contiguity matrix that best fits the USA data regarding productivity and public investment (Munnell, 1990b; Baltagi and Pinnoi, 1995; Baltagi, 2003).

After this introduction, we develop in the following section our methodological approach. We perform some simulation exercises to check the performance of the PMP in uncovering the true contiguity matrix in Section (3). Section (4) exhibits the results of our application. Finally, we will make some concluding remarks.

2 THE MODEL

In fact, the term 'fixed effects' is misleading: both models, fixed and random, have unobserved heterogeneity which is random –the former is correlated with some regressors, whereas the latter is a purely random effect (Cameron and Trivedi, 2005)– Therefore, a Bayesian approach should take into consideration this distinction.

There are two common approaches in Bayesian panel data models with fixed effects: the individual effect model using dummy variables (Koop, 2003) and the hierarchical structure that incorporates a dependence between the fixed effects and the exogenous variables (Greenberg, 2008). Under the recognition that at the methodological level the Frequentist and Bayesian schools of statistics have contributed to statistical practice (Good, 1992; Bayarri and Berger, 2004; Kass, 2011), we follow the most common approach in Frequentist econometrics, where fixed effects are treated as nuisance parameters, so that we get rid of them using an orthonormal eigenvector transformation, but perform a conditioned Bayesian analysis. This approach has two advantages: first it avoids a specific dependency structure between unobserved heterogeneity and the regressors, and second, it improves the mixing properties of the Markov chain Monte Carlo procedure in the presence of unobserved heterogeneity (Best et al., 1999).

Our point of departure is the general nesting spatial model. Although we must recognize that this model has little empirical relevance, this model encompasses the most relevant spatial models (Elhorst, 2014).

$$\begin{aligned} &\mathbf{Y} = \lambda \mathbf{W}_N \mathbf{Y} + \alpha_1 \mathbf{Z}_1 + \cdots + \alpha_K \mathbf{Z}_K + \delta_1 \mathbf{W}_N \mathbf{Z}_1 + \cdots + \delta_K \mathbf{W}_N \mathbf{Z}_K + \mathbf{U} + \mathbf{E} \\ &\mathbf{E} = \rho \mathbf{W}_N \mathbf{E} + \mathbf{V} \end{aligned} \tag{1}$$

where

$$\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1T} \\ y_{21} & y_{22} & \cdots & y_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ y_{N1} & y_{N2} & \cdots & y_{NT} \end{bmatrix}, \mathbf{Z}_k = \begin{bmatrix} z_{k11} & z_{k12} & \cdots & z_{k1T} \\ z_{k21} & z_{k22} & \cdots & z_{k2T} \\ \vdots & \vdots & \ddots & \vdots \\ z_{kN1} & z_{kN2} & \cdots & z_{kNT} \end{bmatrix}, k = 1, 2, \ldots, K,$$

$$\mathbf{U} = \begin{bmatrix} \mu_1 & \mu_1 & \cdots & \mu_1 \\ \mu_2 & \mu_2 & \cdots & \mu_2 \\ \vdots & \vdots & \ddots & \vdots \\ \mu_N & \mu_N & \cdots & \mu_N \end{bmatrix}, \mathbf{E} = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \cdots & \epsilon_{1T} \\ \epsilon_{21} & \epsilon_{22} & \cdots & \epsilon_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{N1} & \epsilon_{N2} & \cdots & \epsilon_{NT} \end{bmatrix}$$

and

$$\mathbf{V} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1T} \\ v_{21} & v_{22} & \cdots & v_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ v_{N1} & v_{N2} & \cdots & v_{NT} \end{bmatrix}$$

Taking into account that for any matrix $\mathbf{L}_{N \times T}$ and scalar γ, $vec(\gamma \mathbf{W}_N \mathbf{L}) = \gamma vec(\mathbf{W}_N \mathbf{L} \mathbf{I}_T) = \gamma(\mathbf{I}_T \otimes \mathbf{W}_N) vec(\mathbf{L})$, we can rewrite our model using the vectorization operator in the following form

$$
\begin{aligned}
vec(\mathbf{Y}) &= \lambda(\mathbf{I}_T \otimes \mathbf{W}_N)vec(\mathbf{Y}) + \alpha_1 vec(\mathbf{Z}_1) + \cdots + \alpha_K vec(\mathbf{Z}_K) + \delta_1(\mathbf{I}_T \otimes \mathbf{W}_N)vec(\mathbf{Z}_1) \\
&\quad + \cdots + \delta_K(\mathbf{I}_T \otimes \mathbf{W}_N)vec(\mathbf{Z}_K) + vec(\mathbf{U}) + vec(\mathbf{E}) \\
vec(\mathbf{E}) &= \rho(\mathbf{I}_T \otimes \mathbf{W}_N)vec(\mathbf{E}) + vec(\mathbf{V})
\end{aligned}
$$

where $vec(\mathbf{U}) = (i_T \otimes I_N)\mu$, $\mu' = [\mu_1, \mu_2, \ldots, \mu_N]$.

This system is equal to

$$
\begin{aligned}
y &= \lambda\left(I_T \otimes W_N\right) y + Z\alpha + \left(I_T \otimes W_N\right) Z\delta + \left(i_T \otimes I_N\right)\mu + \epsilon \qquad (2) \\
\epsilon &= \rho\left(I_T \otimes W_N\right)\epsilon + v
\end{aligned}
$$

where $y = vec(\mathbf{Y})$ and $\epsilon = vec(\mathbf{E})$ are $NT \times 1$ vectors, I_T and I_N are the identity matrices of dimensions $T \times T$ and $N \times N$, W_N is a row standardized contiguity matrix of dimension $N \times N$, $Z = [vec(\mathbf{Z}_1)vec(\mathbf{Z}_2)\ldots vec(\mathbf{Z}_K)]$ is a matrix of regressors of dimension $NT \times K$, μ is an $N \times 1$ vector of fixed errors that are potentially correlated with Z, i_T is a vector of ones whose dimension is T, the parameters $\alpha = [\alpha_1\alpha_2\ldots\alpha_K]'$, $\delta = [\delta_1\delta_2\ldots\delta_K]'$, λ, ρ and σ_v^2 are to be estimated, $v = vec(\mathbf{V}) \sim \mathcal{N}(0, (I_T \otimes I_N)\sigma_v^2)$, N is the number of spatial units and T is the time span.[1]

If we set $\delta = 0$, $\lambda = 0$ and $\rho = 0$, we obtain a panel data model with fixed effects but without spatial effects. We obtain an SAR panel data setting when $\lambda \neq 0$, and $\delta = 0$ and $\rho = 0$, there is a spatial Durbin model (SDM) when $\lambda \neq 0$ and $\delta \neq 0$, and $\rho = 0$. In addition, we obtain an SEM panel setting when $\lambda = 0$ and $\delta = 0$ but $\rho \neq 0$, a spatial Durbin error model (SDEM) when $\lambda = 0$, and $\delta \neq 0$ and $\rho \neq 0$, and a spatial lag of Z model (SLZ) when $\lambda = 0$ and $\rho = 0$ but $\delta \neq 0$. Depending on the presence of local or global spatial effects this general specification should be modified. For instance, local spillover specifications should set $\lambda = 0$, whereas global spillover specifications should set $\lambda \neq 0$ (LeSage, 2014). Introducing simultaneously $\lambda \neq 0$ and $\rho \neq 0$ increases precision but generates complexity in estimation, and may create artificial spa-

[1] Traditional panel data are generally ordered using time as the fast index. On the other hand, spatial panel data are stacked first by time and then by cross sectional units in most of the cases (Millo and Piras, 2012).

tial correlation when the true data generating process is a spatial Durbin model.[2]

To eliminate the issue of incidental parameters, the deviation from the time mean operator, $Q^0 \otimes I_N$, where $Q^0 = (I_T - i_T i_T'/T)$, can be used. However, the resulting stochastic disturbance would be linearly dependent over the time dimension. Thus, Lee and Yu (2010a) propose using the orthonormal eigenvector matrix of Q^0. Set $[F^0, 1/\sqrt{T} i_T]$ to be the orthonormal eigenvector matrix of Q^0, where F^0 is the $T \times (T-1)$ submatrix corresponding to the first $T-1$ eigenvalues such that $F^{0\prime} i_T = 0_{(T-1)\times 1}, F^{0\prime} F^0 = I_{T-1}$ and $F^0 F^{0\prime} = I_T - i_T i_T'/T$. Post–multiplying the system (1) by F^0,

$$
\begin{aligned}
&\mathbf{Y}F^0 = \lambda \mathbf{W}_N \mathbf{Y} F^0 + \alpha_1 \mathbf{Z}_1 F^0 + \cdots + \alpha_K \mathbf{Z}_K F^0 + \delta_1 \mathbf{W}_N \mathbf{Z}_1 F^0 + \cdots + \delta_K \mathbf{W}_N \mathbf{Z}_K F^0 + \mathbf{U}F^0 + \mathbf{E}F^0 \\
&\mathbf{E}F^0 = \rho \mathbf{W}_N \mathbf{E} F^0 + \mathbf{V} F^0
\end{aligned}
$$

where $\mathbf{U}F^0 = \mu i_T' F^0 = \mu 0_{1\times(T-1)} = 0_{N\times(T-1)}$. Thus

$$
\begin{aligned}
&\mathbf{Y}^* = \lambda \mathbf{W}_N \mathbf{Y}^* + \alpha_1 \mathbf{Z}_1^* + \cdots + \alpha_K \mathbf{Z}_K^* + \delta_1 \mathbf{W}_N \mathbf{Z}_1^* + \cdots + \delta_K \mathbf{W}_N \mathbf{Z}_K^* + \mathbf{E}^* \\
&\mathbf{E}^* = \rho \mathbf{W}_N \mathbf{E}^* + \mathbf{V}^*
\end{aligned}
$$

where $\mathbf{Y}^* = \mathbf{Y}F^0, \mathbf{Z}_1^* = \mathbf{Z}_1 F^0, \ldots, \mathbf{Z}_K^* = \mathbf{Z}_K F^0$ and $\mathbf{V}^* = \mathbf{V}F^0$.

