

BAYESIAN SPATIAL HETEROSCEDASTIC AUTOREGRESSIVE MODEL (BSHAR)

Olayede L.

Department of Statistics, University of Ilorin.

Abstract

The study investigates the impact of spatial heteroscedastic autoregressive (SHAR) model on parameter estimations by adopting white (1989) heteroscedastic covariance matrix in Bayesian paradigm. Coordinate references of 94 cities in Nigeria were collected, the study obtained the minimum and maximum values, and these were used to generate coordinate references for the entire simulated data using uniform distribution. The study computed row-standardized spatial contiguity weights with the coordinates' reference matrix, obtained from each sample generated with/without the spatial autocorrelation term. Bias and mean squares error (MSE) criteria were adopted to evaluate the performances of the SHAR model. The study found out that asymptotically, there exist consistency and efficiency of SHAR model, the modified heteroscedastic covariance matrix (HCS) also behaved well and followed the patterns of other HC.

Keywords: SHAR, Heteroscedastic, Spatial and Bayesian inference.

1.0 Introduction

It has been widely observed that observations obtained from varied sources of different coordinate references may depict pattern with contiguity of space. In many regression analyses, more often than none, researchers have overlooked the spatial configuration of data with which it resulted into residuals that vary over space and eventually led to the problems of spatial autocorrelation. Henceforth, affect the inferences of the model parameter estimates. The standard errors also become large and the test of the hypothesis that uses the standard error become inefficient.

The regional scientist often obtained their regional data samples with reference to points in space [1]. Adopting the use of ordinary least squares to analyse such data sample would exhibit spatial autocorrelation. The first order spatial contiguity matrix were adopted in regression analysis of spatial data by [1]. It was described that first order spatial contiguity matrix provides a unified approach, which incorporate the spatial configuration information with respect to the coordinate points in space (latitude and longitude) through which data were obtained. Conventional specification tests are inefficient when there is presence of spatial autocorrelation [2]. A Bayesian approach to the spatial modeling were provided from the works of [3] and [4].

The spatial models were examined in the premise of econometrics perspective [5] whereas [1] examined it in a broader sense with regards to regional sciences. The proponents of spatial expansion model were [6] and [7]. They incorporated spatial influences into the regression relationships. The regional data researcher and analysts should incorporate information in the spatial configuration of the sample data in their study [1] and added that spatial contiguity matrix was adopted to create spatial prior in lieu of the conventional Minnesota prior which had been the famous prior in regional time series forecasting. The contiguity matrix of the sample data where study made use of binary dependent variable that was truncated or censored at some point was extended [1].

Spatial prior was more appropriate for modeling regional time series variables based on the concept of first order spatial contiguity [8]. The Minnesota prior that accommodated regional time series model appropriately was modified [9] and [10]. Bayesian variants of spatial econometrics have been gaining ground recently in the literature, which was as a result of the inadequacy of the frequentist approach that failed to estimate spatial Pooled and Tobin etc, due to their cumbersome nature [11]. Heteroscedastic disturbances in a model estimated using the context of maximum likelihood estimation was proposed by [5], by violating such assumption, there is need for specifying a functional form that can accommodate the non-constant variance over space. Markov Chain Monte Carlo (MCMC) approach is capable of detecting and down-weighting the aberrant observation in the process. The spatial autoregressive limited dependent variable using grid Gibbs sampler would accommodate data embedded with spatial autocorrelation [12] and pointed out that, complication would arise when there exist heteroscedasticity that lead to inconsistency in estimates if ignored during estimation. SAR and SEM probit models with spatial heteroscedasticity that depend on the EM algorithm were examined [13]. Heteroscedastic covariance matrix was injected into spatial autoregressive model and observed asymptotic behaviour via the finite properties of consistency and efficiency [14] and [15]. The Griddy Gibbs Sampler outperformed other sampling algorithms with the intensity of spatial interaction. It was observed that the SAR model spatial autocorrelation was weakly identified [16].

Corresponding Author: Olayede L. Email: olayede.l@unilorin.edu.ng. Tel: +2348064236198

Journal of the Nigerian Association of Mathematical Physics Volume 46, (May, 2018 Issue), 107–114

Generalized Linear Model estimators for the spatial autoregressive model parameter with unknown heteroscedasticity was introduced in [17]. The study established the consistency and asymptotic normality of Instrumental Variable estimators for the spatial autoregressive model parameters.

