Seemingly Unrelated Regression Spatial Autoregressive Bayesian Modeling on Heteroscedasticity Case

Adiatma¹*
¹Mathematics Department, Universitas Islam Negeri Alauddin Makassar, Indonesia

*Email: adiatma.rasyid@uin-alauddin.ac.id

Abstract. The phenomenon encountered occasionally on complications involving spatial data, is that there is a tendency of heteroscedasticity since every region has distinct characteristics. Thus, it requires the approach which is more appropriate with the problem by using the Bayesian method. Bayesian method on spatial autoregressive model to contend the heteroscedasticity by applying prior distribution on variance parameter of error. To detect heteroscedasticity, it is shown from several responses correlating with the predictors. The method able to estimate some responses is Seemingly Unrelated Regression (SUR). SUR is an econometrics model that used to be being utilized in solving some regression equations in which of them has their own parameter and appears to be uncorrelated. However, by correlation of error in differential equations, the correlation would occur among them. With the condition of the Bayesian SUR spatial autoregressive model, it is able to overcome heteroscedasticity cases from the vision of spatial. Further, the model involves four kinds of parameter priors’ distributions estimated by using the process of MCMC.

Keywords: Bayesian, Heteroscedasticity, MCMC, Spatial Autoregressive, SUR

1. Introduction
In spatial analysis especially spatial econometrics, the parameter estimation method that is often used by researchers is the Maximum Likelihood Estimation (MLE) method. One of the assumptions used in the MLE method is that the residuals are normally distributed with constant or identical variants for each observation (homoscedasticity). The phenomenon that often occurs in cases involving spatial/regional data is that there is a tendency for heteroscedasticity, namely the error variance condition is not identical because between regions with other regions are very varied and there are also data outliers that can lead to heteroscedasticity. According to Arbia [1], when a study deals with spatial data (especially with regional data), heteroscedasticity is a general phenomenon that is consistent with the nature of data collection. Thus a method approach is needed that is more in line with these conditions. The approach with the Bayesian method can be used for cases with spatial data because it can accommodate the presence of heteroscedasticity with the addition of priors as initial information. In the problem of heteroscedasticity, the Bayesian method uses a prior distribution on error variance parameters to accommodate error variances that are not constant between observations LeSage [2].

The development of the Bayesian method in the econometric model includes those who developed the Bayesian method in the spatial autoregressive model using priors to handle heteroscedasticity [2]. Lacombe [3] focuses more on the detailed breakdown of mathematical analysis needed in the application of Markov Chain Monte Carlo (MCMC) techniques. The Bayesian method has several advantages over the handling of spatial data mentioned in several previous studies, namely its use is more flexible, conceptually easier to understand, and has a high degree of accuracy. Besides having advantages, it also has weaknesses, namely the time needed to process data longer than the MLE method.

According to Setiawan and Kusrini [4], in addition to a single equation in the spatial econometrics model also involves several interrelated equations in the form of multiple equation models such as simultaneous equations and Seemingly Unrelated Regression (SUR) equations. SUR
was first developed by Zellner [5] which was a development of a linear regression model. SUR is a system of equations consisting of several regression equations, where each equation has a different response variable and it is possible to have a different set of predictor variables.

Setiawan and Suhartono [6], SUR is an econometric model that is widely used to solve several regression equations where each equation has its own parameters and it appears that each equation is not related. But between these equations, there is a connection with each other, namely by the correlation between errors in different equations. Therefore, the advantages of the SUR equation system are able to accommodate the correlation between errors of an equation with other equation errors.

Since the introduction of the SUR model, research related to the SUR model has been carried out by many researchers. Several studies have contributed to the development of SUR estimation problems namely, Kakwani [7], Guilkey and Schmidt [8], and Dwivedi and Srivastava [9]. Zellner [10] was also the first person to introduce SUR estimates with the Bayesian approach. Then Chib and Greenberg [11], Smith and Kohn [12], and Zellner and Ando [13] are researchers who have conducted Bayesian SUR analysis.