Using again the vectorization operator

$$
\begin{aligned}
vec(\mathbf{Y}^*) &= \lambda(\mathbf{I}_{T-1} \otimes \mathbf{W}_N) vec(\mathbf{Y}^*) + \alpha_1 vec(\mathbf{Z}_1^*) + \cdots + \alpha_K vec(\mathbf{Z}_K^*) + \delta_1 (\mathbf{I}_{T-1} \otimes \mathbf{W}_N) vec(\mathbf{Z}_1^*) \\
&\quad + \cdots + \delta_K (\mathbf{I}_{T-1} \otimes \mathbf{W}_N) vec(\mathbf{Z}_K^*) + vec(\mathbf{E}^*) \\
vec(\mathbf{E}^*) &= \rho(\mathbf{I}_{T-1} \otimes \mathbf{W}_N) vec(\mathbf{E}^*) + vec(\mathbf{V}^*)
\end{aligned}
$$

where for any matrix $\mathbf{L}_{N\times T}$, we have that $vec(\mathbf{L}F^0) = vec(I_N \mathbf{L} F^0) = (F^{0\prime} \otimes I_N) vec(\mathbf{L})$. The

[2] We thank an anonymous referee for pointing out this issue.

last system is equal to

$$
\begin{aligned}
y^* &= \lambda \left(I_{(T-1)} \otimes W_N\right) y^* + Z^*\alpha + \left(I_{(T-1)} \otimes W_N\right) Z^*\delta + \epsilon^* \\
\epsilon^* &= \rho \left(I_{(T-1)} \otimes W_N\right) \epsilon^* + v^*
\end{aligned}
$$

where $y^* = Q^*y$, $Z^* = Q^*Z$, $\epsilon^* = Q^*\epsilon$, $v^* = Q^*v$ and $Q^* = F^{0\prime} \otimes I_N$ is an $N(T-1) \times NT$ matrix. Observe that $E\left[v^*v^{*\prime}\right] = E\left[(F^{0\prime} \otimes I_N)vv'(F^0 \otimes I_N)\right] = \sigma_v^2(F^{0\prime}F^0 \otimes I_N) = \sigma_v^2(I_{T-1} \otimes I_N)$.

We set $X = [Z^* \quad \left(I_{(T-1)} \otimes W_N\right) Z^*]$ and $\beta = [\alpha' \quad \delta']'$ to simplify notation. So the last equations can be expressed as

$$
\begin{aligned}
y^* &= \lambda \left(I_{(T-1)} \otimes W_N\right) y^* + X^*\beta + \epsilon^* \\
\epsilon^* &= \rho \left(I_{(T-1)} \otimes W_N\right) \epsilon^* + v^*
\end{aligned}
\tag{3}
$$

Likelihood and Priors

Set $\Sigma(\rho) = \left(I_{T-1} \otimes (I_N - \rho W_N)^{-1}\right)\left(I_{T-1} \otimes (I_N - \rho W_N')^{-1}\right)$, $\tilde{y}^*(\lambda) = y^* - \lambda\left(I_{T-1} \otimes W_N\right) y^*$ and $\Omega(\lambda) = \left(I_{T-1} \otimes (I_N - \lambda W_N)^{-1}\right)\left(I_{T-1} \otimes (I_N - \lambda W_N')^{-1}\right)$. Then, the likelihood function of Equation 3 assuming Gaussian stochastic perturbations is[3]

$$
f\left(\beta, \sigma_v^2, \lambda, \rho\right) = (2\pi)^{-N(T-1)/2}(\sigma_v^2)^{-N(T-1)/2}|\Sigma(\rho)|^{-1/2}|\Omega(\lambda)|^{-1/2} Exp\left\{-\frac{1}{2\sigma_v^2}(\tilde{y}^*(\lambda) - X^*\beta)'\Sigma(\rho)^{-1}(\tilde{y}^*(\lambda) - X^*\beta)\right\} \tag{4}
$$

We employ a Bayesian framework in which the posterior distributions of the parameters are derived using vague conjugate and non-informative priors. We assume that the prior distribution for the parameters $\pi\left(\beta, \sigma_v^2, \lambda, \rho\right)$ has the following structure:

$$
\pi\left(\beta, \sigma_v^2, \lambda, \rho\right) = \pi\left(\beta|\sigma_v^2\right)\pi\left(\sigma_v^2\right)\pi\left(\lambda\right)\pi\left(\rho\right) \tag{5}
$$

[3] The likelihood in Equation (4) is a conditional likelihood on $\bar{y}$, which is the time average for each cross sectional unit stacked in a vector. $\bar{y}$ is a sufficient statistic for μ under the assumption of normality (Lee and Yu, 2010a).

where $\pi\left(\beta|\sigma_v^2\right)$ is $\mathcal{N}(\beta_0, \sigma_v^2 D)$,[4] $\pi\left(\sigma_v^2\right)$ is $\mathcal{IG}(e/2, e/2)$, $\pi(\lambda)$ and $\pi(\rho)$ are $\mathcal{U}(-1, 1)$ (LeSage and Parent, 2007; Kakamu and Wago, 2008).[5] In addition, we set $\beta_0 = 0$, which means that we centre this prior over the hypothesis that the explanatory variables do not have an effect on the dependent variable, and $D = 1,000I$, which implies a vague normal density. Following (Spiegelhalter et al., 2003), we use a 'non-informative' prior distribution in the variance parameter, $\mathcal{IG}(e/2, e/2)$, $e \to 0$. In particular, we use $\sigma_v^2 \sim \mathcal{IG}(0.001, 0.001)$ as the prior density for the variance parameter.[6] Setting 'non-informative' prior distributions over parameters which are common to all models is acceptable when using posterior odds ratios to select models (Koop, 2003).

Posterior inference

The full conditional posteriors are given by the following expressions:
$\beta|\sigma_v^2, \lambda, \rho : y, X, W_N \sim \mathcal{N}(\beta(\lambda, \rho)^*, \sigma_v^2(G'G)^{-1})$ such that $\beta(\lambda, \rho)^* = (X^{*\prime}\Sigma^{-1}X^* + D^{-1})^{-1}(D^{-1}\beta_0 + (X^{*\prime}\Sigma^{-1}X^*)\widehat{\beta}(\lambda, \rho)_{MV})$ where $\widehat{\beta}(\lambda, \rho)_{MV} = (X^{*\prime}\Sigma^{-1}X^*)^{-1}X^{*\prime}\Sigma^{-1}\tilde{y}^*$ and $(G'G)^{-1} = (X^{*\prime}\Sigma^{-1}X^* + D^{-1})^{-1}$.[7]

In addition, $\sigma_v^2|\lambda, \rho : y, X, W_N \sim \mathcal{IG}\left(\frac{e^*}{2}, \frac{\delta^*(\lambda,\rho)}{2}\right)$ where $e^* = N(T-1) + e$ and $\delta^*(\lambda, \rho) = (\tilde{y}^* - X^*\beta(\lambda, \rho)^*)' \, \Sigma^{-1} \, (\tilde{y}^* - X^*\beta(\lambda, \rho)^*) + (\beta_0 - \beta(\lambda, \rho)^*)'B^{-1}(\beta_0 - \beta(\lambda, \rho)^*) + e$.

It is not possible to obtain the conditional posterior distributions of the parameters λ and ρ, so we use the Metropolis–Hastings algorithm (Metropolis et al., 1953; Hastings, 1970; Chib and Greenberg, 1995) nested with the Gibbs sampler (Geman and Geman, 1984; Robert and Casella, 2004): this is known as the 'Metropolis-within-Gibbs' method.[8]

[4] Another common approach for setting the prior covariance matrix in spatial econometric models is to use Zellner's g prior (LeSage and Parent, 2007). We would just need to replace D with $(gX'X)^{-1}$ to achieve this specification.

[5] In fact, the stationarity of the model requires that λ and ρ are within the interval $(1/\omega_{min}, 1)$ for a row standardized contiguity matrix, where ω_{min} is the minimum eigenvalue of W_N (Ord, 1975; Anselin, 1982; Elhorst, 2014). However, the eigenvalues of a non-symmetric contiguity matrix may be complex numbers, so we set $\mathcal{U}(-1, 1)$ as prior distributions for λ and ρ.

[6] Kakamu and Wago (2008) achieve similar outcomes using a formulation like ours and improper priors in a Spatial Autoregressive panel data model.

[7] We omit the dependence on ρ and λ to simplify the notation. In addition, note that $\Sigma^{-1} = I_{T-1} \otimes (I_N - \rho W_N')(I_N - \rho W_N))$.