Maximum likelihood, restricted maximum likelihood, and two instrumental variables were compared when there are correlated regressors in Spatial Autoregressive model and spatial Durbin model in attempt to varying signal-to-noise, spatial dependence, and weight matrix specifications [18]. It was observed that REML outperformed Maximum Likelihood estimation and the instrumental variable performance was affected by the disturbances. The study therefore seek to investigate the impact of spatial autocorrelation and heteroscedastic covariance matrix.

2.0 Methodical Designs

Let $y = \lambda Wy + X\beta + u$ be a spatial autoregressive model where y contains $n \times 1$ dimensional vector of response variable and X denotes $n \times k$ matrix of regressors. W is an exogenous $n \times n$ spatial weight matrix usually with first order contiguity relation, a row-standardized spatial weight matrix, $u \sim N(0, \sigma_u^2 \Omega)$, β is $n \times 1$ vector of model parameter estimates and λ denotes the spatial autocorrelation term.

Definition 1: the contiguity spatial-weight matrix W with neighbour or distance denoted by $\bar{w}_{ij}$ which equal one iff i -th and j -th units are neighbour. Therefore $w_{ij} = \frac{\bar{w}_{ij}}{\sum_{j=1}^n \bar{w}_{ij}}$

Assumption 1: the w_{ij} are identically independently distributed with zero mean and variance $\sigma_w^2 < \infty$ and is heteroscedastic. Therefore, $\sigma_i^2 \neq \sigma_j^2$ being the diagonal element while (1985) heteroscedastic covariance matrix.

Assumption 2: the contiguity spatial weight matrix W_n are uniformly bounded in absolute value in row and column.

Assumption 3: the diagonal elements w_{ii} of the contiguity of spatial weight matrix W_n are zero with off-diagonal elements w_{ij} such that k_n a finite constant has $\sum_{j=1}^n w_{ij} \leq k_n$ and or $\sum_{i=1}^n w_{ij} \leq k_n \forall i = j = 1, \dots, n; \forall n \geq 1$.

Assumption 4: a non-singular contiguity spatial weight matrix $(I - \theta W_n)$ with $\sum_{j=1}^n |w_{ij}|$ and $\sum_{i=1}^n |w_{ij}|$ are uniformly bounded such that $k_n < \infty \forall i = j = 1, \dots, n; \forall n \geq 1$ with w_{ii} and w_{ij} are diagonal and off-diagonal elements of $(I - \theta W_n)$ respectively $\forall \theta < 1$.

Assumption 5: the regressors' matrix X_n is of full rank and uniformly bounded and non-singular. Thus there exist $\lim_{n \rightarrow \infty} X_n' X_n$.

The white Heteroscedastic covariance matrix Σ have been proposed which were constructed by injecting an estimate of type $\hat{\Omega} = \text{diag}(w_1, w_2, \dots, w_n)$ into likelihood equation. The six type of the error structure are shown below:

$$\text{Case: } w_i = \sigma^2 \quad (2.1)$$

$$\text{HC0: } w_i = \hat{\sigma}_i^2 \quad (2.2)$$

$$\text{HC1: } w_i = \frac{\hat{\sigma}_i^2}{n-k} \quad (2.3)$$

$$\text{HC2: } w_i = \frac{\hat{\sigma}_i^2}{1-k} \quad (2.4)$$

$$\text{HC3: } w_i = \frac{\hat{\sigma}_i^2}{(1-k)^2} \quad (2.5)$$

$$\text{HC4: } w_i = \frac{\hat{\sigma}_i^2}{(1-k)^n} \quad (2.6a)$$

$$\text{HC5: } w_i = \frac{\hat{\sigma}_i^2}{\bar{h}(1-k)} \quad (2.6b)$$

where $\hat{h} = H_n$ the diagonal elements of the hat matrix, $\bar{h}$ is their mean and $\bar{h} = \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{1-k} \right)$ the first expression above yields the

standard estimator $\hat{\Sigma}_{\text{case}}$ for equal/or homoscedastic errors. The others produce different kinds of HC estimators. The HC0 estimator was proposed by [14], while the estimators HC1, HC2 and HC3 were proposed by [15] for improving the performance in small samples.

