GIBBS SAMPLING ON EQUAL OBSERVATIONS IN A THREE-EQUATION SEEMINGLY UNRELATED REGRESSION MODEL.

BY

Damian Chibuike ASOGWA

(Matric No: 201702)

A Project submitted to the Department of STATISTICS,
Faculty of Science, University of Ibadan,
in partial fulfilment of the requirements for the Master's
Degree in Science
(M.Sc.)

November, 2018.

CERTIFICATION

This is to certify that this work was carried out by **Damian Chibuike ASOGWA (201702)** in the Department of Statistics, University of Ibadan.

…………………………. …………………………….

Supervisor *Date*

Dr. Oluwayemisi O. ALABA

B.Sc., M.Sc., Ph.D. (Statistics, Ibadan)

Department of Statistics,

University of Ibadan, Nigeria.

………………………….. …………………………….

Ag. Head of Department *Date*

Dr. Angela U. CHUKWU

B.Sc. Mathematics (Calabar), M.Sc., Ph.D. Statistics (Ibadan),

Department of Statistics,

University of Ibadan, Nigeria.

………………………….. …………………………….

External Supervisor *Date*

DEDICATION

I dedicate this work to God Almighty for his favorable mercies, love, kindness and faithfulness for making this work a success.

ACKNOWLEDGEMENTS

Firstly, my profound gratitude goes to the Almighty God, who in His infinite mercy has seen to the completion of this programme. Also I thank Him immensely for giving me a healthy mind and body to carry out my work successfully.

My utmost gratitude goes to my most enlightened parents, Dr. and Mrs. Asogwa for their support, love, prayers, care, sacrifice and encouragement which was the backbone of this programme. May God's blessing never lack from you. God will replenish you abundantly. To my ever-loving siblings, in the person of Chizzy, Onyinye, Chinazo, Chiamaka and Ogochukwu, I thank you all for your love, prayers and encouragement. As you soar high in life, may God be with you all.

I cannot forget to appreciate my amiable and sagacious supervisor in the person of Dr. Oluwayemisi O. Alaba, whom I have nothing but respect for, a woman of great integrity with impeccable qualities and tremendous principles, who despite her tight schedule, she still brought out her time to make this work a success. May God bless you and your family. I thank my father and mentor in academics Dr. Nwaigwe C.C, for his advice and encouragement. Remain blessed Sir.

With joy, I am grateful to all members of staff of the Department of Statistics, University of Ibadan. It has been a wonderful experience learning from you.

To all 2017/2018 MSc colleague, it has been a wonderful experience meeting with you and to my wonderful friends in the person of Oludotun, Chinedu, Nneamaka, Chidinma, and Shamshudeen.

I appreciate you all.

Damian Chibuike ASOGWA

November, 2018

ABSTRACT

Gibbs sampling a randomized MCMC algorithm, mainly for simulating a sequence of observations which are full conditional distributions of each parameter conditioned on the remaining parameters and computing posterior quantities of interest. The Seemingly Unrelated Regression model is a generalization of a linear regression model that are linked by the fact that their disturbances are contemporaneously correlated.

The Gibbs sampling algorithm of MCMC simulation techniques was used in estimating the parameters and distribution in a three-equation SUR model. Simulated data from R package using Gibbs sampling, were used to generate the explanatory variables from a U (-3, 3) distribution for various samples sizes N = 50, 100, 200, 500 and 1000 using 10000 replicates and taking 10% of each sample size as the burn-in period.

The summary of ABIAS for the marginal posterior pdf for $\beta's$ and the summary of ABIAS for the conditional posterior for $\beta|\Sigma$, as the various sample sizes N = 50, 100, 200, 500 and 1000 increases, there was fluctuation in the values of ABIAS obtained.

It was observed that as the sample size increases, the standard deviation values decreased consistently for the parameters. It can also be observed that the larger the sample size, the closer the posterior mean of the parameters gets to its true population parameters.

Contents

List of Tables

4.3d. Summary of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) Convergence diagnostic test statistic (CD) for $\Sigma 's$. N = 200

4.4a. Result for Marginal Posterior pdf for β and Σ. N = 500

4.4b. Result for Conditional Posterior pdf for $\beta | \Sigma$ and $\Sigma | \beta$. N = 500

4.4c. Summary of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) Convergence diagnostic test statistic (CD) for $\beta 's$. N = 500

4.4d. Summary of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) Convergence diagnostic test statistic (CD) for $\Sigma 's$. N = 500

4.5a. Result for Marginal Posterior pdf for β and Σ. N = 1000

4.5b. Result for Conditional Posterior pdf for $\beta | \Sigma$ and $\Sigma | \beta$. N = 1000

4.5c. Summary of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) Convergence diagnostic test statistic (CD) for $\beta 's$. N = 1000

4.5d. Summary of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) Convergence diagnostic test statistic (CD) for $\Sigma 's$. N = 1000

CHAPTER ONE

1.1. BACKGROUND OF THE STUDY

In many areas of research and application, the Seemingly Unrelated Regression (SUR) models, introduced by Zellner (1962), are used as tool to study a wide range of real phenomena. Many studies have contributed to the development of estimation, testing, prediction and other inference techniques for analysis of SUR model including Zellner (1962, 1963), Gallant (1975), Rocke (1989), Neudecker and Windmeijer (1991), Mandy and Martins (1993), Kurata (1999), Liu (2002), Ng (2002), Carroll *et al.* (2006). Also the SUR models and inference techniques for analysing the models are described in almost all Bayesian and non-Bayesian textbooks, providing many references to the literature. The first analysis of the SUR models appeared in Zellner (1962), who employed a generalised least squares approach. Later, likelihood and traditional Bayesian approaches have been developed, followed by various other inference approaches; see for example, the likelihood distributional approach (Fraser *et al.* (2005)), Bayesian analyses and so on.

In the Bayesian analysis of SUR models, one can apply the Gibbs algorithm of Percy (1992). However, we often want to estimate the regression coefficients, the variance matrix of the errors, and a set of variables to be included in the model simultaneously. It is obvious that the traditional method of best subset selection is computationally infeasible for high dimensional data. To solve this issue, Smith and Kohn (2000), recently introduced a Bayesian hierarchical SUR model and developed a Markov Chain Monte Carlo (MCMC) procedure.

Zellner's idea of combining several equations into one model to improve estimation efficiency Zellner (1962), ranks as one of the most successful and lasting innovation in the history of econometrics. The resulting SUR model has generated a wealth of both theoretical and empirical contributions.

Zellner (1971) popularised Bayesian inference in econometrics generally and described the SUR model within the context of Bayesian inference. However, at that time, convenient methods for deriving or estimating marginal posterior density functions and moments for individual SUR coefficient were not generally available. Recently, the application of Markov Chain Monte Carlo (MCMC) methodology to Bayesian inference has made available a new range of numerical methods that make Bayesian estimation of the SUR model more convenient and accessible.

1.2. SEEMINGLY UNRELATED REGRESSION.

In some cases during our research we often found out that, we are dealing with several number of equations which contain an amount of same regressors. Those regressors in some way have known influence on some different regresand. For example we want to learn about investment on two big companies, as illustrated by (Zellner, 1962). The two situation of investment at two companies (say Westinghouse and General Electrics). As we described the situation as two different equations, and assume that in every equation real gross investment were determined by an outstanding shares at the beginning period and the opening value of each companies capital stock. On those conditions as the data for the equation came from the same periods we can expect that error terms of those two equations may be contemporaneously correlated (Zellner, 1962). Correlation of the error terms comes up as given presence of common market forces condition and the likelihood of similar factors which affect the inclusion of error terms in the regression. For instance, if error terms from the first equation is related to some unobservable variables, which mostly also followed by the other equation error terms, so we can find that error terms of each equation are correlated. This condition made those two equations “seemingly unrelated”.

A SUR is a system of regression equations which consists of a set of regression equations, each of which contains different explanatory variables and satisfies the classical assumptions of the standard regression model. A SUR system comprises several individual relationships that are linked by the fact that their disturbances are correlated. Such models have been found useful in many applications. For example, demand functions can be estimated for different households (or household types) for a given commodity. The correlation among the equation disturbances could come from several sources such as correlated shocks to household income.

The SUR model explains the variation of not just one dependent variable, as in the univariate multiple regression model, but the variation of a set of “m” dependent variables, e.g. the monthly consumption expenditures of “m” consumers or the annual voting behavior of “m” voters, in terms of the variation of general and specific input or independent variables and error terms specific to each individual, problems that are frequently encountered in many sciences.

Indeed, Geweke (2003) has written, “The SUR model developed in Zellner (1962) is perhaps the most widely used econometric model after linear regressions. The reason is that it provides a simple and useful representation of systems of demand equations that arise in neoclassical static theories of producer and consumer behavior.” It is the case that a SUR model is a collection of two or more regression relations that can be analysed with data on the dependent and independent variables.