[8] We use the Metropolis–Hastings algorithm using the random walk kernel (Greenberg, 2008). We ob-

Posterior model probability

We use the Posterior Model Probability to find, in the spatial setting, a Bayesian solution to the problem of the uncertainty involving the contiguity matrix (LeSage and Parent, 2007; LeSage and Fischer, 2008; Cotteleer et al., 2011). In particular, with $\mathcal{W} = \{W_1, W_2, \ldots, W_M\}$ denoting the set of contiguity matrices, the PMP is

$$p(W_i|y) = \frac{p(y|W_i)p(W_i)}{\sum_{j=1}^{M} p(y|W_j)p(W_j)} \tag{6}$$

where $p(W_j)$ is the prior contiguity matrix probability, which is equal to $1/M$, making each contiguity matrix equally likely a priori, and $p(y|W_i)$ is called the marginal likelihood, which can be considered as a likelihood function for W_i.

tain good acceptance rates using this strategy in our simulation exercises and application (Approximately 0.6 (Robert and Casella, 2004)). Kakamu and Wago (2008) use the same strategy in a similar framework.

$$
\begin{aligned}
p(y|W_i) =& \int_{(-1,1)\times(-1,1)} \int_{\mathcal{R}^{2K}} \int_{\mathcal{R}^{++}} f\left(\beta, \sigma_v^2, \lambda, \rho\right) \pi\left(\beta|\sigma_v^2\right) \pi\left(\sigma_v^2\right) \pi\left(\lambda\right) \pi\left(\rho\right) d\sigma_v^2 d\beta d(\lambda,\rho) \\
=& \int_{(-1,1)\times(-1/,1)} \int_{\mathcal{R}^{2K}} (2\pi)^{-\left(\frac{N(T-1)+2K}{2}\right)} (|A||B|)^{T-1} |D|^{-1/2} \frac{1}{2^2} \frac{(e/2)^{(e/2)}}{\Gamma(e/2)} \\
& \int_{\mathcal{R}^{++}} (\sigma_v^2)^{-\left(\frac{N(T-1)+2K+e}{2}+1\right)} Exp\left\{-\frac{1}{2\sigma_v^2}\{\left(w - G\beta(\lambda,\rho)^*\right)'(w - G\beta(\lambda,\rho)^*)\right. \\
& \left. + \left(\beta - \beta(\lambda,\rho)^*\right)' G'G\left(\beta - \beta(\lambda,\rho)^*\right)\} - \frac{e}{2\sigma_v^2}\right\} d\sigma_v^2 d\beta d(\lambda,\rho) \\
=& \int_{(-1,1)\times(-1,1)} (2\pi)^{-\left(\frac{N(T-1)+2K}{2}\right)} (|A||B|)^{T-1} |D|^{-1/2} \frac{1}{2^2} \frac{(e/2)^{(e/2)}}{\Gamma(e/2)} \Gamma\left(\frac{N(T-1)+2K+e}{2}\right) \\
& \int_{\mathcal{R}^{2K}} \left(\frac{(w - G\beta(\lambda,\rho)^*)'(w - G\beta(\lambda,\rho)^*) + (\beta - \beta(\lambda,\rho)^*)'G'G(\beta - \beta(\lambda,\rho)^*) + e}{2}\right)^{-\left(\frac{N(T-1)+2K+e}{2}\right)} \\
& d\beta d(\lambda,\rho) \\
=& \int_{(-1,1)\times(-1,1)} (2\pi)^{-\left(\frac{N(T-1)+2K}{2}\right)} (|A||B|)^{T-1} |D|^{-1/2} \frac{1}{2^2} \frac{(e/2)^{(e/2)}}{\Gamma(e/2)} \Gamma\left(\frac{e^* + 2K}{2}\right) \\
& \left(\frac{z(\lambda,\rho)^* e^*}{2}\right)^{-\left(\frac{e^*+2K}{2}\right)} \int_{\mathcal{R}^{2K}} \left(1 + \frac{(\beta - \beta(\lambda,\rho)^*)'\left(\frac{G'G}{z(\lambda,\rho)^*}\right)(\beta - \beta(\lambda,\rho)^*)}{e^*}\right)^{\left(-\frac{e^*+2K}{2}\right)} \\
& d\beta d(\lambda,\rho) \\
=& (\pi)^{-\left(\frac{N(T-1)}{2}\right)} \frac{\Gamma(e^*/2)}{\Gamma(e/2)} (e)^{(e/2)} \frac{1}{2^2} |D|^{-1/2} \int_{(-1,1)\times(-1,1)} (|A||B|)^{T-1} \delta^*(\lambda,\rho)^{-(e^*/2)} \\
& |(G'G)|^{-1/2} d(\lambda,\rho) \\
\propto& \int_{(-1,1)\times(-1,1)} (|A||B|)^{T-1} \delta^{*-(e^*/2)} |(G'G)|^{-1/2} d(\lambda,\rho)
\end{aligned}
\tag{7}
$$

where $w(\lambda,\rho) = \begin{bmatrix} \Sigma(\rho)^{-1/2}\tilde{y}(\lambda)^* \\ D^{-1/2}\beta_0 \end{bmatrix}$, $G(\rho) = \begin{bmatrix} \Sigma(\rho)^{-1/2}X^* \\ D^{-1/2} \end{bmatrix}$, $A(\lambda) = I_N - \lambda W_N$, $B(\rho) = I_N - \rho W_N$ and $z(\lambda,\rho)^* = \delta^*(\lambda,\rho)/e^*$.

Observe that this a general formulation where computation is simplified if there is just local spatial spillovers because in this case $\lambda = 0$. On the other hand, global spatial spillover effects without global diffusion of shocks implies $\rho = 0$. In addition, the selection of the spatial contiguity matrix also impacts the marginal likelihood through the spatial lag of Z which is present in X.

We should clarify that the constant of proportionality in Equation (7) is the same in all models, because the only difference between them is associated with different contiguity matrices. So it is not relevant to calculate this constant for matrix selection. In addition, we can see from this equation that when $\lambda = \rho = 0$ and $T = 2$, $p(y|W_i)$ is the marginal likelihood that is obtained in the Bayesian linear regression (Koop, 2003). However, our expression does not have an analytical solution, in contrast to the Bayesian linear regression, thus we must use numerical integration to solve it, and select the contiguity matrix that best supports the data.[9]

Observe that the PMP is based on Bayes Factors (Jeffreys, 1961), which in turn are firmly grounded in the theory of probability and statistical decision theory, avoiding an *ad hoc* selection of significance levels that do not take into account the implicit balance of losses (Jeffreys, 1961; Bernardo and Smith, 1994). In addition, Bayes Factors minimize the weighted sum of type I and type II error probabilities (DeGroot, 1975; Pericchi and Pereira, 2015). From a methodological perspective, we have statistical arguments to select between non-nested models using Bayes Factors, and an easy way to take into account model uncertainty through Bayesian Model Averaging.

Finally, we can select the model with the highest Posterior Model Probability because it minimizes a 0–1 loss function for correct selection (Clyde and George, 2004). Optionally, our approach allows performing Bayesian Model Averaging in case we want to take into account contiguity matrix uncertainty (LeSage and Parent, 2007; Cotteleer et al., 2011).

3 SIMULATION EXERCISES

We perform two simulation exercises to check the performance of our methodology (see Kakamu and Wago (2008) for extensive simulation exercises comparing ML and Bayesian Spatial Autoregressive

[9] There are different approaches to calculating the marginal likelihood in the Bayesian setting, most of them based on the MCMC outcomes, for instance: Chib (1995) and Chib and Jeliazkov (2001) developed methods to deduce the marginal likelihood from the Gibbs sampler and the Metropolis–Hastings outputs, respectively. Although we might have used these approaches in our paper, we do not, because numerical integration can be applied in our setting without much effort.

panel data models.). In the first, the data generating process of our dependent variable is

$$
\begin{aligned}
y_{it} = 0.9 \sum_{i \sim j} w_j y_{jt} + x_{1it} - 0.5x_{2it} + 3x_{3it} - 2x_{4it} + 2x_{5it} + 0.5x_{6it} - 3x_{7it} + 0.4x_{8it} + 3x_{9it} \\
+ x_{10it} - 2x_{11it} + 2x_{12it} - 0.8x_{13it} + x_{14it} 2x_{15it} - x_{16it} + 0.5x_{17it} + \mu_i + \epsilon_{it}
\end{aligned}
$$

$$
\epsilon_{it} = 0.7 \sum_{i \sim j} w_j \epsilon_{jt} + v_{1it}
$$

where $v_{1it} \overset{i.i.d.}{\sim} \mathcal{N}(0, 0.5^2)$, i, $j = 1, 2, \ldots, 50$, $t = 1, 2, \ldots, 5$ and $i \sim j$ accounts for the j neighbours of i such that $\sum_{i \sim j} w_j = 1$ for each i. In addition, $x_{kit} \overset{i.i.d.}{\sim} \mathcal{N}(10, 1^2)$, $k = 3, 4, \ldots, 9$, $x_{kit} \overset{i.i.d.}{\sim} \mathcal{B}(2, 1)$, $k = 10, 11, \ldots, 14$, $x_{kit} \overset{i.i.d.}{\sim} \mathcal{P}(5)$, $k = 15, 16, 17$ and

$$
\begin{bmatrix} x_{1it} \\ x_{2it} \\ x_{0it} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & -0.5 \\ 0 & 1 & 0.5 \\ -0.5 & 0.5 & 1 \end{bmatrix} \right)
$$

$\mu_i = \frac{1}{5} \sum_{t=1}^{5} x_{0it}$ so by construction the fixed effects are correlated with x_{1it} and x_{2it}.

We draw hypothetical distances (d_{ij}) from an exponential distribution with mean 25 to build our contiguity matrices. The main idea is to approximate our simulation exercise to our application where we use the Euclidean distance to create most of our contiguity matrices. We create 20 contiguity matrices using the same matrix of distances from the exponential distribution. In particular, we define 20 cutoffs (c_l) equally spaced between 5 and 30, then we build 20 binary contiguity matrices such that $w_{lij} = 1$ iff $d_{ij} \leq c_l$, $l = 1, 2, \ldots, 20$. Recall that by definition, $w_{lii} = 0$. Finally, we standardize each matrix by row because this facilitates the interpretation of the weights as constructing a weighted average of the neighbouring values (Anselin, 1988). We use $c_5 = 10.26$ to simulate y_{it}, so this cutoff defines the true contiguity matrix.