GLS Estimation of the SHAR model

$$\text{The SHAR model} \quad y = \lambda Wy + X\beta + u \quad (2.7)$$

$$u \sim N(0, \sigma_u^2 \Omega) \quad (2.8)$$

It can be written as

$$(I - \lambda W)y = X\beta + u \quad (2.9)$$

$$Ay = X\beta + u \quad (2.10)$$

With

$A = (I_n - \lambda W)$, $u \sim N(0, \sigma_u^2 \Omega)$, the model is expressed in likelihood form as:

$$L(\beta, \lambda, \sigma_u^2, \Omega) = (2\pi\sigma_u^2)^{-n/2} |\Omega|^{-1/2} \exp \left\{ -\frac{1}{2\sigma_u^2} (y - A\beta)' \Omega^{-1} (y - A\beta) \right\} \quad (2.11)$$

Obtaining the generalized least squares estimation $b = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} A y$
 $= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} (I_n - A W) y$ (2.12)

$$b = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y - (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} A W y \quad (2.13)$$

$$= b_{OLS} - b_{GLS} \quad (2.14)$$

$$\text{where } b_{OLS} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y \text{ and } b_{GLS} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} A W y \quad (2.15)$$

obtain the two model parameters using generalized least squares, the residuals can be obtained through

$$e_{OLS} = y - X b_{OLS} \quad (2.17)$$

$$e_{GLS} = A W y - X b_{GLS} \quad (2.18)$$

Then the GLS estimate of σ^2 can be expressed as $s^2 = (1/n) \cdot (e_{OLS} - e_{GLS})' (e_{OLS} - e_{GLS})$.

The estimation steps are as follows:

- i. Regress y on X to obtain b_{OLS}
- ii. Compute the residual $e_{OLS} = y - X b_{OLS}$
- iii. Regress $W y$ on X to obtain b_{GLS}
- iv. Compute the residual $e_{GLS} = A W y - X b_{GLS}$
- v. Obtain the λ by maximizing the likelihood function, but in case of the theoretical study λ - values are assumed.
- vi. Compute $b = b_{OLS} - b_{GLS}$ and $s^2 = (1/n) \cdot (e_{OLS} - e_{GLS})' (e_{OLS} - e_{GLS})$

The likelihood function of θ , where $\theta = (\beta, \lambda)$ given sample vectors $X_1, X_2 = (1, 2, \dots, n)'$ and $y = (y_1, y_2, \dots, y_n)'$ is expressed as

$$L(\beta, \lambda, X, y, W) = (2\pi\sigma^2)^{-n/2} \prod_{i=1}^n \Omega^{-1/2} \mu_n - \rho W \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (A y_i - X \beta)' \Omega^{-1} (A y_i - X \beta) \right\} \quad (2.19)$$

where Ω is defined as above in eq 1.... λ being the covariance matrix with w_{ij} as diagonal elements.

Posterior density is derived by conjugating the likelihood function in (2.19) with multivariate normal distribution and inverse gamma distribution. The study obtained Marginal posterior density by integrating the joint posterior distribution with respect to each nuisance parameter.

The study assumed a prior density $\pi(\beta, \lambda, \sigma) = \pi(\beta) \pi(\lambda) \pi(\sigma)$. Thus multivariate normal prior is considered for β , while inverse gamma is chosen for σ and a uniform distribution is deemed for λ such that:

$$\pi(\beta) \propto \frac{\exp(-\frac{1}{2}(\beta - \mu)' \Psi (\beta - \mu))}{(2\pi)^{p/2} |\Psi|^{1/2}} \exp \left\{ -\frac{1}{2}(\beta - \mu)' \Psi (\beta - \mu) \right\}, \beta > 0 \quad (2.20)$$

where Ψ is the precision matrix which is the inverse of the covariance matrix of β .

$$\pi(\sigma^2) \propto (\sigma^2)^{-(a_1+1)} \exp(-b_1/\sigma^2), \sigma^2 > 0 \quad (2.21)$$

$$\pi(\lambda) \propto U(\lambda_{min}, \lambda_{max}) \quad (2.22)$$

$$\Omega = (X' X)^{-1} X' \Phi X (X' X)^{-1}$$

The posterior distribution of $\theta = (\beta, \lambda, \sigma)$, considering independence among the parameters is given by:

$$\pi(\beta, \lambda, \sigma | X, y, W) \propto \frac{(2\pi)^{n/2}}{(2\pi\sigma^2)^{n/2}} (\sigma^2)^{-(a_1+1-n/2)} \exp \left\{ -\frac{1}{2}(\beta - \mu)' \Psi (\beta - \mu) \right\} \prod_{i=1}^n \Omega^{-1/2} \mu_n - \rho W \exp \left\{ -\frac{1}{2\sigma^2} (b_1 + \frac{1}{2} \sum_{i=1}^n (A y_i - X \beta)' \Omega^{-1} (A y_i - X \beta)) \right\} \quad (2.23)$$

where a_1, b_1 are the hyper-parameters for the inverse-gamma distribution. The study excludes Hyper-parameters for β -parameters since its presence may lead to problems which may influence the inferences. The posterior $\pi(\beta, \lambda, \sigma | X, y, W)$ is integrated with respect to σ , when σ stands as nuisance parameter.