1.2.1. DEVELOPMENT OF SUR MODEL

As many statistics models, SUR has experienced modification and development. Most modifications were addressed to improve SUR application to certain situation. Some of these developments were known as follow:

1.2.1.1. Dynamic Seemingly Unrelated Regression (DSUR). One of parameter estimator proposed by Mark, Ogaki and DonggyuSul (2004). The DSUR has proven to be feasible especially for balanced panel data which the number of cointegrating equation (N) were substantially smaller than the number of time series observations T. Dynamic Seemingly Unrelated Regression (DSUR) is also applicable for situation which cointegrating vectors are homogeneous across equations or where they are not. This model has proven to be properly used in case of estimating relation of investment rate over saving rates in European country (Mark *et.al*, 2004).

1.2.1.2. Sparse Seemingly Unrelated Regression (SSUR). As conventional SUR model which is unconstrained model has known to be over parameterized, Wang (2010) has introduced, the Bayesian analysis of SSUR.

Sparse Seemingly Unrelated Regression has the main innovations which include;

- Inferences via Markov Chain Monte Carlo (MCMC) simulations for specific constraints of regression coefficients and errors,
- Evaluations of the marginal likelihoods of restrictions using coupled Candidate's formula approximations, and
- The extension of sparse modelling to dynamic SUR models.

In SUR equation system, the equations are related based on one or both of these following ways:

- The error terms of each equations are related. The error terms are correlated if there are common unobserved factors which influence the dependent variables of the equations.
- There is a relation among the parameters of each different equations. This condition will occur if the same parameters appear in more than one equation.

1.2.2. PROPERTIES OF THE SEEMINGLY UNRELATED REGRESSION MODEL

1. The SUR model is used to gain efficiency when the equations are only related through the error term.
2. The parameters in the SUR model generally vary from equation to equation.
3. Regressors may or may not vary from equation to equation depending on the model.
4. The SUR estimates result in equation-by-equation OLS estimates when:
    - The errors are uncorrelated across equations.
    - Each of the equations contains exactly the same set of regressors.

1.2.3. MOTIVATION FOR SEEMINGLY UNRELATED REGRESSION MODEL

There are two main motivations for use of SUR. The first one is to gain efficiency in estimation by combining information on different equations. The second motivation is to impose and/or test restrictions that involve parameters in different equations (Roger Moon, *et al* (2006)).

1.3. THE MARKOV CHAIN MONTE CARLO METHOD

Markov Chain Monte Carlo (MCMC) methodology provides enormous scope for realistic statistical modelling. Until recently, acknowledging the full complexity and structure in many applications was difficult and required the development of specific methodology and purpose-built software. Now, MCMC methods provide a unifying framework within which many complex problems can be analysed using generic software.

The MCMC is essentially Monte Carlo integration using Markov chains. Bayesians, and sometimes also frequentists, need to integrate over possibly high-dimensional probability distribution to make inference about model parameters or to make predictions. Bayesians need to integrate over the posterior distribution of model parameters given the data, and frequentists may need to integrate over the distribution of observables given parameter values.

One of the most popular approaches for estimating the SUR model in a Bayesian framework involves the use of MCMC method in order to compute posterior densities for parameters and predictive density functions. It allows one to characterise a distribution without knowing all of the distribution's mathematical properties by randomly sampling values out of the distribution, it's also a method that allows one to approximate complex integrals using stochastic sampling routines.

A particular strength of MCMC is that it can be used to draw samples from distributions even when all that is known about the distribution is how to calculate the density for different samples. As MCMC's indicates, the method is composed of two components; Markov chain and Monte Carlo Integration.

1.3.1. MARKOV CHAIN

A Markov chain is a sequential model that transits from one state to another in a probabilistic fashion, where the next state that the chain takes is conditioned on the previous state.

1.3.2. MONTE CARLO INTEGRATION

Monte Carlo integration is a powerful technique that exploits stochastic sampling of the distribution in question in order to approximate the difficult integration. However, in order to use Monte Carlo Integration it is necessary to be able to sample from the probability distribution in question, which may be difficult to sample directly.

The MCMC approaches are so-named because one uses the previous sample values to generate the next sample value, generating a Markov (as the transition probabilities between samples values are only a function of the most recent sample value). The MCMC methods comprise a class of algorithm for sampling from a probability distribution, the Gibbs sampling algorithm and Metropolis-Hasting algorithm. This study is focused on Gibbs sampling.

1.4. THE GIBBS SAMPLER

The Gibbs sampler was introduced as a MCMC tool in the context of images restoration by Geman and Geman (1984). Gelfand and Smith (1990), offered the Gibbs sampler as a very general approach for fitting statistical models. One of the MCMC methods is called Gibbs sampling algorithm. It is mainly based on simulating the full conditional distributions of each parameter vector conditioned on the remaining data parameters and computing posterior quantities of interest. Gibbs sampling is a randomised algorithm (i.e. it is an algorithm that make use of random numbers). As with other MCMC algorithm, Gibbs sampling generates a Markov chain of samples, each of which is correlated with nearby samples. The Gibbs sampler, is very widely applicable to a broad class of Bayesian problems has sparked a major increase in the application of Bayesian analysis, and this interest is likely to continue expanding for some time to come.

To introduce the Gibbs sampler, consider a bivariate random variable (x, y), and suppose we wish to compute one or both marginal, $p(x)$ and $p(y)$. The idea behind the sampler is that it is far easier to consider a sequence of conditional distributions, $p(x \mid y)$ and $p(y \mid x)$, than it is to obtain the marginal by integration of the joint density $p(x, y)$, e.g., $p(x) = \int p(x, y)dy$. The sampler starts with some initial value y_0 for y and obtains x_0 by generating a random variable from the conditional distribution $p(x \mid y = y_0)$. The sampler then uses x_0 to generate a new value of, drawing from the conditional based on the value x_0, $p(y \mid x = x_0)$. The sampler proceeds as follows

$$x_i \sim p(x \mid y = y_{i-1})$$

$$y_i \sim p(y \mid x = x_i) .$$

Thus Gibbs sampling consists purely in sampling from full conditional distributions. The beauty of Gibbs sampling is that it simplifies a complex high-dimensional problems by breaking it down into simple, low-dimensional problems.

1.5. JUSTIFICATION OF STUDY.

In Bayesian Inference, convenient methods for deriving or estimating marginal posterior density function and moments for individual SUR coefficients are not generally available.

However, the application of Markov Chain Monte Carlo in Bayesian Inference has made available a new range of numerical methods that make Bayesian estimation of the SUR model more convenient and accessible.

This work employed, the Gibbs sampling procedure in the estimation of equal observations in a three-equation seemingly unrelated regression model.

1.6. AIM AND OBJECTIVES

The aim of this study is to employ the Gibbs sampling algorithm of MCMC simulation techniques in estimating the parameters and its distribution on equal observations in a three-equation SUR model.

The objectives of this study are;

1. To estimate the marginal posterior probability density function.
2. To estimate the conditional posterior probability density function for $(\beta \mid \Sigma)$ and $(\Sigma \mid \beta)$.

1.7. DATA GENERATION

For the purpose of this study, we simulated data from the R package using Gibbs sampling, to generate the explanatory variables from a U (-3, 3) distribution for various sample sizes N = 50, 100, 200, 500 and 1000, using 10000 replicates and taking 10% of each sample size as the burn-in period.

1.8. DEFINITION OF TERMS

In this section, definition of some terms used in the study will be given, to facilitate the better understanding of it by the readers. Among others, these include:

- **SIMULATION:** is the process of designing of a model of a real system and conducting experiment with the model for the purpose of understanding the behavior for the operation of the system.
- **PARAMETER:** this is an unknown characteristic of the population whose value we are interested to know but may be estimated from sample data.

- **BURN-IN:** is a colloquial term that describes the practice of throwing away some iterations at the beginning of an MCMC run. The burn-in notion says, you start somewhere say at (x), then you run the Markov chain for n steps, from which you throw away all the data (no output). This is the burn-in period. After the burn-in you run normally, using each iteration in your MCMC calculations.
- **PRIOR DISTRIBUTION:** a prior distribution of a parameter is the probability distribution that represents your uncertainty about the parameter before the current examined.
- **NONINFORMATIVE PRIOR:** a prior distribution is non-informative if the prior is "flat" relative to the likelihood function. Thus, a prior $\pi(\theta)$ is non-informative if it has minimal impact on the posterior distribution of θ. Other names for the non-informative prior are vague, diffuse and flat prior. Many statisticians favor non-informative priors because they appear to be more objective. However, it is unrealistic to expect that non-informative priors represent total ignorance about the parameter of interest. In addition, non-informative, priors are often not invariant under transformation, that is, a prior might be non-informative in one parameterisation but not necessarily non-informative if a transformation is applied.
- **LIKELIHOOD FUNCTION:** is always defined as a function of the parameter θ equal to (or sometimes proportional to) the density of the observed data with respect to a common or reference measure, for both discrete and continuous probability distribution.
- **POSTERIOR DISTRIBUTION:** a posterior distribution is obtained from the product of prior distribution and likelihood function. Thus, a combination of prior distribution and likelihood based on the parameter for the given data is known as posterior distribution.