We perform Maximum Likelihood and Bayesian estimations of the spatial panel data with fixed effects model using the right contiguity matrix. In particular, we implement the 'Metropolis-within-Gibbs' algorithm (see Section 2) using 22,000 iterations and a burn-in of 2,000. Then, we draw an observation every 10 iterations to have an effective sample size of 2,000. This last step is done to mitigate the autocorrelation of the chains. We compute several diagnostics to assess the convergence and stationarity of the chains. In particular, we employ the method due to Heidelberger and Welch (1983), the mean difference test proposed by Geweke (1992), and the diagnostic from Raftery et al. (1992). All the chains seem stable, and diagnostics indicate that the chains converge to stationary distributions (outcomes available upon request).

We can see from Table 1 that the Maximum Likelihood and Bayesian estimates are very similar, and close to the population parameters. In general, Frequentist estimates are significant at the 5%, and credible intervals do not embrace zero in the Bayesian case.

We estimate our model using the 20 contiguity matrices to illustrate the relevance of selecting these matrices in our setting. We see in Table 2 the range of variability of the parameter estimates as well as the range of variability of the direct and indirect effects (spillover effects), which are the relevant measures to capture marginal effects (LeSage and Fischer, 2008). For instance, we have that $\hat{\rho}$ varies between -0.214 and 0.951 or $\hat{\beta}_{13}$ between -0.985 and -0.051. Regarding the direct and indirect effects, the range of variability in the difference between the maximum and minimum values oscillates between 0.064 and 1.850 for the former, and 0.037 and 0.374 for the latter. This simple simulation exercise shows the relevance of the contiguity matrix.

To determine the ability of the PMP to identify the right contiguity matrix in the face of different values of the spatial effects (λ and ρ), we perform simulation exercises where the data generating process of y_{it} is determine as above, but the spatial effects change according to the values that are shown in Table 3. We calculate the PMP associated with the right contiguity matrix, the one that uses $c_5 = 10.26$ as cutoff, for each of the 361 scenarios. As we can see in Table 3, when there are meaningful spatial autoregressive effects, that is $\lambda < -0.2$ and $\lambda > 0.1$, our procedure assigns a remarkable PMP to the right contiguity matrix. Obviously, if $\lambda = 0$ and $\rho = 0$ there are no spatial effects, and as a consequence the contiguity matrix does not have any significant role in our econometric framework, so the PMP should be very low.

The second simulation exercise uses the dataset that will be used in our application. In particular, we use the exogenous variables: Public Capital, Private Capital Stock, Employment in non-agricultural payrolls, and the Unemployment rate, all variables in logs, except the latter, from a panel data set consisting of 48 US states from 1970 to 1986 provided by Baltagi (2003). Our data generating process is the following

$$y_{it} = 0.20log(Public\ Capital_{it}) + 0.35log(Private\ Capital_{it}) + 0.55log(Employment_{it}) \\ - 0.007Unemployment\ Rate_{it} + \mu_i + \epsilon_{it}$$

$$\epsilon_{it} = \rho \sum_{i \sim j} w_j \epsilon_{jt} + v_{1it}$$

where $v_{1it} \overset{i.i.d.}{\sim} \mathcal{N}(0, 0.075^2)$, $i, j = 1, 2, \ldots, 48$, $t = 1, 2, \ldots, 17$ and $i \sim j$ accounts for the j neighbours of i such that $\sum_{i \sim j} w_j = 1$ for each i. In addition, $\mu_i = \frac{1}{17} \sum_{t=1}^{17} 0.4log(Public\ Capital_{it}) + \tau_{it}$, $\tau_{it} \sim \mathcal{N}(0, 1)$. This implies that there is a fixed effect that is correlated to an exogenous variable.

We create 18 row standardized contiguity matrices based on the Euclidean distance between the states' centroids. Specifically, we define $w_{lij} = 1$ iff $d_{ij} \leq c_l$, $l = 1, 2, \ldots, 18$ where d_{ij} is the Euclidean distance, and c_l are cutoffs (see the first column of Table 4). In addition, we use the Rook and Queen criteria.

We use the fifth contiguity matrix to create our synthetic dependent variable, and use different levels of spatial autocorrelation in the stochastic shocks ($\rho = 0, 0.1, \ldots, 0.9$). We calculate the Posterior Model Probability (equation 7) associated with the 20 contiguity matrices under different scenarios of spatial autocorrelation. As can be seen in Table 4 the contiguity matrix that was used to simulate our dependent variable, that is the fifth matrix when the cutoff is equal to 11.90, has the highest PMP in all cases, except when $\rho = 0$, that is when there is not spatial autocorrelation. We can see in this table that the PMP associated with the true matrix is higher when the level of spatial dependency is higher. In general, we observe that the PMP identifies well the contiguity matrix.

4 EMPIRICAL ILLUSTRATION

To demonstrate the performance of our methodological proposal in an empirical setting, we analyse Munnell (1990b)'s problem regarding the relation between productivity and public investment. In particular, we use the panel data set consisting of 48 US states from 1970 to 1986 provided by Baltagi (2003). The dependent variable is the Gross State Product, and the explanatory variables are Public Capital, Private Capital Stock, Employment in non-agricultural payrolls, and the Unemployment rate (see Baltagi and Pinnoi (1995) for a detailed description). All variables are taken in logs, except the Unemployment rate, as is done in Baltagi and Pinnoi (1995) (Model 2 in their paper). This data structure was originally treated by Munnell (1990a) using a pooled cross-sectional econometric methodology. This data set is also used in Baltagi and Pinnoi (1995); Croissant and Millo (2008); Millo and Piras (2012).

Our methodology controls by state-specific fixed effects such as the endowment of natural resources, the quality of the public infrastructure, and the ability to attract and use foreign investment. Baltagi and Pinnoi (1995) show that fixed effects should be introduced when public capital is involved in the model's specification. In addition we introduce spatial effects to capture spatial spillover effects as well as spatial dependencies between states. Finally, time effects are not significant at conventional levels given state specific effects (Baltagi and Pinnoi, 1995).

We present some descriptive statistics in Table (5). The average annual gross state product is US$ 61,014.32 with a standard deviation of US$ 69,973.90. The public capital and private capital averaged US$25,036.66 and US$ 58,188.29, respectively. Additionally, the average annual employment was 1,747.10, with a standard deviation of 1,855.99, and the average unemployment rate was 6.60% with a standard deviation equal to 2.23%.

The spatial weights matrices are specified as row standardized. We define $w_{lij} = 1$ iff $d_{ij} \leq c_l$, $l = 1, 2, \ldots, 18$ where d_{ij} is the Euclidean distance between the states' centroids, and c_l are cutoffs (see the first column of Table 7). In addition, we use the k–nearest neighbours ($k = \{1, 2, 3, 4, 5\}$), and Rook and Queen criteria.

We perform Moran's I and Local Moran's I tests to identify the presence of local or global spillover effects in gross state product. We can see in Table 6 that there is no statistical evidence supporting global spatial effects, whereas Local Moran's I tests indicate that gross Delaware product may have local spatial correlation. As a consequence we specify SDMA models using the 25 different contiguity matrices.

We can see in Table 7 that the highest Posterior Model Probabilities are 0.54 and 0.46, and they are associated with the contiguity matrices rook and queen. This implies an average of 4.4 links. It is remarkable that the other contiguity criteria (Euclidean distance and k–nearest neighbours) exhibit PMPs approximately equal to 0. So, the use of these two criteria could generate misleading outcomes regarding parameter estimates as well as calculations of direct and indirect effects.

We estimate our model using Markov chain Monte Carlo techniques (Robert and Casella, 2004). In particular, we use the 'Metropolis-within-Gibbs' method. After running the chains for 62,000 iterations, we discard the first 2,000 and draw an observation every 15 iterations to get an effective sample size of 4,000. We show, in Table (8) in Appendix 6, that in general all the chains achieve convergence and stationarity.[10]

We present in Figures (1), (2), (3), (4), (5), (6), (7), (8), (9) and (10) the main outcomes using the contiguity matrix with the highest Posterior Model Probability (similar outcomes are obtained using queen criteria, and are available upon request). As we can see, all median posterior estimates yield results similar to those of the Maximum Likelihood estimates. In addition,

[10]All the simulation exercises and posterior analyses were performed using R (R Core Team, 2014).

we present in each figure the posterior 0.025 and 0.975 quantiles, as well as the histogram and density estimates.

The 95% credible interval of the spatial autoregressive disturbances parameter is (0.68, 0.77) (see Figure (1)), and $2 \times log\ B_{10} = \infty$ for $H_0 : \rho \in (-\infty, 0]$ and $H_1 : \rho \in (0, \infty)$. This is decisive evidence against H_0 (Jeffrey, 1935; Kass and Raftery, 1995; Greenberg, 2008). As a consequence there are very strong positive unobservable spatial effects that affect the gross state product. The mean and median posterior estimates of this parameter are both 0.73.

The mean and median posterior parameter estimates of the elasticity of public capital are both approximately equal to 0.16 (see Figure (2)). In addition, $2 \times log\ B_{10} = \infty$ for $H_0 : \beta_1 \in (-\infty, 0]$ and $H_1 : \beta_1 \in (0, \infty)$. This is decisive evidence against H_0. Regarding the elasticity of private capital, $2 \times log\ B_{10} = \infty$ for $H_0 : \beta_2 \in (-\infty, 0]$ and $H_1 : \beta_1 \in (0, \infty)$. This is decisive evidence to support a positive elasticity whose posterior mean and median are both 0.33 with a 95% credible interval equal to (0.31, 0.35) (see Figure (3)). The mean as well as the median posterior employment elasticity is 0.57. We find, again using the Bayes Factor, decisive evidence of a positive elasticity. Its 95% credible interval is (0.54, 0.60) (see Figure (4)). On the other hand, there is decisive evidence of a negative effect of the unemployment rate on the gross state product, although its magnitude is small due to its mean being equal to -0.0078 (see Figure (5)). Regarding the spatial lags of the independent variables, we found that there are not statistical significant effects (see figures 6, 7, 8 and 9). This implies that there are not local spatial spillover effects due to these variables. Finally, the mean and median posterior estimates of the standard deviation are both equal to 0.059, the 95% credible interval being (0.056, 0.062) (see Figure (10)).