The development of estimation and testing of economic models with the Bayesian approach in many cases has been carried out. Most cases only applied spatial interactions have not discussed the study of spatial econometrics has not thoroughly explored yet in detail in terms of unobserved heterogeneity as in Anselin's [14] study of spatial econometrics. Therefore, in this study spatial autoregressive SUR Bayesian modeling was able to overcome the case of heteroscedasticity from a spatial/regional perspective.

2. Literature review
2.1 seemingly unrelated regression model
The first developed the SUR model which was a development of a linear regression model [5]. In general, the SUR model for $k$ is an equation where each equation consists of the $pj$ predictor variables can be written as follows:

$$
y_{1i} = \beta_{10} + \beta_{11}X_{1i1} + \cdots + \beta_{1p_1}X_{1ip_1} + \epsilon_{1i}$$
$$y_{2i} = \beta_{20} + \beta_{21}X_{2i1} + \cdots + \beta_{2p_2}X_{2ip_2} + \epsilon_{2i}$$
$$\vdots$$
$$y_{ki} = \beta_{k0} + \beta_{k1}X_{ki1} + \cdots + \beta_{kp_k}X_{kip_k} + \epsilon_{ki}$$

where $i = 1, 2, ..., n$ and $j = 1, 2, ..., k$

In SUR modeling systems, [5] has assumed that the matrix structure of variance-covariance is as follows:

$$\Omega = \text{E}(\epsilon_{i}^T \epsilon_{i}^T)$$

where $\Omega$ is an identity matrix measuring $(n \times n)$. Zellner [5] was also the first to introduce SUR estimates with the Bayesian approach. Then Chib and Greenberg [11], Smith and Kohn [12], and Zellner and Ando [13] are researchers who have conducted Bayesian SUR analysis.

The development of estimation and testing of economic models with the Bayesian approach in many cases has been carried out. Most cases only applied spatial interactions have not discussed the study of spatial econometrics has not thoroughly explored yet in detail in terms of unobserved heterogeneity as in Anselin's [14] study of spatial econometrics. Therefore, in this study spatial autoregressive SUR Bayesian modeling was able to overcome the case of heteroscedasticity from a spatial/regional perspective.
2.2 Spatial regression model

Anselin [14] developed a general model of spatial regression using cross-section data. This regression model is a spatial econometrics model which is the development of a simple regression model that has accommodated spatial autocorrelation phenomena. The general model of spatial regression is expressed in the form of the following equation:

\[ y = \rho W_1 y + X \beta + u \]

\[ y = (I_n - \rho W_1)^{-1} X \beta + (I_n - \rho W_1)^{-1} u \]

with

\[ u = \lambda W_2 u + \varepsilon \]

\[ u = (I_n - \lambda W_2)^{-1} \varepsilon \]

\[ \varepsilon \sim N(0, \sigma^2 I_n) \]

2.3 Spatial regression model in cross-sectional data

From the general form of spatial regression, several models can be formed in cross-sectional data, namely:

1) When \( \rho = 0 \) and \( \lambda = 0 \) then the equation becomes:

\[ y = X \beta + \varepsilon \]

This equation is called the classical regression model by ignoring the spatial effect.

2) When \( \rho \neq 0 \) and \( \lambda = 0 \) then the equation becomes:

\[ y = \rho W_1 y + X \beta + \varepsilon \]

\[ y = (I_n - \rho W_1)^{-1} X \beta + (I_n - \rho W_1)^{-1} \varepsilon \]

This equation is called the Spatial Lag Model or Spatial Autoregressive Model (SAR).

3) When \( \rho = 0 \) and \( \lambda \neq 0 \) then the equation becomes:

\[ y = X \beta + u \]

\[ y = X \beta + \lambda W_2 u \]

\[ y = X \beta + (I_n - \lambda W_2)^{-1} \varepsilon \]

This equation is called the Spatial Error Model (SEM).

4) When \( \rho \neq 0 \) and \( \lambda \neq 0 \) then the equation becomes:

\[ y = \rho W_1 y + X \beta + u \]

\[ y = \rho W_1 y + \lambda W_2 u + \varepsilon \]

\[ y = (I_n - \rho W_1)^{-1} X \beta + (I_n - \rho W_1)^{-1} u \]

This equation is called the General Spatial Model or Spatial Autocorrelation (SAC).