The posterior distribution summarises the current state of knowledge about all the uncertain quantities, (including unobservable parameters and also missing, latent, and observed potential data) in a Bayesian analysis.

- **MARGINAL DISTRIBUTION:** is where we are only interested in one of the random variables. In other words, either X or Y. In statistics, the marginal distribution of a subset of a collection of random variables is the probability distribution of the variables contained in the subset. It gives the probabilities of various values of the variables in the subset without reference to the values of the other variables.
- **CONDITIONAL DISTRIBUTION:** is where we are only interested in a particular sub-population of our entire data set. More generally, one can refer to the conditional distribution of a subset of a set of more than two variables; this conditional distribution is contingent on the values of all the remaining variables, and if more than one variable is included in the subset then this conditional distribution is the conditional joint distribution of the included variables.

CHAPTER TWO

2.0 LITERATURE REVIEW.

In econometrics, the Seemingly Unrelated Regression (SUR), proposed by (Zellner, 1962) and in (1963), Stewart (1980) and Parks (1967) is a generalisation of a linear regression model that consist of several regression equations, each having its own dependent variable and potentially different sets of exogenous explanatory variables. Each equation is a valid linear regression on its own and could be estimated separately, this justifies the system of equations being called seemingly unrelated.

Parks (1967), considered the problem of obtaining efficient estimates for the parameters of a system of regression equation when the disturbance terms are normally distributed. When the number of equations is greater than two equations. The disturbance terms of this system are assumed to be related by contemporaneous correlation in the different equations under the assumption he developed an estimator that is consistent and has the same normal distribution as the two least squares estimators which assumed the covariance matrix to be known.

Kmenta and Gilbert (1968), described the small sample behavior of these estimator in a number of specific cases by means of Monte Carlo experiment, discussed the small sample properties of five alternative estimators of set of linear regression equation when disturbances are correlated. Two of these procedures are least squares and Aitken two stage least squares. They examined these estimator with the different samples sizes, studied properties of these estimators and compared the results of five procedures, thus they explained that in many cases, Zellner's estimator of using Aitken two stage least squares are better than the other estimators when replacing the variance-covariance matrix of disturbance terms Σ by its consistent estimate ($\Sigma = S$). The estimate of the regression coefficient can be used for calculating a new set of residuals so as to obtain a new estimate of variance-covariance matrix Σ which can be used for obtaining new estimates of the regression coefficient.

Revankar (1974), considered a system of two SUR equations model and examined some finite sample properties of coefficient estimators based on the variance of residuals, derived the exact variance covariance matrix of Zellner's estimator with the normality distribution and additional condition, where the explanatory variable in the second equation is a subset of the explanatory variable of first equation. He showed that the coefficient of the estimators based on the unrestricted estimators of the variance-covariance of disturbances term depend upon the residuals from the OLS regression by applying regression of dependent variables on all explanatory variables. Revankar, posited that the coefficient estimators are Best Linear Unbiased Estimator and derived the exact moments of Zellner's estimator while showing that the efficiency increased rapidly as the sample size increases.

Chib and Greenberg (1994), developed practical and exact methods of analyzing ARMA (p, q) regression error models in Bayesian framework by using the Gibbs sampling and Metropolis – Hasting algorithms, and they also prove that the kernel of the proposed Markov chain sampler converges to the true density. The procedures can be applied to pure ARMA time series models and determine features of the likelihood function by choosing appropriate diffuse priors.

Chib and Greenberg (1996), also explained MCMC methods in some details and illustrated their application to problems in econometrics. Their purpose was to explain how these methods work both in theory and in practical applications. Many problems in Bayesian statistics (such as the computation of posterior moments and marginal density functions) can be solved by simulating the posterior distribution, they emphasise Bayesian applications, but these tools are also valuable in frequentist inference, where they can be used to explore the likelihood surface and to find modal estimates or maximum likelihood estimates with diffuse priors.

Smith and Kohn (2000), considered a system of regression equations that can seem unrelated, but actually are because their errors are correlated. Such a system of equations is called a setoff seemingly unrelated regression, or a SUR model.

Liu (2002), derived simpler expressions under a certain structure of design matrices for the two-stage Aitken estimates of the regression coefficients of two SUR equations. The estimates are shown to have smaller variance than the OLS estimates for sufficiently large samples.

Griffiths and Valenzuela (2006), were concerned with estimating a model that contains within its several SUR models. They added a third dimension to the conventional SUR model that typically has two dimensions, a number of equations and repeated observations on the variables in these equations. They called the model that contains several SUR models 'a set of seemingly unrelated regressions', also they were interested in deriving convenient conditional posterior densities that can be used within a Gibbs sampler for sampling from the joint posterior density.

Yahya, *et al* (2008), examined the relative gain/loss in efficiency of SUR estimators when one or more pair of the predictors in the system of equations are correlated (non-orthogonal). Literature has revealed that multicollinearity often affects the efficiency of SUR estimators. However, it addresses such challenges by determining the 'Tolerable Non-orthogonal Correlation Points' (TNCP) among the predictors at which the efficiency of SUR estimators will still be preserved. Results from their simulation studies showed that SUR estimators are still efficient up to the range of the TNCP values, especially when the predictors have Gaussian distribution.

Wang (2010), developed a Sparse Seemingly Unrelated Regression (SSUR) model with Gaussian errors; that is, a set of regressions in which both the regression coefficients and the error precision matrix have many zeros. Zeros in regression coefficients arise when each response possibly only depends on a subset of different predictors; zeros in a precision matrix arise when the error terms satisfy a set of conditional independence restrictions consistent with an underlying graphical model. He studied and proposed a fully Bayesian analysis of the SSUR model, and provided effective methods for marginal likelihood computation using a specified subset of variables and a specified graphical model to structure the covariance matrix.

Wang, *et al* (2010), this vignette focuses on the Gibbs sampler. He provided a review of its origins and its crossover into the mainstream statistical literature. He then attempted an assessment of the impact of Gibbs sampling on the research community, on both statisticians and subject area scientists. Finally, he offered some thoughts on where the technology is headed and what needs to be done as we move into the next millennium.

Zellner and Ando (2010), gave a description of computationally efficient methods for the Bayesian analysis of Student-t SUR models with unknown degrees of freedom. The method combines a Direct Monte Carlo (DMC) approach with an importance sampling procedure to calculate Bayesian estimation and prediction results using a diffuse prior. This approach is employed to compute Bayesian posterior densities for the parameters, as well as predictive densities for future values of variables and the associated moments, intervals and other quantities that are useful to both forecasters and others. The results obtained using our approach are compared to those yielded by the use of DMC for a standard normal SUR model.

Zellner (2010), presented new operational convergence criteria for the Gibbs sampler (GS) and related procedures that are useful for determining whether they not, only have converged but also have converged to provide reliable results. The three Gibbs sampler convergence criteria are presented and applied: the difference convergence criterion (DC^2), the ratio convergence criterion (RC^2), and the anchored ratio convergence criterion (ARC^2). Their uses and properties are discussed and examples are analysed to illustrate their application.

Ando (2011), used the Bayesian Variable Selection for the SUR models with a large number of predictors. Computationally efficient methods for Bayesian analysis of SUR models with large number of predictors are developed. One of the most crucial problems in Bayesian modelling of SUR models is how to determine the optimal combination of predictors. Under a Bayesian hierarchical framework where each regression function is represented as a linear combination of a large number of basic functions, the regression coefficients, the variance matrix of the errors, and a set of predictors to be included in the model were estimated simultaneously. Usually, the Bayesian model estimation problem is solved using MCMC techniques. Here, they showed how a DMC techniques can be employed to solve the variable selection and model parameter estimation problems more efficiently.

Adepoju and Akinwumi (2017), examined the effect of outliers on the performances of SUR and OLS estimators using Monte Carlo simulation method. The Cholesky method was used to partition the variance-covariance matrix Σ by decomposing it into the upper and lower non-singular triangular matrices. Varying degree of outliers; 0%, 5%, and 10% were each introduced into five sample sizes; 20, 40, 60, 100 and 500 respectively. The performances of the estimators were evaluated using Absolute Bias (ABIAS) and Mean Square Error (MSE). The results showed that at 0% outliers (when outliers were absent), the ABIAS and MSE of the SUR and OLS estimators showed similar results. At 5% and 10% outliers, the magnitude in ABIAS and MSE for both estimators increased but the SUR estimator showed better performance than the OLS estimator. As the sample size increases, ABIAS and MSE of the estimators decreased consistently for the various degrees of outliers considered with SUR consistently better than OLS.