Our findings show that public capital has a remarkable effect on a state's productivity in the USA. This fact agrees with the previous literature on this topic (Munnell, 1990b; Baltagi and Pinnoi, 1995; Baltagi, 2003). In particular, we obtained similar outcomes to those shown in Table 1 in Baltagi and Pinnoi (1995), which in turn are those from Munnell (1990b). As a consequence, these results suggest that although there are some common explanations for the decline in Multifactor Productivity in the 1970s, such as the 'Baby boom' effect, and decreasing in investment in private capital, R&D expenditures as well as maturation in some industries, a remarkable component is associated with the near cessation of investment in public infrastructure. In addition, we find that the latter effect could be exacerbated by the global diffusion of unobserved spatial interactions that generated positive effects.

5 CONCLUDING REMARKS

We have proposed a Bayesian framework using a Frequentist transformation to estimate a general nesting spatial panel model. Our procedure takes into consideration the advantages of both statistical schools, avoiding specific dependency structures between unobserved heterogeneity and the regressors, and improving the mixing properties of the Markov chain Monte Carlo procedure in the presence of unobserved heterogeneity. We show, through simulation exercises and an application, that our methodological proposal obtains similar outcomes as a Maximum Likelihood estimation. However, our procedure has the methodological advantages of the Bayesian framework. For instance, it identifies the contiguity matrix that best supports the data based on a strong statistical foundation. This allows measuring model uncertainty or performing Bayesian Model Averaging.

We applied our methodology to analyse the relation between productivity and public capital in 48 states from 1970 to 1986 in the USA. Our procedure identifies that two of the most used contiguity criteria, Rook and Queen, have high support from the Posterior Model Probability. In particular, it seems that there is not support of global and local spatial spillover effects on the gross state product, whereas there is a significant effect of the global diffusion process of unobserved shocks, which implies a positive effect. In addition, we found decisive statistical evidence that public and private capital, employment, and the unemployment rate are relevant factors in determining productivity.

Future research should include model selection and model uncertainty regarding covariates.

In addition, we should carry out the statistical inference of functions of parameter estimates such as the Multifactor Productivity in our application.

References

Anselin, L. (1982). A note to Small Sample properties of estimators in a First-Order Spatial Autoregressive model. *Environment and Planning A*, 14(8):1023–1030.

Anselin, L. (1988). *Spatial Econometrics: Methods and models*. Kluwer, The Netherlands.

Anselin, L. (2001). Spatial effects in econometric practice in environmental and resource economics. *American Journal of Agricultural Economics*, 83(3):705–710.

Anselin, L. (2002). Under the hood: Issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, 27(3):247–267.

Anselin, L., Gallo, L. L., and Jayet, J. (2008). *Spatial Panel Econometrics. The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*. Springer–Verlag, Berlin.

Baltagi, B. and Li, D. (2006). Prediction in the Panel Data model with Spatial Correlation: The case of liquor. *Spatial Economic Analysis*, 1(2):175–185.

Baltagi, B., Song, S., and Koh, W. (2007). Testing for Serial Correlation, Spatial Autocorrelation and Random Effects using Panel data. *Journal of Econometrics*, 140(1):5–51.

Baltagi, B. H. (2003). *Econometric Analysis of Panel Data*. Wiley, New York.

Baltagi, B. H. and Pinnoi, N. (1995). Public capital stock and state productivity growth: Further evidence. *Empirical Economics*, 20:351–359.

Bayarri, M. J. and Berger, J. (2004). The interplay of Bayesian and Frequentist analysis. *Statistical Science*, 19(1):58–80.

Bernardo, J. and Smith, A. (1994). *Bayesian Theory*. Wiley, Chichester, England.

Best, N. G., Waller, L. A., Thomas, A., Conlon, E. M., and Arnold, R. A. (1999). Bayesian models for spatially correlated diseases and exposure data. In Bernardo, J. M., editor, *Bayesian Statistics*, pages 131–156. Oxford University Press.

Cameron, A. and Trivedi, P. (2005). *Microeconometrics Methods and Applications*. Cambridge University Press.

Chib, S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Assocation*, 90:1313–1321.

Chib, S. (2008). Panel data modeling and inference: a Bayesian premier. In *The Econometrics of Panel Data*, pages 479–515. Springer Berlin Heidelberg.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4):327–335.

Chib, S. and Jeliazkov, I. (2001). Marginal likelihood from the Metropolis–Hastings output. *Journal of American Statistical Association*, 96(453):270–282.

Clyde, M. and George, E. (2004). Model uncertainty. *Statistical Science*, 19(1):81–94.

Cotteleer, G., Stobbe, T., and Kooten, C. V. (2011). Bayesian model averaging in the context of spatial hedonic pricing: An application to farmland values. *Journal of Regional Science*, 51(3):540–557.

Cressie, N. (1993). *Statistics for Spatial Data*. Willey, New York, revised edition.

Croissant, Y. and Millo, G. (2008). Panel data econometrics in R: The plm package. *Journal of Statistical Software*, 27(2):1–43.

DeGroot, M. H. (1975). *Probability and Statistics*. Addison-Wesley, London.

Elhorst, J. (2001). Dynamic models in space and time. *Geographical Analysis*, 33:119–140.

Elhorst, J. (2003). Specification and estimation of Spatial Panel Data models. *International Regional Sciences Review*, 26(3):244–268.

Elhorst, J. (2014). *Spatial Econometrics from Cross–Sectional Data to Spatial Panels*. Springer–Verlag, Berlin.

Florax, R. J. and De Graaff, T. (2004). The performance of diagnostic tests for spatial dependence in linear regression models: a meta-analysis of simulation studies. In Anselin, L., Florax, R. J., and Rey, S., editors, *Advances in Spatial Econometrics*, pages 29–65. Springer, Berlin.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M., editors, *Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting*, pages 169–193. Oxford University Press.

Good, I. J. (1992). The Bayes/non Bayes compromise: A brief review. *Journal of the American Statistical Association*, 87(419):597–606.

Greenberg, E. (2008). *Introduction to Bayesian Econometrics*. Cambridge University Press.

Griffith, D. A. and Lagona, F. (1998). On the quality of likelihood-based estimators in spatial autoregressive models when the data dependence structure is misspecified. *Journal of Statistical Planning and Inference*, 69(1):153–174.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.

Heidelberger, P. and Welch, P. D. (1983). Simulation run length control in the presence of an initial transient. *Operations Research*, 31(6):1109–1144.

Jeffrey, H. (1935). Some test of significance, treated by the theory of probability. *Proceedings of the Cambridge Philosophy Society*, 31:203–222.

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

Kakamu, K. and Wago, H. (2008). Small-sample properties of panel spatial autoregressive models: Comparison of the Bayesian and Maximum Likelihood methods. *Spatial Economic Analysis*, 3(3):305–319.

Kass, R. (2011). Statistical inference: The big picture. *Statistical Science*, 26(1):1–9.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.

Koop, G. (2003). *Bayesian Econometrics*. Wiley, New York.

Lee, L. and Yu, J. (2010a). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154:165–185.

Lee, L. and Yu, J. (2010b). Some recent developments in spatial panel data models. *Regional Science and Urban Economics*, 40(5):255–271.

LeSage, J. (2014). What regional scientists need to know about spatial econometrics. *The Review of Regional Studies*, 44(1):13–32.

LeSage, J. and Fischer, M. (2008). Spatial growth regressions: Model specification, estimation and interpretation. *Spatial Economic Analysis*, 3(3):275–304.

LeSage, J. and Parent, O. (2007). Bayesian model averaging for spatial econometric models. *Geographical Analysis*, 39:241–267.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.

Millo, G. and Piras, G. (2012). splm: Spatial Panel Data models in R. *Journal of Statistical Software*, 47(1):1–38.

Munnell, A. H. (1990a). How does public infrastructure affect regional economic performance? *The New England Economic Review*, pages 11–32.

Munnell, A. H. (1990b). Why has productivity declined? Productivity and public investment. *The New England Economic Review*, pages 3–22.

Ohtsuka, Y., Oga, T., and Kakamu, K. (2010). Forecasting electricity demand in Japan: A Bayesian spatial autoregressive ARMA approach. *Computational Statistics and Data Analysis*, 54:2721–2735.

Ord, J. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70:120–297.

Paez, A., Scott, D. M., and Volz, E. (2008). Weight matrices for social influence analysis: An investigation of measurement errors and their effect on model identification and estimation quality. *Social Networks*, 30(4):309–317.

Parent, O. and LeSage, J. (2010). A spatial dynamic panel model with random effects applied to commuting times. *Transportation Research Part B*, 44:633–645.

Parent, O. and LeSage, J. (2011). A space-time filter for panel data models containing random effects. *Computational Statistics and Data Analysis*, 55:475–490.

Parent, O. and LeSage, J. (2012). Spatial dynamic panel data models with random effects. *Regional Science and Urban Economics*, 42:727–738.

Pericchi, L. and Pereira, C. (in press 2015). Adaptative significance levels using optimal decision rules: Balancing by weighting the error probabilities. *Brazilian Journal of Probability and Statistics*.

Piras, G. and Lozano-Garcia, N. (2012). Spatial J-test: some Monte Carlo evidence. *Statistics and Computing*, 22(1):169–183.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.

Raftery, A. E., Lewis, S., et al. (1992). How many iterations in the Gibbs sampler? *Bayesian Statistics*, 4(2):763–773.