2.4 Bayesian spatial econometrics

In this model estimation, the researchers used an estimate. The estimate is known as the Markov Chain Monte Carlo (MCMC) estimation.

2.5 Bayesian spatial autoregressive

The general model of spatial autoregressive regression (SAR) for each 1st observation is:

\[ y_i = \rho \sum_{j=1}^n W_{ij} y_j + \sum_{r=1}^k \beta_r x_{ri} + \varepsilon_i \]

The equation can be presented in the form of matrix notation as follows:

\[ y = \rho Wy + X \beta + \varepsilon \]

\[ y = (I_n - \rho W)^{-1} X \beta + (I_n - \rho W)^{-1} \varepsilon \]

\[ \varepsilon \sim N(0, \sigma^2 I_n) \]

Furthermore, to apply the existence of heteroscedasticity in the SAR Bayesian model, the scalar factor \( V \) variance is added to the assumption of the residual error, where \( V \) is a diagonal matrix containing the parameters \( (v_1, v_2, ..., v_n) \) so its variant-covariant matrix is \( \sigma^2 V \) indicates that the error variant is not constant. The description is as follows:

\[ y = \rho Wy + X \beta + \varepsilon \]

\[ y = (I_n - \rho W)^{-1} X \beta + (I_n - \rho W)^{-1} \varepsilon \]

\[ \varepsilon \sim N(0, \sigma^2 V) \]

\[ v_{ii} = v_i , \text{ for } i=1,...,n \text{ and } v_{ij} = 0 , \text{ for } i\neq j \]
As mentioned in the previous section, the Bayesian method is based on the likelihood and prior distribution in forming a posterior distribution for estimation and conclusion.

3. Research method

This research is theoretical research that can be applied to identified cases of heteroscedasticity, especially in the case of econometrics. To achieve the objectives of this research, the steps taken in SUR Spatial Autoregressive Bayesian modeling in heteroscedasticity cases are as follows:

1. Creating a SUR spatial autoregressive model
2. Making the likelihood function
3. Determining the prior parameter distribution
4. Determining the posterior join distribution
5. Determining the distribution of the full conditional parameter
6. Estimating parameters by performing the MCMC process by taking samples from the full conditional distribution
7. Making conclusions

4. Result and explanation

Following are the stages of Bayesian modeling SUR Spatial Autoregressive in the case of heteroscedasticity:

4.1 Creating a SUR spatial autoregressive model

\[ y_{gi} = \rho_g \sum_{j=1}^{n} W_{ij} y_{gj} + \sum_{r=1}^{k} \beta_r X_{ri} + \varepsilon_{gi} \]  \hspace{1cm} (12)

4.2 Making the likelihood function

The likelihood function of the spatial heteroscedasticity model with errors is identically independent N(0, \( \sigma^2 \mathbf{V} \)):

\[
p(D | \beta, \sigma, \rho, \mathbf{V}) = (2\pi \sigma^2)^{-\frac{n}{2}} |\mathbf{I}_n - \rho \mathbf{W}| |\mathbf{V}|^{-\frac{1}{2}} \exp \left( -\frac{1}{2\sigma^2} (\mathbf{A} \mathbf{y} - \mathbf{X} \beta)' \mathbf{V}^{-1} (\mathbf{A} \mathbf{y} - \mathbf{X} \beta) \right)
\]

\[
= (2\pi \sigma^2)^{-\frac{n}{2}} |\mathbf{A}| |\mathbf{V}|^{-\frac{1}{2}} \exp \left( -\frac{1}{2\sigma^2} (\mathbf{e}' \mathbf{V}^{-1} \mathbf{e}) \right)
\]

\[
\propto \sigma^{-n} |\mathbf{A}| \prod_{i=1}^{n} |\mathbf{V}_{i}^{-\frac{1}{2}} \exp \left( -\sum_{i=1}^{n} \frac{\varepsilon_{i}^2}{2\sigma^2 \mathbf{V}_{i}} \right)
\]  \hspace{1cm} (13)

where the \( \varepsilon_{i} \) notation represents the \( i^{th} \) element of the vector \( \mathbf{e} = \mathbf{A} \mathbf{y} - \mathbf{X} \beta \), and \( \mathbf{A} = \mathbf{I}_n - \rho \mathbf{W} \)