Huang (2017), developed a practical sampling scheme for Bayesian analysis of correlated censored data using the seemingly unrelated Tobit regressions model. Posterior inference was performed via the Gibbs sampler with data augmentation algorithm. In particular, the relevant full conditional distributions needed in the use of Gibbs sampler are derived. The method is then applied to a real set on the determination of the payments of cash and dividends.

However, based on the existing literature, in this study we employ the Gibbs sampling algorithm of MCMC simulation techniques in estimating the parameters and its distribution on equal observations in a three-equation SUR model.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1. INTRODUCTION

The SUR model was introduced by Zellner (1962) to accommodate and take advantage of contemporaneous correlation in the errors of linear models that might otherwise appear unrelated. Since then, it has been studied extensively and has become commonplace in economic applications involving joint estimation of a number of equations. In order to gain estimation efficiency, Arnold Zellner combined several equations into one model and now this tool is used to study the impact of a wide range of phenomena, especially in econometrics and biometrics.

3.2. SUR MODEL SPECIFICATION

In this paper, we considered a set of *M* equations, this *M* number of seemingly unrelated regression equations can be represented in a matrix form as:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & X_M \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_M \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_M \end{bmatrix} \tag{3.1}$$

This system of equations in (3.1) can further be presented in a more compact form and when stacked together the whole system becomes:

$$y = X\beta + \varepsilon \tag{3.2}$$

Then algebraically, the SUR model is represented as:

$$y_i = X_i \beta_i + \varepsilon_i \tag{3.3}$$

Where $i = 1, 2, \cdots M$, y_i is an $n \ x \ 1$ vector of observations on the i^{th} response variable, X_i is an $n \ x \ p_i$ matrix of explanatory variables, β_i is a $p_i \ x \ 1$ vector of regression parameter and ε_i is the corresponding $n \ x \ 1$ vector of distribution. $\varepsilon \sim N(0, \Sigma \otimes I_T)$, we assume that the errors are heteroscedastic (different variances) and correlated across the equations. The regression equation in (3.1), appear independent (seemingly unrelated) with one another because they do not have common variables or parameter.

3.2.1. MODEL ASSUMPTIONS.

It is mainly assumed that the vector of disturbances are contemporaneously but not serially correlated. However, in SUR model we assume that:

1. $E(\varepsilon_i \mid X) = 0$
2. $Var(\varepsilon_i \mid X) = \sigma_1^2 I_T = \sigma_{ii} I_T \quad groupwise\ heteroscedasticity.$
3. $E(\varepsilon_{it}\varepsilon_{jt} \mid X) = \sigma_{ij} \quad comtemporaneous\ correlation.$
4. $E(\varepsilon_{it}\varepsilon_{is} \mid X) = 0 \quad (t \neq s) \quad no\ autocorrelation.$
5. $E(\varepsilon_{it}\varepsilon_{js} \mid X) = 0 \quad (t \neq s) \quad no\ time-varying\ cross\ correlation.$

However, the assumption placed on the variance-covariance matrix of the disturbance in (3.2) is that

$$E(\varepsilon\varepsilon') = \Sigma \otimes I_T \qquad (3.4)$$

In addition, the standard assumption that

$$\mathrm{E}(\varepsilon) = 0 \qquad (3.5)$$

is also maintained.

Applying the Kronecker product notation to the covariance matrix of ε , since $\varepsilon \sim N(0, \Sigma \otimes I_T)$

Then, we have that

$$E(\varepsilon\varepsilon') = \begin{bmatrix} E(\varepsilon_1\varepsilon_1') & E(\varepsilon_1\varepsilon_2')\ldots\ldots\ldots E(\varepsilon_1\varepsilon_M') \\ E(\varepsilon_2\varepsilon_1') & E(\varepsilon_2\varepsilon_2')\ldots\ldots\ldots E(\varepsilon_2\varepsilon_M') \\ \vdots & \vdots \qquad\qquad\qquad \vdots \\ E(\varepsilon_M\varepsilon_1') & E(\varepsilon_M\varepsilon_2')\ldots\ldots\ldots E(\varepsilon_M\varepsilon_M') \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{11} & \sigma_{12} \cdots\cdots\cdots \sigma_{1M} \\ \sigma_{21} & \sigma_{22} \cdots\cdots\cdots \sigma_{2M} \\ \vdots & \vdots \qquad\qquad \vdots \\ \sigma_{M1} & \sigma_{M2} \cdots\cdots\cdots \sigma_{MM} \end{bmatrix} \otimes \quad I_T$$

Multiplying variance-covariance matrix by identity matrix of order (*T x T*), we have

$$= \begin{bmatrix} \sigma_{11}I_T & \sigma_{12}I_T \cdots\cdots\cdots \sigma_{1M}I_T \\ \sigma_{21}I_T & \sigma_{22}I_T \cdots\cdots\cdots \sigma_{2M}I_T \\ \vdots & \vdots \qquad \vdots \qquad \vdots \\ \sigma_{M1}I_T & \sigma_{M2}I_T \cdots\cdots\cdots \sigma_{MM}I_T \end{bmatrix} = \Sigma \otimes I_T \qquad (3.6)$$

Where $\otimes$ denotes the matrix kronecker (tensor) product, Σ is the variance-covariance matrix of the error which is an (*M x M*) matrix and I_T is an identity matrix of order (*T x T*).

The variance-covariance matrix (Σ) is given as:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \cdots\cdots\cdots \sigma_{1M} \\ \sigma_{21} & \sigma_{22} \cdots\cdots\cdots \sigma_{2M} \\ \vdots & \vdots \qquad\qquad \vdots \\ \sigma_{M1} & \sigma_{M2} \cdots\cdots\cdots \sigma_{MM} \end{bmatrix} \qquad (3.7)$$

3.4. ESTIMATION OF PARAMETER (β).

If the *M x M* positive definite variance-covariance matrix $\Sigma \otimes I_T$ of (3.4) is denoted by Ω and all values of elements of Σ are known, then the SUR formulation of the regression models in (3.1) through (3.2) produces more efficient regression parameter estimates using Generalised Least Squares (GLS) estimator given by

$$\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y \qquad (3.8)$$

Where $\Omega^{-1} = \Sigma^{-1} \otimes I_T$

Employing the Gibbs sampling algorithm of MCMC simulation techniques in estimating the parameters and its distribution on equal observations in a three-equation SUR model, the Gibbs sampler, directly samples iteratively from all of the complete conditional posterior distribution. Using *f* (.) as generic notation for probability density function (pdf).

3.5. LIKELIHOOD FUNCTION.

The likelihood function for β and Σ can be written as

$$f\,(y\,|\,\beta,\Sigma) = (2\pi)^{-MT/2}\,|\,\Sigma\,|^{-T/2}\exp[-\frac{1}{2}(y - X\beta)'(\Sigma^{-1} \otimes I_T)(y - X\beta)] \qquad (3.9)$$

This pdf in (3.9) can also be written as

$$f\,(y\,|\,\beta,\Sigma) = (2\pi)^{-MT/2}\,|\,\Sigma\,|^{-T/2}\exp[-\frac{1}{2}tr(A\Sigma^{-1})] \qquad (3.10)$$

Where *A is* an *(M x M)* matrix with $(i, j)^{th}$ element given by $[A]_{ij} = (y_i - X_i\beta_i)'(y_j - X_j\beta_j)$

"*tr*" denotes the trace of a matrix

$|\Sigma| = \det(\Sigma)$, denotes the value of the determinant of Σ.

3.6. PRIOR.

Priors express vague or general information about a parameter (objective prior). The simplest and oldest rule for determining a non-informative prior is the *principle of indifference*, which assigns equal probabilities to all possibilities.

In absence of prior knowledge, Bayesian analysis with non-informative priors is very common in practice. One of the most widely used non-informative priors, introduced by Jeffreys (1946, 1961), is Jeffreys's invariant prior.

We considered a non-informative prior (Jeffrey's invariant prior), given as

$$f(\beta, \Sigma) = f(\beta) f(\Sigma) \propto |\Sigma|^{-(M+1)/2} \tag{3.11}$$

Which is proportional to the square root of the determinant of the Fisher information matrix.

The advantages of the use of Jeffreys's prior is that, it is invariant under any one-to-one reparameterisation of the model, and it also conveys a considerable information about Σ.