Robert, C. P. and Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer–Verlag, Berlin, 2. edition.

Seya, H., Yamagata, Y., and Tsutsumi, M. (2013). Automatic selection of a spatial weight matrix in spatial econometrics: Application to a spatial hedonic approach. *Regional Science and Urban Economics*, 43(3):429–444.

Spiegelhalter, D., Best, T., Gilks, W., and Lunn, D. (2003). BUGS: Bayesian inference using Gibbs sampling. Technical report, MRC Biostatistics Unit, England. www.mrc-bsu.cam.ac.uk/bugs/.

Stakhovych, S. and Bijmolt, T. H. (2009). Specification of spatial models: A simulation study on weights matrices. *Papers in Regional Science*, 88(2):389–408.

Stetzer, F. (1982). Specifying weights in spatial forecasting models: The results of some experiments. *Environment and Planning A*, 14(A):571–584.

6 APPENDIX

Table 1: Comparison between Bayesian and Maximum Likelihood estimates of a GNS panel model.

Parameter	d.g.p.	Maximum Likelihood	Posterior Mean
λ	0.900	0.858	0.858
		(0.008)	(0.009)
ρ	0.700	0.540	0.503
		(0.132)	(0.152)
β_1	1.000	0.997	0.999
		(0.033)	(0.076)
β_2	-0.500	-0.519	-0.515
		(0.032)	(0.076)
β_3	3.000	2.954	2.953
		(0.031)	(0.075)
β_4	-2.000	-2.031	-2.030
		(0.031)	(0.072)
β_5	2.000	1.996	1.990
		(0.033)	(0.076)
β_6	0.500	0.506	0.506
		(0.032)	(0.073)
β_7	-3.000	-3.007	-3.007
		(0.031)	(0.071)
β_8	0.400	0.309	0.310
		(0.032)	(0.075)
β_9	3.000	3.011	3.012
		(0.031)	(0.074)
β_{10}	1.000	0.836	0.837
		(0.138)	(0.329)
β_{11}	-2.000	-2.217	-2.202
		(0.136)	(0.316)
β_{12}	2.000	2.178	2.165
		(0.132)	(0.312)
β_{13}	-0.800	-0.781	-0.784
		(0.127)	(0.291)
β_{14}	1.000	0.730	0.737
		(0.145)	(0.335)
β_{15}	2.000	2.004	2.002
		(0.013)	(0.032)
β_{16}	-1.000	-0.999	-0.999
		(0.015)	(0.035)
β_{17}	0.500	0.497	0.499
		(0.014)	(0.033)
σ_v^2	0.500	0.430	0.486
		(0.023)	(0.024)

Standard deviation in parentheses.

Table 2: Variability of estimates using different contiguity matrices.

	Parameter		Direct Effect		Indirect Effect	
	Minimum	Maximum	Minimum	Maximum	Minimum	Maximum
λ	0.376	0.864	NA	NA	NA	NA
ρ	-0.214	0.951	NA	NA	NA	NA
β_1	0.997	1.155	1.069	1.278	0.014	0.144
β_2	-0.649	-0.492	-0.719	-0.543	-0.082	-0.007
β_3	2.825	2.954	2.878	3.264	0.036	0.366
β_4	-2.116	-1.985	-2.345	-2.022	-0.266	-0.025
β_5	1.865	1.996	1.874	2.205	0.023	0.248
β_6	0.406	0.656	0.410	0.722	0.006	0.079
β_7	-3.049	-2.920	-3.336	-2.948	-0.382	-0.036
β_8	0.231	0.325	0.233	0.358	0.003	0.040
β_9	2.926	3.245	2.965	3.595	0.036	0.410
β_{10}	-0.008	0.957	-0.008	1.017	0.000	0.102
β_{11}	-2.783	-2.198	-3.088	-2.246	-0.354	-0.027
β_{12}	1.965	2.436	2.089	2.696	0.028	0.309
β_{13}	-0.985	-0.051	-1.088	-0.051	-0.119	-0.001
β_{14}	-0.845	0.861	-0.935	0.915	-0.105	0.089
β_{15}	1.927	2.016	1.995	2.214	0.024	0.247
β_{16}	-1.009	-0.969	-1.113	-1.001	-0.127	-0.012
β_{17}	0.490	0.528	0.516	0.580	0.006	0.066

Table 3: Posterior Model Probability: First Simulation Exercise

										λ										
		-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	-0.9	1.000	1.000	1.000	1.000	1.000	0.999	0.852	0.235	0.105	0.037	0.377	0.959	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.8	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.856	0.237	0.103	0.027	0.343	0.971	1.000	1.000	1.000	1.000	1.000	1.000
	-0.7	1.000	1.000	1.000	1.000	1.000	0.999	0.861	0.240	0.102	0.021	0.315	0.979	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.6	1.000	1.000	1.000	1.000	1.000	0.999	0.866	0.246	0.102	0.016	0.293	0.984	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.5	1.000	1.000	1.000	1.000	1.000	0.999	0.872	0.254	0.104	0.014	0.281	0.987	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.4	1.000	1.000	1.000	1.000	1.000	0.999	0.879	0.264	0.106	0.012	0.277	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.3	1.000	1.000	1.000	1.000	1.000	0.999	0.887	0.276	0.110	0.011	0.284	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.2	1.000	1.000	1.000	1.000	1.000	0.999	0.895	0.292	0.120	0.011	0.303	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	-0.1	1.000	1.000	1.000	1.000	1.000	0.999	0.903	0.311	0.047	0.012	0.334	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ρ	0.0	1.000	1.000	1.000	1.000	1.000	0.999	0.912	0.333	0.135	0.014	0.380	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.1	1.000	1.000	1.000	1.000	1.000	0.999	0.921	0.358	0.153	0.018	0.438	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.2	1.000	1.000	1.000	1.000	1.000	0.999	0.930	0.387	0.171	0.026	0.506	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.3	1.000	1.000	1.000	1.000	1.000	0.999	0.938	0.420	0.186	0.039	0.583	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	1.000	1.000	1.000	1.000	1.000	1.000	0.946	0.457	0.206	0.063	0.643	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	1.000	1.000	1.000	1.000	1.000	1.000	0.954	0.498	0.227	0.100	0.719	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	1.000	1.000	1.000	1.000	1.000	1.000	0.962	0.543	0.254	0.155	0.772	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	1.000	1.000	1.000	1.000	1.000	1.000	0.969	0.592	0.286	0.224	0.817	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	1.000	1.000	1.000	1.000	1.000	1.000	0.975	0.643	0.321	0.298	0.853	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.9	1.000	1.000	1.000	1.000	1.000	1.000	0.981	0.694	0.353	0.365	0.880	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4: Posterior Model Probability: Second Simulation Exercise

	ρ									
Contiguity Criteria	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
7.50	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8.60	0.05	0.06	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00
9.70	0.06	0.10	0.09	0.05	0.01	0.00	0.00	0.00	0.00	0.00
10.80	0.05	0.09	0.10	0.08	0.04	0.02	0.01	0.00	0.00	0.00
11.90	0.09	0.26	0.52	0.75	0.89	0.95	0.98	0.99	1.00	1.00
12.90	0.05	0.09	0.10	0.08	0.05	0.02	0.01	0.00	0.00	0.00
14.00	0.05	0.06	0.05	0.02	0.01	0.00	0.00	0.00	0.00	0.00
15.10	0.04	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00
16.20	0.04	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17.30	0.03	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
18.40	0.04	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19.50	0.06	0.05	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
20.60	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
21.60	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22.70	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23.80	0.05	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
24.90	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
26.00	0.05	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Rook	0.09	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Queen	0.08	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 5: Descriptive Statistics

Variable		Mean	Std. Dev.	Min	Max
Gross State Product (US$)	overall	61,014.32	69,973.90	4,354.00	464,550.00
	between		69,484.36		
	within		12,769.31		
Public Capital (US$)	overall	25,036.66	27,780.40	2,627.12	140,217.30
	between		27,905.86		
	within		2,881.21		
Private Capital (US$)	overall	58,188.29	59,770.78	4,052.71	375,341.60
	between		58,697.23		
	within		13,957.80		
Employment (People)	overall	1,747.10	1,855.99	108.30	11,258.00
	between		1,847.81		
	within		311.95		
Unemployment (%)	overall	6.60	2.23	2.80	18.00
	between		1.27		
	within		1.84		

Source: Author's calculations

Table 6: Moran and Local Moran tests: Gross State Product (USA, 1970–1986)