Recall

$$\int_0^\infty x^{-(p+1)} \exp\left(-\frac{x}{q}\right) dx = 1 \quad (2.24)$$

Re-Scaling the joint posterior density with inverted gamma $q, \dots$

$$\frac{(2\pi)^{n/2}}{(2\pi\sigma^2)^{n/2}} (\sigma^2)^{-(a_1+1-n/2)} \exp \left\{ -\frac{1}{2}(\beta - \mu)' \Psi (\beta - \mu) \right\} \prod_{i=1}^n \Omega^{-1/2} \mu_n - \rho W \exp \left\{ -\frac{1}{2\sigma^2} (b_1 + \frac{1}{2} \sum_{i=1}^n w^{-\delta} (y - X \beta)^2) \right\} \quad (2.25)$$

By setting $p = a_1 - 1 - n/2$

$$q = \exp \left\{ -\frac{1}{2}(\beta - \mu)' \Psi (\beta - \mu) \right\} \prod_{i=1}^n \Omega^{-1/2} \mu_n - \rho W \exp \left\{ -\frac{1}{2\sigma^2} (b_1 + \frac{1}{2} \sum_{i=1}^n \Omega^{-1} (y - X \beta)^2) \right\} \quad (2.26)$$

where $\sigma = \sigma^2$

$$\int_0^\infty x^{-(p+1)} \exp\left(-\frac{x}{q}\right) dx = \frac{\Gamma(p)}{q^p} \quad (2.27)$$

Let p is a constant

$$\text{Therefore,} \quad \int_0^\infty x^{-(p+1)} \exp\left(-\frac{x}{q}\right) dx \propto q^{-p} \quad (2.28)$$

Thus the marginal posterior distribution for β is obtained by substitution into q^{-p}

This yields the following full marginal density of the parameters β and σ :

$$\pi(\beta | \delta, X, y, W) \propto \exp \left\{ -\frac{1}{2}(\beta - \mu)' \Psi (\beta - \mu) \right\} \prod_{i=1}^n \Omega^{-1/2} \mu_n - \rho W \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \Omega^{-1} (y - X \beta)^2 \right\}^{-(a_1+1-n/2)} \quad (2.29)$$

$$\exp \left\{ -\frac{1}{2}(\beta - \mu)' \Psi (\beta - \mu) \right\} \prod_{i=1}^n \Omega^{-1/2} \mu_n - \rho W \exp \left\{ -\frac{1}{2\sigma^2} (b_1 + \frac{1}{2} \sum_{i=1}^n \Omega^{-1} (y - X \beta)^2) \right\}^{-(a_1+1-n/2)} \quad (2.30)$$

$$\pi(\lambda | \beta, X, y, W) \propto \prod_{i=1}^n \Omega^{-1/2} \mu_n - \rho W \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \Omega^{-1} (y - X \beta)^2 \right\}^{-(a_1+1-n/2)} \quad (2.31)$$

2.1 Markov Chain Monte Carlo Simulation

Simulating β

The Gibbs simulation for model parameter β begins with $\beta^{(1)} | \theta_{-p}^{(0)} \sim p(\beta | \theta_{-p}^{(0)}, Ay, X)$ where the superscript r in $\beta^{(r)}$ denotes the r th simulated β , thus $r = 0$ denoted the initial value, for $r = 1, \dots, R$, and R is the number of simulations in the Markov Chain Monte Carlo. The full marginal distribution of β given θ_{-p} is $\beta | (\theta_{-p}, y, X) \sim N(\bar{\beta}, \sigma^{-2}(X\Omega^{-1}X)^{-1})$ where $\bar{\beta} = (X\Omega^{-1}X)^{-1}X\Omega^{-1}Ay$ (2.32)

2.2 Simulating σ^{-2}

The kernel of inverted gamma probability density function

$\sigma^{-2} \propto IG(a_1 + \frac{n}{2}, b_1 + \frac{1}{2} \sum_{i=1}^n (Ay_i - x\beta)^2 \Omega^{-1} (Ay_i - x\beta))$ is set as the initial value of θ_{-p} with succeeding simulation of

$$\sigma^{(r)} | \theta_{-p}^{(r-1)} \sim IG\left(a_1 + \frac{n}{2}, b_1 + \frac{e^{(r-1)} A^{-1} e^{(r-1)}}{2}\right), \quad (2.33)$$

$$\text{where } e^{(1)} = Ay - X\beta^{(1)}, \quad A = I_n - \rho W \quad (2.34)$$

3.0 Generation of Sample and Analysis Procedures

The study adopted six sample sizes specified as: 25, 50, 100, 200, 500 and 1000. [14], [15], [19] and proposed HCS heteroscedastic covariance matrix were injected into Spatial Autoregressive Model and observed asymptotic behaviour via the finite properties of consistency and efficiency. The weight matrix was tagged as exogenous variable and construct its contiguity spatial matrix as $c_{ij} = c_{ji}$ be it symmetry, for $i = j$, thus $c_{ii} = 0$ for $i = j$. Then weight $w_{ij} = \frac{c_{ij}}{\sum_{j=1}^n c_{ij}}$, $i, j = 1, \dots, n$.