3.7. JOINT POSTERIOR PDF FOR (β, Σ)

According to Bayes' theorem, the joint posterior pdf is proportional to the product of the likelihood and the prior, that is

$$f(\theta \mid y) \propto f(y \mid \theta) \times f(\theta) \tag{3.12}$$

Where θ is the parameter of interest and y is the data, $f(\theta \mid y)$ is the joint posterior pdf, $f(y \mid \theta)$ is the likelihood function and $f(\theta)$ is the prior pdf.

Now, applying Bayes' theorem to the likelihood function in (3.9) and the prior pdf in (3.11), yields the joint posterior pdf for β and Σ, written as

$$f(\beta, \Sigma \mid y) \propto f(y \mid \beta, \Sigma) f(\beta, \Sigma)$$

$$f(\beta, \Sigma \mid y) \propto \frac{1}{(2\pi)^{\frac{MT}{2}} \mid \Sigma \mid^{\frac{T}{2}}} \exp[-\frac{1}{2}(y - X\beta)'(\Sigma^{-1} \otimes I_T)(y - X\beta)] \times \mid \Sigma \mid^{\frac{-M+1}{2}}$$

Multiplying the likelihood function with prior distribution above, we have

$$f(\beta, \Sigma \mid y) \propto \frac{1}{(2\pi)^{\frac{MT}{2}} \mid \Sigma \mid^{\frac{T+M+1}{2}}} \exp[-\frac{1}{2}(y - X\beta)'(\Sigma^{-1} \otimes I_T)(y - X\beta)]$$

In this joint posterior distribution given above, we have multiplicative constant that can be safely removed without affecting the shape of the function. Removing it, we can see that the joint posterior pdf is given as

$$f(\beta, \Sigma \mid y) \propto \mid \Sigma \mid^{-\frac{T+M+1}{2}} \exp[-\frac{1}{2}(y - X\beta)'(\Sigma^{-1} \otimes I_T)(y - X\beta)] \qquad (3.13)$$

By taking into account the properties of the determinant and trace operators, the pdf in (3.13), can be written as

$$f(\beta, \Sigma \mid y) = \mid \Sigma \mid^{-(T+M+1)/2} \exp[-\frac{1}{2} tr(A\Sigma^{-1})] \qquad (3.14)$$

3.8. CONDITIONAL POSTERIOR PDF FOR $(\beta \mid \Sigma)$

The most common MCMC algorithm is a Gibbs sampler, used to draw from the full conditional posterior pdf for $(\beta \mid \Sigma)$ and $(\Sigma \mid \beta)$.

The term in the exponent of equation (3.13), written as

$$(y - X\beta)'(\Sigma^{-1} \otimes I_T)(y - X\beta)$$

When expanded can be written as

$$(y - X\beta)'(\Sigma^{-1} \otimes I_T)(y - X\beta)$$

$$= (y - X\beta + X\hat{\beta} - X\hat{\beta})'(y - X\beta + X\hat{\beta} - X\hat{\beta})$$

$$= (y - X\hat{\beta})'(y - X\hat{\beta}) + (\beta - \hat{\beta})'X'X(\beta - \hat{\beta})(\Sigma^{-1} \otimes I_T)$$

$$= (y - X\hat{\beta})'(\Sigma^{-1} \otimes I_T)(y - X\hat{\beta}) + (\beta - \hat{\beta})'(\Sigma^{-1} \otimes I_T)(\beta - \hat{\beta}) \quad (3.15)$$

Where $\hat{\beta} = [X'(\Sigma^{-1} \otimes I_T)X]^{-1} X'(\Sigma^{-1} \otimes I_T)y$, given in (3.8) to be the generalized least squares estimator. It follows that the conditional posterior pdf for β given Σ is the multivariate normal pdf written as

$$f(\beta \mid \Sigma, y) \propto \exp[-\frac{1}{2}(\beta - \hat{\beta})'X'(\Sigma^{-1} \otimes I_T)X(\beta - \hat{\beta})] \quad (3.16)$$

3.8.1. CONDITIONAL POSTERIOR PDF FOR $(\Sigma \mid \beta)$

Then, viewing the joint posterior pdf in (3.13) as a function of only Σ yields the conditional posterior pdf for Σ given β

$$f(\beta,\Sigma \mid y) \propto |\Sigma|^{-\frac{T+M+1}{2}} \exp[-\frac{1}{2}(y - X\beta)'(\Sigma^{-1} \otimes I_T)(y - X\beta)]$$

It follows that the conditional posterior pdf for $(\Sigma \mid \beta)$ is the inverse Wishart pdf, written as

$$f(\Sigma \mid \beta, y) \propto |\Sigma|^{-\frac{(T+M+1)}{2}} \exp[-\frac{1}{2} tr(A\Sigma^{-1})] \quad (3.17)$$

3.9. MARGINAL POSTERIOR PDF FOR β.

The marginal posterior pdf for β, is obtained by combining the prior pdf in (3.11) with the likelihood function in (3.9), is written as

$$f(\beta,\Sigma \mid y) \propto |\Sigma|^{-\frac{T+M+1}{2}} \exp[-\frac{1}{2}(y - X\beta)'(\Sigma^{-1} \otimes I_T)(y - X\beta)]$$

And integrating out Σ, the marginal posterior pdf for β, is given as

$$f(\beta \mid y) = \int f(\beta,\Sigma \mid y) \partial\Sigma$$

$$f(\beta \mid y) \propto |A|^{-T/2} \quad (3.18)$$

The integral is performed using properties of the inverse Wishart distribution.

3.9.1. MARGINAL POSTERIOR PDF FOR Σ

Then, the marginal posterior pdf for Σ, is obtained by using the result in (3.15),

$$(y - X\beta)'(\Sigma^{-1} \otimes I_T)(y - X\beta)$$

$$= (y - X\hat{\beta})'(\Sigma^{-1} \otimes I_T)(y - X\hat{\beta}) + (\beta - \hat{\beta})'(\Sigma^{-1} \otimes I_T)(\beta - \hat{\beta})$$

$$f(\Sigma \mid y) = \int f(\beta, \Sigma \mid y) \partial \beta$$

And integrating out β using the properties of multivariate normal pdf, given by

$$f(\Sigma \mid y) \propto \mid X'(\Sigma^{-1} \otimes I_T) X \mid^{-1/2} \mid \Sigma \mid^{-(T+M+1)/2} \exp[-\frac{1}{2}(y - X\hat{\beta})'(\Sigma^{-1} \otimes I_T)(y - X\hat{\beta})]$$

The marginal posterior pdf for Σ, is given as

$$f(\Sigma \mid y) = \mid X'(\Sigma^{-1} \otimes I_T) X \mid^{-1/2} \mid \Sigma \mid^{-\frac{(T+M+1)}{2}} \exp[-\frac{1}{2} tr(\hat{A}\Sigma^{-1})] \qquad (3.19)$$

3.10. SIMULATION STUDIES

Our simulation work, considers a system of SUR equations containing three distinct linear regression equations with one of them having a pair of correlated covariates. Let the three equations be distributed as follows;

$(Y_1 \mid X_1) \sim N_n(X_1\beta_1, \sigma_{11}^2), (Y_2 \mid X_2) \sim N_n(X_2\beta_2, \sigma_{22}^2)$ and $(Y_3 \mid X_3) \sim N_n(X_3\beta_3, \sigma_{33}^2)$.

Thus, with M = 3, we have

$$\begin{aligned} y_1 &= X_1\beta_1 + \varepsilon_1 \\ y_2 &= X_2\beta_2 + \varepsilon_2 \\ y_3 &= X_3\beta_3 + \varepsilon_3 \end{aligned} \qquad (3.20)$$

and the whole system in (3.20) is assumed to be distributed as $(Y \mid X) \sim N_{3n}(X\beta, \Sigma \otimes I_T)$.

Therefore, the structural parameters of the SUR equations with correlated errors, for our simulation is

$$\begin{aligned} y_1 &= -2 + 0.5X_{11} + 3X_{12} + 1.8X_{13} + \varepsilon_1 \\ y_2 &= \ \ 4 + 2X_{21} + \varepsilon_2 \\ y_3 &= \ \ 0.8 - 3X_{31} + 2.4X_{33} + \varepsilon_3 \end{aligned} \tag{3.21}$$

Gibbs sampler was used to simulate the explanatory variables $X_{11}, X_{12}, X_{13}, X_{21}, X_{31}, X_{33}$ from a U (-3, 3) distribution for various sample sizes N = 50, 100, 200, 500 and 1000, using 10000 replicates and taking 10% of each sample size as the burn-in period.

When M = 3 in (3.7), Σ is then a positive definite 3 x 3 variance-covariance matrix of the errors ε_i , i = 1, 2, 3, defined by

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \tag{3.22}$$

The error terms, $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3) \sim N_{3n}(0, \Sigma \otimes I_T)$ is independently and identically distributed as defined in (3.22).