4.3 Determining the prior parameter distribution

According to LeSage and Pace [15], the prior distribution used in this heteroscedasticity spatial model are:

\[
\pi(\beta) = \mathcal{N}(\mathbf{c}, \mathbf{T})
\]

\[
\pi(\sigma^2) = \mathcal{IG}(a, b)
\]

\[
\pi(\rho) \sim \mathcal{U}(\lambda_{min}, \lambda_{max})
\]

\[
\pi(r/v_i) \sim \mathcal{IID}(2), i = 1, 2, ..., n
\]

Hepple [16] states that it would be more beneficial to obtain a resolution of the Bayesian SAR model in the context of simplification offered by prior diffuse or non-informative. For uninformative prior for prior normal and gamma-inverses, use parameter values \( a, b = 0 \) and \( \mathbf{c} = 0 \), \( \mathbf{T} = \mathbf{I}_k \cdot 10^{12} \), and assume independence between priors so \( \pi(\beta, \sigma, \mathbf{V}, \rho) = \pi(\beta) \pi(\sigma) \pi(\mathbf{V}) \pi(\rho) \).

In detail the prior distribution for each parameter is as follows:

a) The prior distribution for \( \beta \) is \( \pi(\beta) = \mathcal{N}(\mathbf{c}, \mathbf{T}) \) which is a multivariate normal distribution with parameter \( \mathbf{e} \) which is the average vector, and \( \mathbf{T} \) which is the variance-covariance matrix.
\[
\pi(\beta) = \frac{1}{(2\pi)^{2n/2}} \exp \left( -\frac{1}{2} \left[ (\beta - c)' T^{-1}(\beta - c) \right] \right) \\
\propto \left| \exp \left( -\frac{1}{2} \left[ (\beta - c)' T^{-1}(\beta - c) \right] \right) \right|
\]
When using uninformative (diffuse) priors with parameters \( c = 0 \), \( \beta = T^{-1} \cdot b \) then, \( \pi(\beta) \propto \) constant.

b) Prior distribution for \( \sigma^2 \) is \( \pi(\sigma^2) = IG(a,b) \) where \( a \) and \( b \) is a parameter of the gamma-inverse distribution
\[
\pi(\sigma^2) = \frac{b^a}{\Gamma(a)} (\sigma^2)^{-(a+1)} \exp(-b/\sigma^2) \\
\propto (\sigma^2)^{-(a+1)} \exp(-b/\sigma^2)
\]
When using uninformative (diffuse) priors with parameters \( a = b = 0 \) then, \( \pi(\sigma) \propto \frac{1}{\sigma} \), \( 0 < \sigma < \infty \)

c) The prior distribution for \( V \) where \( v_i \) is independent identical, with \( \left( \frac{r}{2} \right) \) Chi-square distribution with \( r \) is the degree of freedom.
\[
\pi \left( \frac{r}{2} \right) \sim iid \chi^2(r), i = 1,2, ..., n \\
\pi \left( \frac{r}{2} \right) = \frac{(r)^{r/2}}{2^{r/2} \Gamma(r/2)} V_i^{r/2} \exp \left( -\frac{r}{2V_i} \right)
\]
or
\[
\pi(V) = \left( \frac{r}{2} \right)^{nr/2} [\Gamma \left( \frac{r}{2} \right)]^{-n} \prod_{i=1}^{n} V_i^{r/2} \exp \left( -\frac{r}{2V_i} \right) \\
\propto \prod_{i=1}^{n} V_i^{r/2} \exp \left( -\frac{r}{2V_i} \right)
\]
The prior distribution for \( \rho \) is \( \pi(\rho) \sim U(\lambda_{min}^{-1}, \lambda_{max}^{-1}) \), that is a uniform distribution with parameters \( \lambda_{min} \) and \( \lambda_{max} \) which is the minimum and maximum value of the eigenvalue of the spatial weighting matrix \( (W) \), so that \( \pi(\rho) \propto \) constant.