Coordinate references of 94 cities in Nigeria were collected, minimum and maximum values of the coordinate references were obtained, and these were used to generate coordinate reference for the entire simulated data using uniform distribution. The study computed row-standardized spatial weights with the coordinates' reference obtained from each sample generated. ρ denoted the spatial autocorrelation term, of 0.33 and 0.77.

The error term ϵ_i were generated based on seven different heteroscedastic covariance error structure with spatial autocorrelated values. Thereafter, the study injected ϵ_i into the SAR model and the parameters β_0 , β_1 and β_2 were hypothetically set at 1, 2.5 and 1.8 respectively to generate variable y variable which was conjugated with contiguity of the spatial matrix W . The number of replications of the experiment was set to 11,000 iterations with a burn-in of 1000 and thinning of 5.

4.0 Analysis and Results

Table 4.1 showing heteroscedasticity $w_i = \text{constant}$

samples		ρ		
		β_0	β_1	β_2
ρ	T_{true}	1.0000	2.5000	1.8000
0.33	25	9282(9398)	4191(6078)	3786(3118)
	50	2468(8937)	1636(1118)	1684(8935)
	100	2852(1008)	1098(2956)	0573(8360)
	200	0183(0089)	0637(0199)	2011(0573)
	500	0225(0038)	0266(0074)	0068(0037)
	1000	0427(0036)	0128(0033)	0523(0061)
	25	9282(9398)	6591(6078)	3586(3118)
	50	2468(8937)	1636(1118)	1684(8935)
	100	2852(1008)	5496(2956)	0573(8360)
	200	0183(0089)	0637(0200)	2011(0573)
0.77	500	0225(0038)	0266(0074)	0000(0000)
	1000	0000(0000)	0128(0033)	0523(0061)

Table 4.1 gave the outcome from the estimation of model with constant variance which does not account for the presence of heteroscedastic disturbance term. It revealed that the absolute bias for the intercept β_0 , slopes β_1 and β_2 decrease and increase as the sample size augmented, this is due to the presence of spatial autocorrelation. The pattern for both ρ at 0.33 and 0.77 are similar for both the intercept and slopes. The study concluded that there is inconsistency due to the presence of spatial autocorrelation. The mean squares error of estimation of constant variance decrease as the sample sizes increase from 25 to 50, but suddenly increase as the sample size increases to 100, but remain stable in decreasing word as the sample sizes increase.

Table 4.2: showing HCS: $w_i = \bar{u}_i^2$

samples		ρ		
		β_0	β_1	β_2
0.33	25	5741(3969)	3028(2501)	4582(4214)
	50	0449(0464)	1026(0878)	0840(0752)
	100	0135(0180)	0074(0304)	0212(0346)
	200	0150(0094)	0216(0357)	0144(0181)
	500	0034(0035)	0058(0062)	0014(0058)
	1000	0010(0037)	0106(0031)	0100(0033)
	25	3688(3913)	2992(2467)	4451(4080)
	50	0478(0498)	1071(0884)	0928(0828)
	100	0112(0191)	0049(0521)	0252(0367)
	200	0178(0180)	0262(0169)	0027(0184)
0.77	500	0000(0000)	0055(0065)	0013(0057)
	1000	0003(0019)	0082(0031)	0000(0000)

Table 4.2 showed the outcome from the estimation of model HCO which account for heteroscedasticity when the diagonal elements from the squares of the residuals. The absolute bias for intercept and slopes decrease algebraically as the sample size increase from small sample to large sample sizes. Though, there exist inconsistency in sample size 200 for slope β_1 and sample size 1000 for β_2 . The study observed similar pattern for both spatial autocorrelation ρ of 0.33 and 0.77 respectively. The study found out that at large sample sizes, the observed model parameters estimates are close to true values. The mean square error of estimation of HCO decreases as the sample sizes augmented from 25 to 1000. The study was able to establish asymptotic efficiency and opined that large sample sizes will be appropriate when spatial autocorrelation and heteroscedasticity are at level of HCO.