The true value of the variance-covariance (Σ) used in ε , as $\varepsilon \sim N(0,1)$ for our simulation is

$$\Sigma = \begin{pmatrix} 4.1 & 0.7 & 0.8 \\ 0.7 & 3.4 & 0.2 \\ 0.8 & 0.2 & 2.6 \end{pmatrix} \tag{3.23}$$

3.11. GIBBS SAMPLING PROCEDURE

To introduce the Gibbs sampler, consider a three-equation random variable $f(x) = f(x_1, x_2, x_3)$. We can write down the set of full conditionals as $f(x_1 | x_2, x_3)$, $f(x_2 | x_1, x_3)$, $f(x_3 | x_1, x_2)$ and can be sampled from.

Gibbs sampler generates a sequence $\{X^{(t)} : t \geq 0\}$ by iteratively sampling from the conditionals.

$$X_1^{(t)} \sim f(x_1 | X_2^{(t-1)}, X_3^{(t-1)})$$

$$X_2^{(t)} \sim f(x_2 | X_1^{(t)}, X_3^{(t-1)})$$

$$X_3^{(t)} \sim f(x_3 | X_1^{(t)}, X_2^{(t)})$$

Then Gibbs sequence now forms a Markov chain, by drawing the explanatory variables $X_{11}, X_{12}, X_{13}, X_{21}, X_{31}, X_{33}$ from a U (-3, 3) distribution.

3.11.1. GIBBS SAMPLING WITH β and Σ

We applied the Gibbs sampling algorithm in order to generate draws of β and Σ from their conditional posterior distribution. Given a starting value for β (assuming that is $\beta^{(0)}$), the j^{-th} draw from the Gibbs sampler $(\beta^{(0)}, \Sigma^{(0)})$ is obtained by simulating the next two steps.

1. Draw $\beta^{(j)}$ from $p(\beta | \Sigma^{j-1}, y)$
2. Draw $\Sigma^{(j)}$ from $p(\Sigma | \beta^{j-1}, y)$

The basic idea behind Gibbs sampling is that we simulate successively each component θ_j from its (posterior) full conditional distributions. After every sufficient draws from the conditional distributions, the Markov chain would converge to the desired posterior distribution whereas "burn-in" are discarded from the simulation because they are not from the stationary distribution of the MCMC.

CHAPTER FOUR

DATA, RESULTS AND DISCUSSION

4.1. INTRODUCTION

This chapter summarizes the results of the simulation as described in the methodology when there is an equal observation in a three-equation SUR model. Also, relevant discussions of the simulation results obtained in this work are presented.

4.2. MODEL

The model for our simulation is

$$y_1 = -2 + 0.5X_{11} + 3X_{12} + 1.8X_{13} + \varepsilon_1$$

$$y_2 = 4 + 2X_{21} + \varepsilon_2$$

$$y_3 = 0.8 - 3X_{31} + 2.4X_{33} + \varepsilon_3$$

When M = 3 in (3.7), Σ is then a positive definite 3 x 3 variance-covariance matrix of the errors ε_i , i = 1, 2, 3, defined by

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}$$

The true value of the variance-covariance (Σ) used in ε , as $\varepsilon \sim N(0,1)$ for our simulation is

$$\Sigma = \begin{pmatrix} 4.1 & 0.7 & 0.8 \\ 0.7 & 3.4 & 0.2 \\ 0.8 & 0.2 & 2.6 \end{pmatrix}$$

Using the above model, Gibbs sampler was used to simulate explanatory variables from U (-3, 3) distribution for five sample sizes N = 50, 100, 200, 500, and 1000, using 10000 replicates and 10% of each sample size as the burn-period.

4.3. TABLES OF RESULTS OBTAINED

The tables below display the results for the parameters and its distribution on equal observations in a three-equation SUR model.

Table 4.1a: *Result for Marginal Posterior pdf for* β *and* Σ ***N = 50***

Marg. Posterior pdf for β	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
	0.5493	3.2186	1.9169	1.8755	0.0396	0.2003	-2.7415	0.2672	1.8252
ABIAS	0.0493	0.2186	0.1169	0.1245	0.0396	0.2003	0.2585	0.2672	0.5748
Marg. Posterior pdf for Σ	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
	154.49	-65.43	-89.50	-65.43	83.90	-17.33	-89.50	-17.33	181.02

Table 4.1b: *Result for Conditional Posterior pdf for* $\beta \mid \Sigma$ *and* $\Sigma \mid \beta$. ***N = 50***

Cond. Posterior pdf for $\beta \mid \Sigma$	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
	0.6317	3.4148	1.4457	1.9758	1.0932	-0.0463	-3.5658	1.7937	0.6333
ABIAS	0.1317	0.4148	0.3543	0.0242	1.0932	0.0463	0.5658	1.7937	1.7667
Cond. Posterior pdf for $\Sigma \mid \beta$	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
	215.52	-62.76	-162.47	-62.47	67.31	10.65	-162.47	10.62	257.31

Table 4.1c: *Summary of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) convergence diagnostic test statistic (CD).* ***N = 50***

Beta(s)	Mean	ABIAS	95% CIs		CD
			1st Q	3rd Q	
β_{11}	0.5492	0.0492	0.1933	0.9104	-0.0130
β_{12}	3.2240	0.2240	2.8880	3.5600	-1.4107
β_{13}	1.9124	0.1124	1.5610	2.2679	-0.5269
β_{21}	1.8721	0.1279	1.5659	2.1852	-1.3402
β_{22}	0.0399	0.0399	-0.2955	0.3761	-1.0480
β_{23}	0.1984	0.1984	-0.0872	0.4824	1.4839
β_{31}	-2.7466	0.2534	-3.2599	-2.2352	-1.5055
β_{32}	0.2695	0.2695	-0.1971	0.7354	0.8109
β_{33}	1.8220	0.5780	1.2350	2.3800	0.0262

Table 4.1d: *Result of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) convergence diagnostic test statistic (CD).* ***N = 50***

sigma(s)	Mean	95%CIs 1st Q	95%CIs 3rd Q	CD
σ_{11}	168.00	142.60	188.00	2.057
σ_{12}	-71.26	-83.40	-56.28	-0.762
σ_{13}	-97.22	-114.59	-75.59	-1.626
σ_{21}	-71.26	-83.40	-56.28	-0.762
σ_{22}	91.22	77.12	102.17	0.209
σ_{23}	-18.90	-31.20	-5.76	1.438
σ_{31}	-97.22	-114.59	-75.59	-1.626
σ_{32}	-18.90	-39.20	-5.76	1.438
σ_{33}	197.00	167.44	167.44	0.2374

Table 4.2a: *Result for Marginal Posterior pdf for* β *and* Σ ***N = 100***

Marg. Posterior pdf for β	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
	0.3583	3.0222	1.9777	1.7526	-0.1663	-0.3575	-3.0690	0.0342	2.2536
ABIAS	0.1417	0.0222	0.1777	0.2474	0.1663	0.3575	0.0690	0.0342	0.1464
Marg. Posterior pdf for Σ	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
	141.21	-62.29	-29.31	-62.29	88.01	1.41	-29.31	1.41	189.59

Table 4.2b: *Result for Conditional Posterior pdf for* $\beta \mid \Sigma$ *and* $\Sigma \mid \beta$. ***N = 100***

Cond. Posterior pdf for $\beta \mid \Sigma$	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
	0.4552	3.0349	1.6477	1.5250	-0.8391	-0.4063	-2.8870	0.1984	3.2368
ABIAS	0.0448	0.0349	0.1523	0.4750	0.8391	0.4063	0.1130	0.1984	0.8368
Cond. Posterior pdf for $\Sigma \mid \beta$	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
	146.18	-56.59	-7.89	-56.59	86.39	9.67	-7.89	9.67	207.92

Table 4.2c: *Summary of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) convergence diagnostic test statistic (CD).* ***N = 100***

Beta(s)	Mean	ABIAS	95%CIs		CD
			1st Q	3rd Q	
β_{11}	0.3519	0.1481	0.0064	0.6928	0.3213
β_{12}	3.0241	0.0241	2.6727	3.3659	0.6804
β_{13}	1.9820	0.1820	1.6300	2.3452	-0.8253
β_{21}	1.7424	0.2576	1.4241	2.0687	-1.2047
β_{22}	-0.1687	0.1687	-0.5035	0.1701	-1.9358
β_{23}	-0.3600	0.3600	-0.6578	-0.0659	0.3832
β_{31}	-3.0678	0.0678	-3.5937	-2.5336	-2.0414
β_{32}	0.0314	0.0314	-0.5271	0.5868	1.7185
β_{33}	2.2622	0.1378	1.7735	2.7412	0.9538