4.4 Determining the joint posterior distribution
Based on the Bayes Theorem with independent assumptions between priors
\[
\pi(\beta, \sigma, V, \rho) = \pi(\beta) \pi(\sigma) \pi(V) \pi(\rho)
\]
then the posterior joint distribution uses prior uninformative (diffuse) to be as follows:
\[
p(\beta, \sigma, V, \rho | D) \propto p(D | \beta, \sigma, V, \rho) \pi(\beta) \pi(\sigma) \pi(V) \pi(\rho)
\]
\[
\propto \sigma^{-n} |A| \prod_{i=1}^{n} V_i^{-\frac{1}{2}} \exp \left( -\sum_{i=1}^{n} e_i^2 \right) \cdot \frac{1}{\sigma} \cdot \prod_{i=1}^{n} V_i^{r+2} \exp \left( -\frac{r}{2V_i} \right) \\
\propto \sigma^{-n} |A| \prod_{i=1}^{n} V_i^{r+2} \exp \left( -\Sigma_{i=1}^{n} \frac{\sigma^2 e_i^2 + r}{2\sigma^2 V_i} \right)
\]
where the \( e_i \) notation denotes the \( i \)th element of the vector \( e = Ay - X\beta \) and \( A = I_n - \rho W \).

Whereas if using informative priors, then:
\[
p(\beta, \sigma^2, V, \rho | D) \propto p(D | \beta, \sigma^2, \rho \pi(\beta) \pi(\sigma^2) \pi(V) \pi(\rho)
\]
\[
\propto \sigma^{-n} |A| \prod_{i=1}^{n} V_i^{-\frac{1}{2}} \exp \left( -\frac{1}{2\sigma^2} (Ay - X\beta)'V^{-1}(Ay - X\beta) \right) \\
\times \exp \left( -\frac{1}{2} \left[ (\beta - c)' T^{-1}(\beta - c) \right] \right) \times (\sigma^2)^{-(a+1)} \exp \left( -\frac{b}{\sigma^2} \right) \\
\times \prod_{i=1}^{n} V_i^{r+2} \exp \left( -\frac{r}{2V_i} \right)
\]
where $A = I_n - \rho W$.

### 4.5 Determining the distribution of the full conditional parameter

From the posterior joint equation above, we can find the full conditional distribution for each parameter in the model. After the data is generated from the full conditional distribution of all the parameters in sequence, it becomes a set of sample parameters that will be used to construct the posterior joint distribution so that the inference of each parameter is obtained (such as the mean, standard deviation, etc.).

#### 4.5.1 Full conditional posterior distribution for $\beta$

The full conditional posterior distribution can be obtained by simply selecting the portion of the posterior joint distribution which contains the relevant parameters $[3]$ so that the full conditional posterior distribution for $\beta$ is:

$$p(\beta|\sigma^2, \rho, V) \propto \exp \left( -\frac{1}{2\sigma^2} \left[ (Ay - X\beta)'V^{-1}(Ay - X\beta) + (\beta - c)'\sigma^2T^{-1}(\beta - c) \right] \right)$$

(20)

Thus the full conditional posterior distribution for $\beta$ is a multivariate normal distribution with the average vector $c^*$ and the following $T^*$ covariance variance:

$$p(\beta|\rho, \sigma^2, V) \sim N(c^*, T^*)$$

(21)

$$c^* = (X'V^{-1}X + \sigma^2T^{-1})^{-1}(X'V^{-1}Ay + \sigma^2T^{-1}c)$$

$$T^* = \sigma^2(X'V^{-1}X + \sigma^2T^{-1})^{-1}$$

where $A = I_n - \rho W$.

#### 4.5.2 Full conditional posterior distribution for $\sigma^2$

From the posterior joint distribution, we get a full conditional posterior distribution for $\sigma^2$ is:

$$p(\sigma^2|\beta, \rho, V) \propto (\sigma^2)^{-\left(\frac{1}{2} + \frac{n}{2} + 1\right)} \exp \left( -\frac{1}{2\sigma^2} \left[ (e'V^{-1}e) + 2b \right] \right)$$

(22)

with $a$ and $b$ are parameters of prior $\text{IG}(a, b)$.