	Year																
	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986
Moran test	0.774	0.757	0.742	0.743	0.734	0.744	0.754	0.765	0.754	0.745	0.744	0.731	0.719	0.694	0.681	0.673	0.661
State	Moran Local test																
ALABAMA	0.493	0.490	0.486	0.483	0.473	0.468	0.459	0.459	0.456	0.460	0.466	0.470	0.469	0.453	0.453	0.450	0.447
ARIZONA	0.722	0.690	0.669	0.661	0.652	0.668	0.668	0.649	0.627	0.591	0.576	0.583	0.606	0.574	0.541	0.504	0.478
ARKANSAS	0.647	0.643	0.635	0.639	0.627	0.622	0.622	0.621	0.617	0.616	0.609	0.603	0.596	0.590	0.589	0.590	0.589
CALIFORNIA	0.989	0.985	0.981	0.978	0.972	0.973	0.972	0.971	0.966	0.963	0.959	0.964	0.966	0.959	0.956	0.949	0.951
COLORADO	0.493	0.488	0.485	0.481	0.478	0.473	0.473	0.471	0.467	0.461	0.456	0.457	0.453	0.452	0.452	0.452	0.453
CONNECTICUT	0.326	0.339	0.345	0.347	0.348	0.367	0.375	0.379	0.386	0.387	0.383	0.377	0.365	0.338	0.326	0.313	0.296
DELAWARE	0.002	0.002	0.002	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.001
FLORIDA	0.266	0.256	0.244	0.232	0.230	0.238	0.239	0.236	0.230	0.222	0.220	0.217	0.216	0.200	0.189	0.180	0.173
GEORGIA	0.363	0.355	0.350	0.354	0.360	0.367	0.361	0.362	0.364	0.364	0.365	0.363	0.360	0.364	0.362	0.367	0.376
IDAHO	0.665	0.677	0.680	0.698	0.675	0.694	0.691	0.694	0.685	0.689	0.693	0.695	0.704	0.693	0.694	0.690	0.698
ILLINOIS	0.750	0.749	0.725	0.713	0.703	0.698	0.694	0.697	0.686	0.679	0.665	0.659	0.653	0.648	0.663	0.678	0.673
INDIANA	0.530	0.524	0.516	0.519	0.514	0.514	0.511	0.513	0.514	0.513	0.513	0.508	0.503	0.498	0.493	0.492	0.487
IOWA	0.490	0.475	0.476	0.489	0.487	0.509	0.503	0.500	0.492	0.485	0.473	0.485	0.450	0.406	0.411	0.400	0.397
KANSAS	0.526	0.528	0.533	0.543	0.540	0.532	0.538	0.555	0.566	0.559	0.570	0.569	0.565	0.573	0.578	0.581	0.577
KENTUCKY	0.460	0.459	0.460	0.459	0.458	0.458	0.459	0.460	0.462	0.462	0.465	0.463	0.464	0.469	0.468	0.472	0.474
LOUISIANA	0.369	0.365	0.367	0.377	0.387	0.396	0.387	0.393	0.400	0.407	0.414	0.419	0.417	0.412	0.413	0.414	0.420
MAINE	0.543	0.553	0.544	0.531	0.515	0.508	0.511	0.513	0.511	0.505	0.502	0.509	0.512	0.509	0.512	0.520	0.505
MARYLAND	0.353	0.354	0.361	0.371	0.375	0.376	0.384	0.394	0.403	0.408	0.407	0.407	0.402	0.400	0.402	0.400	0.401
MASSACHUSETTS	0.419	0.421	0.427	0.433	0.434	0.433	0.440	0.451	0.461	0.463	0.466	0.467	0.463	0.459	0.461	0.457	0.451
MICHIGAN	0.893	0.899	0.903	0.915	0.911	0.912	0.911	0.917	0.917	0.911	0.891	0.880	0.855	0.853	0.852	0.853	0.854
MINNESOTA	0.673	0.672	0.664	0.676	0.670	0.667	0.664	0.665	0.655	0.654	0.646	0.640	0.646	0.652	0.665	0.668	0.680
MISSISSIPPI	0.614	0.611	0.614	0.616	0.625	0.623	0.606	0.601	0.607	0.609	0.614	0.606	0.600	0.605	0.598	0.596	0.603
MISSOURI	0.247	0.245	0.246	0.249	0.258	0.263	0.259	0.257	0.261	0.264	0.276	0.278	0.277	0.269	0.266	0.266	0.266
MONTANA	0.721	0.732	0.733	0.737	0.736	0.735	0.754	0.761	0.755	0.753	0.736	0.722	0.707	0.696	0.705	0.706	0.698
NEBRASKA	0.433	0.425	0.420	0.415	0.421	0.420	0.417	0.416	0.414	0.416	0.425	0.432	0.436	0.431	0.431	0.432	0.426
NEVADA	0.394	0.398	0.394	0.405	0.380	0.386	0.384	0.387	0.374	0.371	0.376	0.383	0.389	0.373	0.373	0.367	0.366
NEW HAMPSHIRE	0.276	0.277	0.272	0.272	0.268	0.253	0.274	0.285	0.280	0.272	0.260	0.251	0.251	0.276	0.284	0.300	0.308
NEW JERSEY	0.594	0.595	0.598	0.603	0.595	0.595	0.593	0.595	0.603	0.607	0.610	0.607	0.610	0.613	0.608	0.606	0.612
NEW MEXICO	0.473	0.469	0.465	0.464	0.458	0.441	0.450	0.450	0.448	0.446	0.444	0.447	0.457	0.477	0.483	0.494	0.514
NEW YORK	0.151	0.146	0.148	0.139	0.151	0.158	0.171	0.188	0.191	0.199	0.221	0.204	0.202	0.211	0.213	0.225	0.226
NORTH CAROLINA	0.412	0.411	0.413	0.412	0.415	0.416	0.416	0.413	0.411	0.406	0.401	0.393	0.389	0.391	0.390	0.391	0.390
NORTH DAKOTA	0.998	0.997	0.997	0.994	0.995	0.994	0.995	0.996	0.994	0.994	0.994	0.992	0.993	0.994	0.994	0.994	0.994
OHIO	0.621	0.619	0.621	0.635	0.641	0.653	0.653	0.652	0.664	0.665	0.653	0.642	0.629	0.602	0.592	0.590	0.583
OKLAHOMA	0.490	0.489	0.486	0.482	0.481	0.480	0.480	0.480	0.479	0.478	0.480	0.480	0.486	0.484	0.481	0.479	0.474
OREGON	0.433	0.436	0.443	0.442	0.447	0.447	0.454	0.457	0.459	0.461	0.457	0.445	0.433	0.432	0.428	0.423	0.425
PENNSYLVANIA	0.563	0.556	0.560	0.544	0.548	0.525	0.562	0.571	0.560	0.551	0.556	0.542	0.529	0.563	0.578	0.594	0.599
RHODE ISLAND	0.434	0.431	0.424	0.424	0.416	0.396	0.403	0.401	0.399	0.394	0.392	0.397	0.399	0.428	0.435	0.445	0.464
SOUTH CAROLINA	0.483	0.480	0.477	0.474	0.475	0.474	0.476	0.476	0.478	0.478	0.479	0.480	0.482	0.481	0.480	0.481	0.479
SOUTH DAKOTA	0.875	0.869	0.868	0.862	0.865	0.864	0.881	0.876	0.883	0.883	0.885	0.872	0.866	0.863	0.850	0.842	0.842
TENNESSE	0.415	0.406	0.394	0.391	0.393	0.395	0.393	0.390	0.388	0.391	0.399	0.403	0.404	0.403	0.402	0.401	0.396
TEXAS	0.987	0.987	0.987	0.988	0.989	0.992	0.991	0.991	0.991	0.992	0.992	0.993	0.993	0.990	0.988	0.986	0.982
UTAH	0.711	0.701	0.695	0.693	0.697	0.705	0.710	0.709	0.711	0.718	0.719	0.715	0.712	0.703	0.690	0.682	0.679
VERMONT	0.851	0.849	0.848	0.850	0.857	0.861	0.870	0.876	0.868	0.869	0.874	0.881	0.886	0.888	0.891	0.886	0.884
VIRGINIA	0.842	0.850	0.854	0.864	0.880	0.892	0.891	0.893	0.895	0.897	0.906	0.909	0.915	0.919	0.914	0.913	0.917
WASHINGTON	0.563	0.554	0.550	0.555	0.570	0.579	0.582	0.583	0.599	0.607	0.608	0.599	0.591	0.601	0.596	0.601	0.612
WEST VIRGINIA	0.270	0.265	0.259	0.245	0.237	0.247	0.247	0.251	0.243	0.243	0.236	0.215	0.210	0.193	0.187	0.177	0.174
WISCONSIN	0.507	0.509	0.513	0.515	0.515	0.512	0.511	0.512	0.508	0.507	0.502	0.497	0.492	0.489	0.490	0.485	0.488
WYOMING	0.116	0.107	0.097	0.097	0.096	0.101	0.092	0.092	0.087	0.090	0.099	0.104	0.102	0.081	0.077	0.076	0.061

Table 7: Productivity and Public Investment: Contiguity Matrix Selection.

Distance		
Criteria	Posterior Model Probability	Average number of links
7.5	0.00	8.0
8.6	0.00	10.0
9.7	0.00	11.8
10.8	0.00	13.8
11.9	0.00	15.7
12.9	0.00	17.2
14.0	0.00	19.6
15.1	0.00	21.5
16.2	0.00	23.3
17.3	0.00	24.6
18.4	0.00	26.3
19.5	0.00	27.8
20.6	0.00	29.2
21.6	0.00	30.7
22.7	0.00	31.8
23.8	0.00	33.0
24.9	0.00	34.0
26.0	0.00	35.0
k near neighbour		
Criteria	Posterior Model Probability	Average number of links
1	0.00	1
2	0.00	2
3	0.00	3
4	0.00	4
5	0.00	5
Traditional		
Criteria	Posterior Model Probability	Average number of links
Rook	0.54	4.4
Queen	0.46	4.5

Table 8: Stationarity and Convergence diagnostics

Parameter	Heidelberger (1st Part/p-value)[a]	Heidelberger (2nd Part)[b]	Geweke[c]	Raftery[d]
Spatial Lag Disturbances	0.47	7.08e-04	0.17	1.00
Public Capital	0.48	4.98e-04	0.60	0.99
Private Capital	0.68	3.22e-04	-1.53	1.00
Employment	0.37	4.31e-04	0.28	1.00
Unemployment Rate	0.99	3.70e-05	-0.22	0.99
Spatial Lag Public Capital	0.71	1.04e-03	-0.61	0.96
Spatial Lag Private Capital	0.44	6.08e-04	0.84	0.97
Spatial Lag Employment	0.11	8.50e-04	-0.54	0.99
Spatial Lag Unemployment Rate	0.91	7.99e-05	-0.54	0.96
Standard Deviation	0.92	4.81e-05	0.34	1.04