Table 4.2d: *Result of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) convergence diagnostic test statistic (CD).* ***N = 100***

sigma(s)	Mean	95%CIs		CD
		1st Q	3rd Q	
σ_{11}	147.35	132.24	160.25	0.2412
σ_{12}	-65.06	-73.54	-55.48	1.0227
σ_{13}	-30.45	-41.69	-18.40	-1.7233
σ_{21}	-65.06	-73.54	-55.48	1.0227
σ_{22}	91.65	82.09	99.73	-0.7192
σ_{23}	1.47	-7.70	10.57	2.3579
σ_{31}	-30.45	-41.69	-18.40	-1.7233
σ_{32}	1.47	-7.70	10.57	2.3579
σ_{33}	197.80	177.80	214.60	-0.1910

Table 4.3a: *Result for Marginal Posterior pdf for* β *and* Σ ***N = 200***

Marg. Posterior pdf for β	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
	0.5431	2.9254	1.7755	2.0629	-0.1394	-0.0060	-3.0884	0.1104	2.4700
ABIAS	0.0431	0.0746	0.0245	0.0629	0.1394	0.0060	0.0884	0.1104	0.0700
Marg. Posterior pdf for Σ	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
	124.20	-69.93	48.26	-69.93	103.32	39.81	48.26	39.81	186.12

Table 4.3b: *Result for Conditional Posterior pdf for* $\beta \mid \Sigma$ *and* $\Sigma \mid \beta$. ***N = 200***

Cond. Posterior pdf for $\beta \mid \Sigma$	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
	0.8081	3.0803	2.1594	1.8913	-.0.2001	-0.2998	-2.8160	0.1588	2.7625
ABIAS	0.3081	0.0803	0.3594	0.1087	0.2001	0.2998	0.1840	0.1588	0.3625
Cond. Posterior pdf for $\Sigma \mid \beta$	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
	109.56	-59.511	46.70	-59.51	90.54	34.60	46.70	34.60	174.61

Table 4.3c: *Result of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) convergence diagnostic test statistic (CD).* ***N = 200***

Beta(s)	Mean	ABIAS	95%CIs		CD
			1st Q	3rd Q	
β_{11}	0.5423	0.0423	0.3713	0.7123	-0.3665
β_{12}	2.9280	0.0720	2.7490	3.1030	0.1410
β_{13}	1.7774	0.0226	1.5931	1.9646	0.2103
β_{21}	2.0610	0.0610	1.8810	2.2420	-0.5502
β_{22}	-0.1422	0.1422	-0.3070	0.0227	-0.9994
β_{23}	-0.0071	0.0071	-0.1816	0.1624	0.0989
β_{31}	-3.0900	0.0090	-3.3580	-2.8190	-2.3796
β_{32}	0.1056	0.1056	-0.1702	0.3791	1.2155
β_{33}	2.4657	0.0657	2.1877	2.7399	1.4205

Table 4.3d: *Result of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) convergence diagnostic test statistic (CD).* ***N = 200***

sigma(s)	Mean	95%CIs		CD
		1^{st} Q	3^{rd} Q	
σ_{11}	126.9	117.90	134.80	-0.5034
σ_{12}	-71.43	-77.61	-64.53	-0.1023
σ_{13}	49.31	41.31	56.84	-2.6559
σ_{21}	-71.43	-77.61	-64.53	-0.1023
σ_{22}	105.43	97.82	112.15	0.7045
σ_{23}	40.68	33.54	47.42	1.7807
σ_{31}	49.31	41.31	56.84	-2.6559
σ_{32}	40.68	33.54	47.42	1.7807
σ_{33}	190.10	176.80	202.0	-1.2070

Table 4.4a: *Result for Marginal Posterior pdf for β and Σ* ***N = 500***

Marg. Posterior pdf for β	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
	0.3834	3.0723	1.7771	2.0396	-0.0291	-0.1925	-3.0071	-0.0916	2.3140
ABIAS	0.1166	0.0723	0.0229	0.0396	0.0291	0.1925	0.0071	0.0916	0.0860
Marg. Posterior pdf for Σ	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
	120.96	-51.03	-13.88	-51.03	97.46	11.72	-13.88	11.72	184.68

Table 4.4b: *Result for Conditional Posterior pdf for $\beta \mid \Sigma$ and $\Sigma \mid \beta$.* ***N = 500***

Cond. Posterior pdf for $\beta \mid \Sigma$	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
	0.3161	3.6837	1.8332	2.0568	0.1349	0.1899	-2.5694	-0.3524	2.0023
ABIAS	0.1839	0.6837	0.0332	0.0568	0.1349	0.1899	0.4306	0.3524	0.3977
Cond. Posterior pdf for $\Sigma \mid \beta$	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
	117.38	-56.46	-9.90	-56.46	102.4	7.96	-9.90	7.96	180.80

Table 4.4c: *Result of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) convergence diagnostic test statistic (CD).* ***N = 500***

Beta(s)	Mean	ABIAS	95%CIs		CD
			1st Q	3rd Q	
β_{11}	0.3815	0.1185	0.2168	0.5466	-01831
β_{12}	3.0740	0.0740	2.9060	3.2390	-0.2565
β_{13}	1.7741	0.0259	1.6037	1.9475	-0.1354
β_{21}	2.0400	0.0400	1.8860	2.1950	0.3068
β_{22}	-0.0289	0.0289	-0.1756	0.1173	0.0520
β_{23}	-0.1939	0.1939	-0.3481	-0.0376	-2.8051
β_{31}	-3.0070	0.0070	-3.2390	-2.7770	-0.9068
β_{32}	-0.0909	0.0909	-0.3288	0.1455	0.8440
β_{33}	2.3100	0.0900	2.0680	2.5510	1.8665

Table 4.4d: *Result of the posterior means, the standard deviations (SDs), 95% Credible intervals (95% CIs), and Geweke's (1992) convergence diagnostic test statistic (CD). N =500*

sigma(s)	Mean	95%CIs		CD
		1st Q	3rd Q	
σ_{11}	122.00	116.66	126.95	0.5970
σ_{12}	-51.54	-55.05	-47.79	-0.1850
σ_{13}	-13.94	-18.43	-9.37	-1.5515
σ_{21}	-51.54	-55.05	-47.79	-0.1850
σ_{22}	98.27	93.93	102.35	0.4111
σ_{23}	11.89	7.82	15.95	2.1420
σ_{31}	-13.89	-18.43	-9.37	-1.5515
σ_{32}	11,89	7.82	15.95	2.1420
σ_{33}	186.30	178.10	193.90	0.1944

Table 4.5a: *Result for Marginal Posterior pdf for β and Σ* ***N = 1000***

Marg. Posterior pdf for β	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
	0.5139	2.9906	1.7713	2.0544	0.0138	-0.0846	-3.0042	0.0299	2.3890
ABIAS	0.0139	0.0094	0.0287	0.0544	0.0138	0.0846	0.0042	0.0299	0.0110
Marg. Posterior pdf for Σ	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
	119.08	-61.34	-6.80	-61.34	100.85	36.93	-6.80	36.93	178.79

Table 4.5b: *Result for Conditional Posterior pdf for $\beta \mid \Sigma$ and $\Sigma \mid \beta$.* ***N = 1000***

Cond. Posterior pdf for $\beta \mid \Sigma$	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
	0.6137	2.8260	1.9725	2.0986	-0.3530	-0.3867	-3.2216	-0.1725	2.1903
ABIAS	0.1137	0.1740	0.1725	0.0986	0.3530	0.3867	0.2216	0.1725	0.2097
Cond. Posterior pdf for $\Sigma \mid \beta$	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
	120.71	-59.99	-16.65	-59.99	95.21	37.05	-16.05	37.05	169.65

Table 4.5c: *Result of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) convergence diagnostic test statistic (CD).* ***N = 1000***

Beta(s)	Mean	ABIAS	95%CIs 1st Q	95%CIs 3rd Q	CD
β_{11}	0.5145	0.0145	0.4064	0.6243	-0.4307
β_{12}	2.9930	0.0070	2.8840	3.0990	-0.1901
β_{13}	1.7730	0.2270	1.6610	1.8860	0.7506
β_{21}	2.0540	0.0540	1.9540	2.1540	-0.9767
β_{22}	0.0143	0.0143	-0.0819	0.1111	-0.2497
β_{23}	-0.0843	0.0843	-0.1804	0.0119	-1.2824
β_{31}	-3.0040	0.0040	-3.1580	-2.8480	-1.4080
β_{32}	0.0302	0.0302	-0.1206	0.1817	-1.5094
β_{33}	2.3850	0.0150	2.2220	2.5450	1.1458