Table 4.3 showing HC1: $w_i = \frac{n}{n-k} \hat{u}_i^2$

	samples	ρ	β_0	β_1	β_2
0.33	25	Bias(MSE)	5745(.4031)	3036(.2638)	4291(.4404)
	50		0448(.0482)	1020(.0902)	0839(.0780)
	100		0135(.0184)	0074(.0514)	0213(.0553)
	200		0150(.0095)	0216(.0158)	0544(.0183)
	500		0036(.0035)	0059(.0062)	0014(.0058)
	1000		0010(.0017)	0106(.0054)	0100(.0053)
0.77	25	Bias(MSE)	5779(.4121)	3036(.2638)	4563(.4404)
	50		0448(.0540)	0999(.0913)	0915(.0878)
	100		0134(.0208)	0047(.0535)	0283(.0793)
	200		0144(.0104)	0252(.0161)	0000(.0000)
	500		0000(.0000)	0027(.0064)	0000(.0000)
	1000		0000(.0000)	0079(.0070)	0000(.0000)

Table 4.3 showed the outcome from the estimation of model HC1 which accounts for heteroscedasticity when the diagonal elements from the squares of the residuals multiplied by sample size and divided by sample size less rank of the independent variables. The absolute bias for intercept and slopes decrease algebraically as the sample sizes increase from small sample to large sample sizes, though there exist inconsistency in sample size 200 for slope β_0 and sample size 1000 for β_2 . The study observed similar pattern for both spatial autocorrelation ρ of 0.33 and 0.77 respectively. The study found out that at large sample sizes the observed model parameters estimates are close to true values particularly when the spatial autocorrelation is high.

The mean square error of estimation of HC1 decreases as the sample sizes augmented from 25 to 1000, the study was able to establish asymptotic efficiency and opined that large sample sizes will be appropriate when spatial autocorrelation and heteroscedasticity are at level of HC1. This was attested to by the estimated parameters that are close to true values.

Table 4.4 showing HC2: $w_i = \frac{\hat{u}_i^2}{1 - \hat{h}_i}$

ρ	samples	β_0	β_1	β_2
0.33	25	5812(.4210)	3648(.3416)	4215(.4541)
	50	0445(.0515)	1010(.0965)	0842(.0833)
	100	0138(.0192)	0079(.0537)	0214(.0369)
	200	0150(.0096)	0216(.0161)	0143(.0186)
	500	0036(.0036)	0059(.0062)	0014(.0058)
	1000	0010(.0017)	0106(.0054)	0100(.0053)
0.77	25	5831(.4247)	3806(.3754)	4078(.4739)
	50	0440(.0541)	1057(.1005)	0859(.0939)
	100	0141(.0203)	0085(.0576)	0256(.0409)
	200	0164(.0103)	0256(.0175)	0099(.0201)
	500	0000(.0000)	0014(.0063)	0000(.0000)
	1000	0000(.0000)	0066(.0030)	0000(.0000)

Table 4.4 showed the outcome from the estimation of model HC2 which account for heteroscedasticity when the diagonal elements from the squares of the residuals divided one less hat matrix of the independent variables. The absolute bias for intercept and slopes decreases algebraically as the sample sizes increase from small sample to large sample sizes, though there exist inconsistency in sample size 200 for slope β_0 , β_1 and sample size 1000 for β_2 . The study observed similar pattern for both spatial autocorrelation ρ of 0.33 and 0.77 respectively. The study found out that at large sample sizes, the observed model parameters estimates closed to true values at parameters β_0 and β_2 particularly when the spatial autocorrelation is high. The mean square error of estimation of HC2 decreases as the sample sizes augmented from 25 to 1000. The study was able to establish asymptotic efficiency and opined that large sample sizes will be appropriate when spatial autocorrelation and heteroscedasticity are at level of HC2. This was attested to by the estimated parameters that are closed to true values.

Table 4.5 showing HC3: $w_i = \frac{\hat{u}_i^2}{(1 - \hat{h}_i)^2}$

ρ	samples		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
0.33	25	Bias(mse)	.5878(.4410)	.4188(.4317)	.3906(.4831)
	50		.9443(.9550)	.0931(.1494)	.0946(.0890)
	100		.0141(.0201)	.0083(.0941)	.0216(.0387)
	200		.0148(.0098)	.0236(.0165)	.0143(.0190)
	500		.0036(.0034)	.0059(.0063)	.0013(.0059)
	1000		.0018(.0017)	.0106(.0011)	.0100(.0033)
0.77	25	Bias(mse)	.5837(.4448)	.4333(.4443)	.3627(.4823)
	50		.0400(.0378)	.0842(.1012)	.0973(.0982)
	100		.0115(.0216)	.0000(.0000)	.0393(.0433)
	200		.0184(.0107)	.0311(.0174)	.0000(.0000)
	500		.0000(.0000)	.0000(.0000)	.0000(.0000)
	1000		.0000(.0000)	.0074(.0031)	.0000(.0000)