Notes: [a] Null hypothesis is stationarity of the chain, [b] Half-width to mean ratio (threshold of 0.1), [c] Mean difference test z-score, [d] Dependence factor (threshold of 5)

Figure 1: Posterior Estimates: Spatial Lag Disturbances Parameter

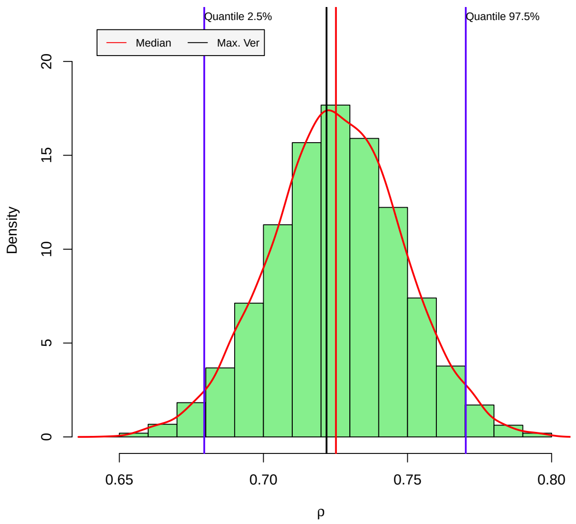

Figure 2: Posterior Estimates: Public Capital Parameter

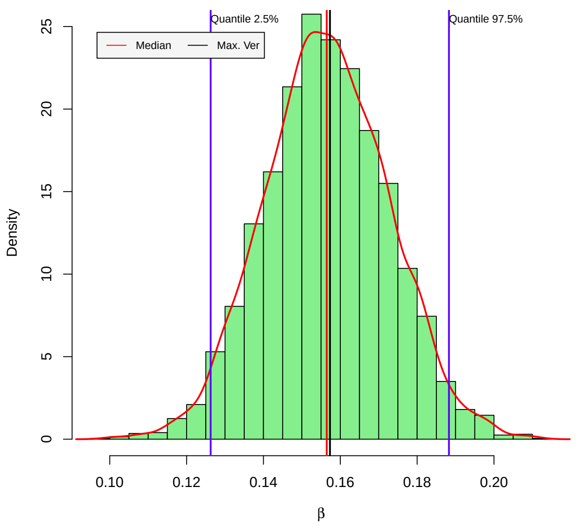

Figure 3: Posterior Estimates: Private Capital Parameter

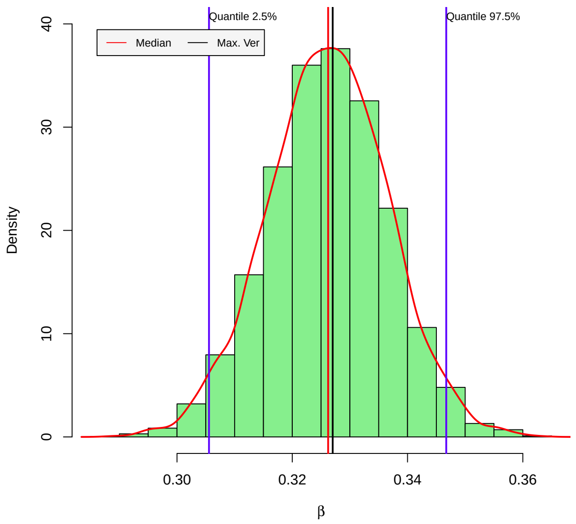

Figure 4: Posterior Estimates: Employment Parameter

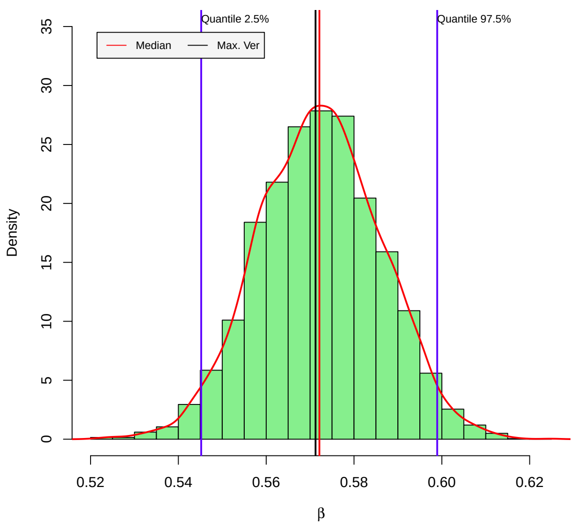

Figure 5: Posterior Estimates: Unemployment Parameter

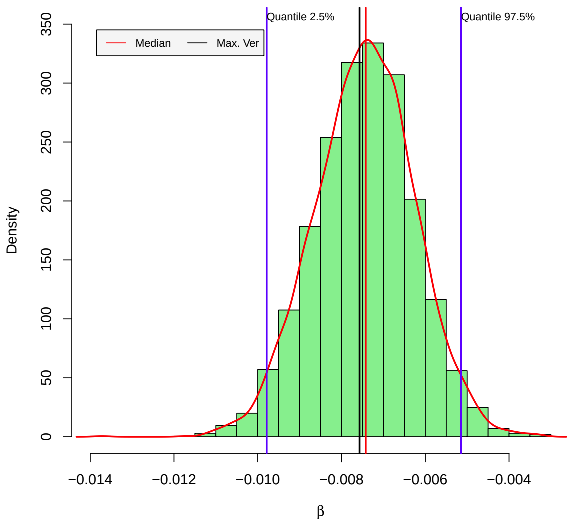

Figure 6: Posterior Estimates: Spatial Lag Public Capital Parameter

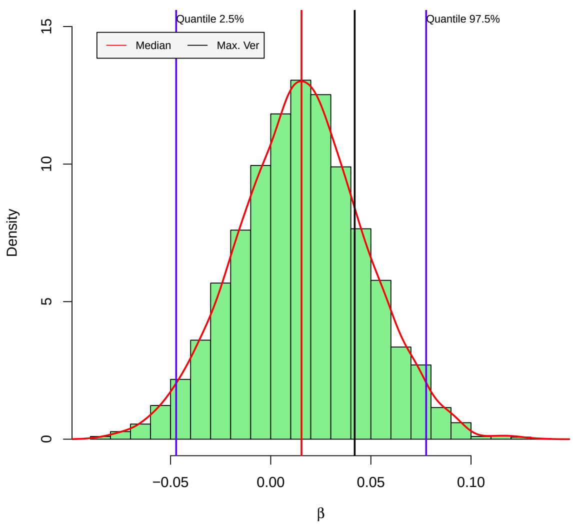

Figure 7: Posterior Estimates: Spatial Lag Private Capital Parameter

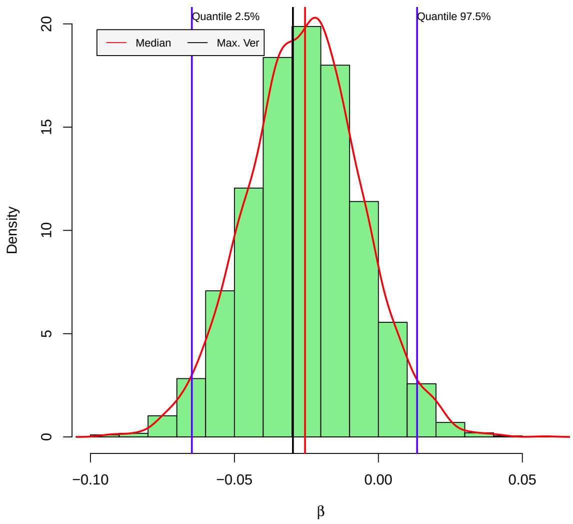

Figure 8: Posterior Estimates: Spatial Lag Employment Parameter

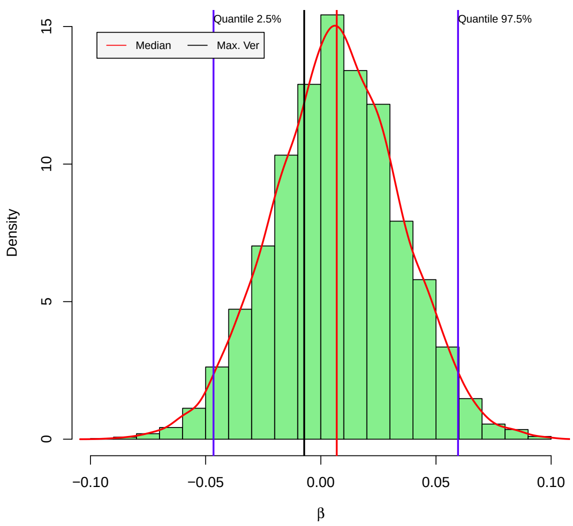

Figure 9: Posterior Estimates: Spatial Lag Unemployment Parameter

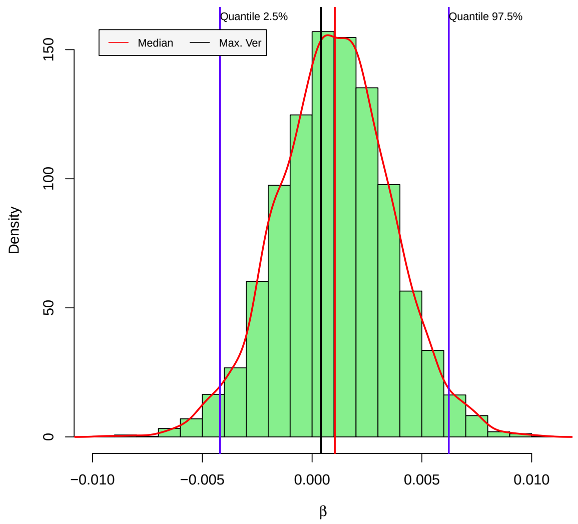

Figure 10: Posterior Estimates: Standard Deviation Parameter