From this distribution, it can be stated that the full conditional distribution is distributed chi-square ($\chi^2$) with df $(n + 2a)$. Specifically, it can be stated from the conditional distribution of each $\nu_i$ as follows:

$$p\left(\frac{(e'V^{-1}e)+(2b)}{\sigma^2}\mid \beta, V, \rho\right) \sim \chi^2_{(n^*)}$$

(23)

where $n^* = n + 2a$, and $e = Ay - X\beta$, with $A = I_n - \rho W$.

#### 4.5.3 Full conditional posterior distribution for $V$

From the posterior joint distribution we get a full conditional posterior distribution for $V$ is:

$$p(V|\beta, \sigma^2, \rho) \propto \prod_{i=1}^{n} V_i^{-\frac{r+3}{2}} \exp \left( -\sum_{i=1}^{n} \frac{\sigma^{-2}e_i^2 + r}{2V_i} \right)$$

(24)

The full conditional posterior distribution for $V$ is Chi-square ($\chi^2$) distribution with df $r + 1$ has been shown by Geweke [17]. Specifically, it can be stated from the conditional distribution of each $\nu_i$ as follows:

$$p\left(\frac{\sigma^{-2}e_i^2 + r}{\nu_i}\mid \beta, \sigma^2, V_{-i}\right) \sim \chi^2_{(r+1)}$$

(25)

where $V_i = (V_1, \ldots, V_{i-1}, V_{i+1}, \ldots, V_n)$ for each $i$. Thus we take samples from each scalar variant to the other variance scalars. The $e_i$ notation denotes the $i_{th}$ element of the vector $e = Ay - X\beta$, and $A = I_n - \rho W$. 
4.5.4 **Full conditional posterior distribution for \( \rho \).** Based on the joint posterior distribution, the fully conditional posterior distribution obtained for \( \rho \) is:

\[
p(\rho | \beta, \sigma^2, V) \propto |\ln - \rho W| \exp\left(\frac{(\mathbf{e}'V^{-1}\mathbf{e})}{2\sigma^2}\right)
\]

\[
\propto |\Lambda| \exp\left(-\frac{1}{2\sigma^2}[(\mathbf{Ay} - \mathbf{X}\beta)'V^{-1}(\mathbf{Ay} - \mathbf{X}\beta)]\right)
\]

where \( \mathbf{e} = \mathbf{Ay} - \mathbf{X}\beta \), and \( \Lambda = \mathbf{I}_n - \rho \mathbf{W} \).

To get a sampling from the parameter \( \rho \), the Metropolis-Hastings algorithm will be used because the full conditional form of the posterior distribution is unknown (nonstandard).

The Metropolis-Hastings algorithm requires the distribution of proposals (candidates) for the parameter \( \rho \) called \( \rho^* \). [15] have used standard normal distribution as a proposal distribution with a random walk procedure to produce candidate values for \( \rho \). The procedure involves the value of \( \rho \) is \( \rho^t \), a random number from the generation of standard normal distribution, and tuning parameter \( c \) as shown in the equation below:

\[
\rho^* = \rho^t + c \cdot N(0,1)
\]

Then this candidate value \( \rho^* \) and the current value \( \rho^t \) are evaluated to calculate the acceptance opportunities using the equation below:

\[
\alpha_t(\rho^t, \rho^*) = \min\left(1, \frac{p(\rho^t | \beta, \sigma^2, V)}{p(\rho^* | \beta, \sigma^2, V)}\right)
\]

Then accept of \( \rho^{t+1} = \rho^* \) with the probability \( \alpha_t(\rho^t, \rho^*) \), if not then \( \rho^{t+1} = \rho^t \), which means it remains with the values at this time.

4.5.5 **Markov chain monte carlo estimation (MCMC).** The Bayesian method with the MCMC approach is a method of generating sample data sequentially from a series of conditional distributions of all parameters in the model to produce a set of estimates converging near the original distribution of the posterior joint of the parameter model. [15] state that this approach is used in estimating spatial regression parameters that are related to variants that are not constant (varied) in each of their observations. In [18] adds that this method is a numerical method that parses complex estimation problems to be simpler based on the conditional distribution of each parameter in the model.