Table 4.5 showed the outcome from the estimation of model HC3 which account for heteroscedasticity when the diagonal elements form the squares of the residuals divided squares of one less hat matrix of the independent variables. The absolute bias for intercept and slopes decreases algebraically as the sample sizes increase from small sample to large sample sizes, though there exist inconsistency in sample size 200 for slope $\hat{\beta}_0$, $\hat{\beta}_1$ and sample size 1000 for $\hat{\beta}_2$. The study observed similar pattern for both spatial autocorrelation ρ of 0.33 and 0.77 respectively. The study found out that at large sample sizes the observed model parameters estimates closed to true values at parameters $\hat{\beta}_0$ and $\hat{\beta}_2$ particularly when the spatial autocorrelation is high.

The mean square error of estimation of HC3 decreases as the sample sizes augmented from 25 to 1000. The study was able to established asymptotic efficiency and opined that large sample sizes will be appropriate when spatial autocorrelation is high and heteroscedasticity are at level of HC3. This was attested to by the estimated parameters that are close to true values.

Table 4.6 showing HC4: $w_i = \frac{\hat{u}_i^2}{(1 - \hat{h}_i)^\lambda}$

ρ	samples		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
0.33	25	Bias(mse)	.5773(.4098)	.3734(.3435)	.4440(.4430)
	50		.0447(.0493)	.1011(.0824)	.0840(.0798)
	100		.0136(.0187)	.0076(.0323)	.0213(.0338)
	200		.0150(.0045)	.0216(.0139)	.0144(.0184)
	500		.0120(.0093)	.0216(.0139)	.0144(.0184)
	1000		.0010(.0017)	.0106(.0031)	.0109(.0033)
0.77	25	Bias(mse)	.5750(.4079)	.3393(.3853)	.4236(.4238)
	50		.0516(.0535)	.1105(.0996)	.1007(.0832)
	100		.0163(.0207)	.0082(.0582)	.0560(.0282)
	200		.0143(.0103)	.0256(.0176)	.0000(.0000)
	500		.0093(.0093)	.0023(.0068)	.0000(.0000)
	1000		.0000(.0000)	.0093(.0032)	.0000(.0000)

Table 4.6 showed the outcome from the estimation of model HC4 which account for heteroscedasticity when the diagonal elements form the squares of the residuals divided lambda power one less hat matrix of the independent variables, lambda has been defined in the methodology. The absolute bias for intercept and slopes decreases algebraically as the sample sizes increase from small sample to large sample sizes, though there exist inconsistency in sample size 200 for slope $\hat{\beta}_0$, and $\hat{\beta}_1$. The study observed similar pattern for both spatial autocorrelation ρ of 0.33 and 0.77 respectively. The study found out that at large sample sizes the observed model parameters estimates closed to true values at parameters $\hat{\beta}_0$ and $\hat{\beta}_2$ particularly when the spatial autocorrelation is high.

The mean square error of estimation of HC4 decreases as the sample sizes augmented from 25 to 1000. The study was able to established asymptotic efficiency and opined that large sample sizes will be appropriate when spatial autocorrelation is high and heteroscedasticity are at level of HC4. This was attested to by the estimated parameters that are close to true values.

Table 4.7 showing proposed HC5w: $w_i = \frac{\hat{u}_i^2}{\lambda(1 - \hat{h}_i)}$

ρ	samples		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
0.33	25	Bias(MSE)	.5779(.4119)	.3340(.3812)	.4396(.4458)
	50		.0447(.0498)	.1015(.0933)	.0820(.0800)
	100		.0136(.0188)	.0076(.0323)	.0213(.0361)
	200		.0150(.0046)	.0216(.0160)	.0144(.0185)
	500		.0036(.0036)	.0058(.0062)	.0014(.0038)
	1000		.0010(.0017)	.0106(.0031)	.0100(.0033)
0.77	25	Bias(MSE)	.5766(.4106)	.3477(.3683)	.4201(.4390)
	50		.0435(.0519)	.0981(.0824)	.0825(.0878)
	100		.0113(.0191)	.0029(.0340)	.0278(.0374)
	200		.0182(.0101)	.0273(.0169)	.0104(.0190)
	500		.0000(.0000)	.0012(.0061)	.0000(.0000)
	1000		.0011(.0017)	.0071(.0038)	.0000(.0000)