In this study, the Gibbs Sampling method is used to estimate the parameters \( \beta, \sigma^2 \), and \( V \) because the conditional distribution of the parameters \( \beta, \sigma^2 \), and \( V \) has a general form known (standard), while the parameter \( \rho \) will be solved by the Metropolis-Hastings method algorithm (M-H) because the form of distribution of parameters is not commonly known.

5. **Conclusion**

Based on the results of SUR Spatial Autoregressive Bayesian modeling, it can be concluded that in overcoming cases of heteroscedasticity four prior parameters are mutually independent. Then the parameters in the model are estimated using MCMC namely Gibbs Sampling and Metropolis-Hasting.

**References**

[1] Arbia, G. (2006), Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence, Springer, Berlin.

[2] LeSage, J.P. (1997), “Bayesian Estimation of Spatial Autoregressive Models”, International Regional Science Review, Vol. 20, No.1&2, page. 113–129.

[3] Lacombe, D.J. (2008), An Introduction to Bayesian Inference in Spatial Econometrics, West Virginia University, United States.

[4] Setiawan and Kusrini, D.E. (2010), Ekonometrika, CV. Andi Yogyakarta, Yogyakarta.

[5] Zellner, A. (1962), “An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests for Aggregation Bias”, Journal of the American Statistical Association, Vol.57, page.348-368.
[6] Setiawan dan Suhartono, (2014), “S-GSTAR-SUR Model for Seasonal Spatio Temporal Data Forecasting”, Proceedings of the 10th IMT-GT ICMSA 2014, Kuala Terengganu, Malaysia.

[7] Kakwani, N.C. (1967), “The unbiasedness of Zellner’s seemingly unrelated regression equations estimators”, Journal of the American Statistical Association, Vol. 62, No. 317, hal. 141-142.

[8] Guilkey, D.K. and Schmidt, P. (1973), “Estimation of seemingly unrelated regressions with vector autoregressive errors”. Journal of the American Statistical Association, Vol.68, No.343, page.642-647.

[9] Dwivedi, T.D. dan Srivastava, V.K. (1978), “Optimality of least squares in the seemingly unrelated regression equation model”, Journal of Econometrics, Vol. 7, page. 391-395.

[10] Zellner, A. (1971), An Introduction to Bayesian Inference in Econometrics, John Wiley & Sons, New York.

[11] Chib, S. and Greenberg, E. (1995), “Hierarchical Analysis of SUR Models with Extensions to Correlated Serial Errors and Time-Varying Parameter Models”, Journal of Econometrics, Vol. 68, page. 339-360.

[12] Smith, M. dan Kohn, R. (2000), Nonparametric Seemingly Unrelated Regression. Journal of Econometrics, Vol. 98, page. 257-282.

[13] Zellner, A. dan Ando, T. (2010), “Hierarchical Bayesian Analysis of the Seemingly Unrelated Regression and Simultaneous Equations Models Using a Combination of Direct Monte Carlo and Importance Sampling Techniques”, Journal of International Society for Bayesian Analysis, Vol.5, page.6.

[14] Anselin, L., (1988), Spatial Econometrics: Methods and Models, Kluwer Academic, Dordrecht.

[15] LeSage, J.P. dan Pace, R.K. (2009), Introduction to Spatial Econometrics, CRC Press (Taylor and Francis Group), Boca Raton [FL], London and New York.

[16] Hepple, L.W. (1999), “Bayesian Model Choice in Spatial Econometrics”, Proceedings of LSU Spatial Econometrics Conference, Baton Rouge, 7-11, November 2003.

[17] Geweke, J. (1993), “Bayesian Treatment of the Independent Student-t Linear Model,” Journal of Applied Econometrics, Vol. 8, page. 19–40.

[18] Priyono, (2011), Pendekatan Bayesian Pada Model Spasial Ekonometrika: Pemodelan Migrasi Penduduk di Provinsi Jawa Barat, Tesis Magister, Institut Teknologi Sepuluh Nopember, Surabaya.

Acknowledgment
The author would like to thank Dr. Aswi for her valuable suggestions.