Table 4.7 showed the outcome from the estimation of model HC5 which account for heteroscedasticity when the diagonal elements form the squares of the residuals divided by squares root of one less hat matrix of the independent variables. The absolute bias for intercept and slopes decrease algebraically as the sample sizes increase from small sample to large sample sizes, though there exist inconsistency in sample size 200 for slope β_1 and β_2 . The study observed similar pattern for both spatial autocorrelation ρ of 0.33 and 0.77 respectively. The study found out that at large sample sizes, the observed model parameter estimates closed to true values at parameters β_1 and β_2 particularly when the spatial autocorrelation is high. The proposed heteroscedastic model equally performed better as other models of heteroscedastic.

The mean square error of estimation of HC5 decreases as the sample sizes augmented from 25 to 1000. The study was able to established asymptotic efficiency and opined that large sample sizes will be appropriate when spatial autocorrelation is high and heteroscedasticity are at level of HC5. This was attested to by the estimated parameters that are closed to true values.

4.0 Conclusion

In this paper, the study presented an efficient simple way of modeling and estimating spatial heteroscedastic Autoregressive model using conditional Gibbs sampler. Gibbs sampler was adopted since it has been established that Graddy Gibbs sampler outperform other MCMC algorithms [16]. The study observed asymptotically that the SHAR model performed consistently and efficiently across the seven heteroscedastic covariance matrix. The study established the same scenario for both absolute bias and mean squares errors criteria. Our approach can be applied to further studies in the area of linear mixed model, multi-level modeling and other econometrics models.

References

- [1] LeSage P. (1997): Regression analysis of Spatial data, *The Journal of regional Analysis and Policy* 27:83-94
- [2] Anselin L. and Griffith D.A. (1988): "Do Spatial Effect Really Matter in Regression Analysis Papers in the regional Science Association, 65, pp. 11-34
- [3] LeSage J.P. and Stevens J.J. (1994): "A Bayesian Approach to Spatial Expansion" University of Toledo working paper
- [4] LeSage J.P. and Murawski W.A. (1994): "Spatial Modeling of Gypsy Moth Movements," of Toledo working paper
- [5] Anselin L. (1988b): *Spatial Econometrics: Methods and Models*. Dordrecht, Kluwer Academic Publisher
- [6] Caselli E. (1992): "Bayesian Regression and Expansion Methods," *Geographical Analysis*, 4, pp. 51-91
- [7] Caselli E. and Jones J.P. (1987): "Spatial Application Methods of Expansion Paradigm," *Quantitative Analysis in Geography*, pp.121-136
- [8] Krivchyn A. (1997): "A Spatial Prior for Bayesian Vector Autoregressive Models. Paper presented at the Mid-Continent Regional Science Association meeting.
- [9] Doan T.R, Litterman B. and Sims C. (1984): Forecasting and conditional Projection using Realistic Prior Distributions, *Economic Review*, 3, pp. 113-129
- [10] Pan Z. and LeSage J.P. (1995): "Using Spatial Contiguity as Prior Information in Vector Autoregressive Models," *Economic Letters*, 47 pp. 137-142
- [11] Donald J. L. (2008): An Introduction to Bayesian Inference in Spatial Econometrics, West Virginia University Regional Research Institute <http://www.com/abstract=1244261>
- [12] LeSage P.L. (2010): Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models, *Geography Analysis* Vol 32, No 1, The Ohio State University
- [13] McMullen D.P. (1992): Probit with Spatial Autocorrelation, *Journal of Regional Science* vol. 32 No. 3 Pp417-438
- [14] White H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix and a Diagnostics Test for Heteroskedasticity," *Econometrica*, 48, 817-838
- [15] MacKinnon J.G. and White H. (1985): "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties," *Journal of Econometrics*, 29, 105-125.
- [16] Yoshihiro O. and Kazuhiko K. (2015) Comparison of the Sampling Efficiency in Spatial Autoregressive Model, *Open Journal of Statistics*, 2015, 5, 10-20 <http://www.scirp.org/journal/ojs> <http://dx.doi.org/10.4236/ojs.2015.54002>

- [17]. Harry H. K. and Ingmar R. P. (2010): Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances. *Journal of Econometrics* Volume 157, Issue 1, Pages 53-67. <http://dx.doi.org/10.1016/j.jeconom.2009.10.025>
- [18]. Pace, R. K. and LeSage, J. P. and Zhu S. (2011), Spatial Dependence in Regressors and its Effect on Estimator Performance. SSRN: <http://ssrn.com/abstract=1801241> or <http://dx.doi.org/10.2139/ssrn.1801241>