Table 4.5d: *Result of the posterior means, 95% Credible intervals (95% CIs), and Geweke's (1992) convergence diagnostic test statistic (CD).* ***N =1000***

sigma(s)	Mean	95%CIs		CD
		1st Q	3rd Q	
σ_{11}	119.6	116.00	123.00	-0.5387
σ_{12}	-61.62	-64.21	-58.84	4.0195
σ_{13}	-6.82	-9.89	-3.68	4.6099
σ_{21}	-61.62	-64.21	-58.84	4.0195
σ_{22}	101..28	98.13	104.22	4.5837
σ_{23}	37.13	34.08	40.07	4.4629
σ_{31}	-6.82	-9.89	-3.683	4.6099
σ_{32}	37.13	34.08	40.07	4.4629
σ_{33}	179.60	174.10	184.90	8.0567

4.3.1. SUMMARY OF TABLES

Table 1: *Summary of the Standard Deviations (SDs) for* $\beta's$

N	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
50	0.5308	0.5253	0.5386	0.4754	0.5099	0.4453	0.7949	0.8282	0.8539
100	0.5148	0.5050	0.5198	0.4569	0.5026	0.4185	0.7697	0.6951	0.7345
200	0.2515	0.2655	0.2781	0.2635	0.2433	0.2551	0.4008	0.4112	0.4172
500	0.2413	0.2506	0.2534	0.2232	0.2188	0.2306	0.3460	0.3500	0.3591
1000	0.1608	0.1613	0.1643	0.1456	0.1435	0.1427	0.2317	0.2259	0.2399

Table 2: *Summary of the Standard Deviations (SDs) for* $\Sigma's$

N	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
50	35.88	21.68	30.67	21.68	19.75	20.27	30.67	20.27	41.56
100	21.60	13.77	17.74	13.77	13.49	13.92	17.74	13.92	28.46
200	12.78	9.76	11.66	9.76	10.78	10.58	11.66	10.58	19.19
500	7.72	5.42	6.75	5.42	6.29	6.12	6.75	6.12	11.87
1000	5.32	4.01	4.60	4.01	4.58	4.46	4.60	4.46	8.05

Table 3: *Summary of Absolute Bias for the Marginal Posterior pdf for* β*'s*

N	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
50	0.0493	0.2186	0.1169	0.1245	0.0396	0.2003	0.2585	0.2672	0.5748
100	0.1417	0.0222	0.1777	0.2474	0.1663	0.3575	0.0690	0.0342	0.1464
200	0.0431	0.0746	0.0246	0.0629	0.1394	0.0060	0.0884	0.1104	0.0700
500	0.1166	0.0723	0.0229	0.0396	0.0291	0.1925	0.0071	0.0916	0.0860
1000	0.0139	0.0094	0.0287	0.0544	0.0138	0.0846	0.0042	0.0299	0.0110

Table 4: *Summary of Absolute Bias for the Conditional Posterior pdf for* $\beta \mid \Sigma$

N	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
50	0.1317	0.4148	0.3543	0.0242	1.0932	0.0463	0.5658	1.7937	1.7667
100	0.0448	0.0349	0.1523	0.4750	0.8391	0.4063	0.1130	0.1984	0.8368
200	0.3081	0.0803	0.3594	0.1087	0.2001	0.2998	0.1840	0.1588	0.3625
500	0.1839	0.6837	0.0332	0.0568	0.1349	0.1899	0.4306	0.3524	0.3979
1000	0.1137	0.1740	0.1725	0.0986	0.3530	0.3867	0.2216	0.1725	0.2097

Table 5: *Summary of Absolute Bias for the Posterior Means pdf for* β*'s*

N	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}	β_{31}	β_{32}	β_{33}
50	0.0492	0.2240	0.1124	0.1279	0.0399	0.1984	0.2534	0.2695	0.5780
100	0.1481	0.0241	0.1820	0.2576	0.1687	0.3600	0.0678	0.0314	0.1378
200	0.0423	0.0720	0.0226	0.0610	0.1422	0.0071	0.0090	0.1056	0.0657
500	0.1185	0.0740	0.0259	0.0400	0.0289	0.1939	0.0070	0.0909	0.0900
1000	0.0145	0.0070	0.2270	0.0540	0.0143	0.0843	0.0040	0.0302	0.0150

4.4. DISCUSSION OF RESULTS

Discussion of the simulation results obtained in this work are presented in this section.

Table 4.1a, shows the result for the marginal posterior pdf for β and Σ, for sample size N = 50. It can observed that when the sample size N = 50, β_{11} and β_{22} has the lowest ABIAS to be 0.0493 and 0.0396 respectively. Table 4.1b, shows the result for conditional posterior pdf for $\beta \mid \Sigma$ and $\Sigma | \beta$ when N = 50. It can be seen that when the sample size N = 50, β_{21} and β_{23} has the lowest ABIAS to be 0.0242 and 0.0463 respectively. Table 4.1c and d, shows the summary of the posterior means, 95% credible intervals (95% CIs) and Geweke's (1992) convergence diagnostic test statistic (CD) for β and Σ, when the sample size N = 50. It can be observed that in table 4.1c, β_{11}, β_{12}, β_{13}, β_{22}, and β_{33} are positively significant and β_{31} is negatively significant.

Table 4.2a, shows the result for marginal posterior pdf for β and Σ , for sample size N = 100. It can observed that when the sample size N = 100, β_{12} and β_{32} has the lowest ABIAS to be 0.0222 and 0.0342 respectively. Table 4.2b, shows the result for conditional posterior pdf for $\beta \mid \Sigma$ and $\Sigma | \beta$ when N = 100. It can be seen that when the sample size N = 100, β_{11} and β_{12} has the lowest ABIAS to be 0.0448 and 0.0349 respectively. Table 4.2c, showed that β_{11}, β_{12}, β_{13}, β_{21} and β_{33} are positively significant and β_{23}, β_{31} are negatively significant.

Table 4.3a, affirms that the result for marginal posterior pdf for β and Σ, when the sample size N = 200. It can be observed that β_{13} and β_{23}, has the lowest ABIAS to be 0.0245 and 0.0060 respectively. Table 4.3b, shows the result for conditional posterior pdf for $\beta \mid \Sigma$ and $\Sigma | \beta$, it can be seen that β_{12} has the lowest ABIAS to be 0.0803, when the sample size N = 200. In table 4.3c, it can also be observed that β_{11}, β_{12}, β_{13}, β_{21} and β_{33} are positively significant and β_{31} is negatively significant.

Table 4.4a, showed that in the result for marginal posterior pdf for β and Σ, sample size N = 500, it can be observed that β_{13} and β_{22} has a lower ABIAS, β_{31} has the lowest ABIAS to be 0.0071. Table 4.4b, shows the result for conditional posterior pdf for $\beta \mid \Sigma$ and $\Sigma|\beta$. It can be seen that β_{13} and β_{21} has the lowest ABIAS to be 0.0332 and 0.0568 respectively. It can be observed that in table 4.4c, β_{11}, β_{12}, β_{13}, β_{21} and β_{33} are positively significant, β_{23} and β_{31} are negatively significant. It can be observed that parameters with positive significant had a greater number than parameters with negative significant.

It can be observed that in table 4.5a, as the sample size increases N = 1000, the values of ABIAS obtained was lower when N = 1000, compare to other sample sizes. Specifically, sample sizes N = 1000 had a good number of lower ABIAS compare to other sample sizes. It was observed that the larger the sample size, the closer posterior mean of the parameter get to its true population parameter.

From the summary of the tables above, table 1, showed that in the summary of ABIAS for the marginal posterior pdf for $\beta's$, the various sample sizes N = 50, 100, 200, 500 and 1000 has a lower value of the ABIAS to be 0.0396, 0.0222, 0.0060, 0.0071 and 0.0042 respectively, it can also be observed that in table 1, sample size N = 500 and 1000 has the lowest ABIAS in β_{31} to be 0.0071 and 0.0042 respectively. Table 2, showed that in the summary of the ABIAS for conditional posterior pdf for $\beta \mid \Sigma$, the sample sizes N = 50, 100, 200, 500, and 1000 with the following parameters β_{21}, β_{12}, β_{12}, β_{13} and β_{21} has a lower value of ABIAS respectively. In the summary of ABIAS for the marginal posterior pdf for $\beta's$ and the summary of ABIAS for the conditional posterior for $\beta|\Sigma$, as the sample sizes increases, there was fluctuation in the values of ABIAS obtained.

From the tables above, tables 4.1c, 4.2c, 4.3c, 4.4c and 4.5c it can be observed that, for the various sample sizes N = 50, 100, 200, 500 and 1000, β_{31} was negatively significant for the posterior means for $\beta's$ and tables 4.1d, 4.2d, 4.3d, 4.4d and 4.5d showed that N = 50, 100, 500 and 1000, σ_{31} was negatively significant for the posterior means, expect for N =200 which can be observed to be positively significant. Tables 4.1c, 4.2c, 4.3c, 4.4c and 4.5c affirms that the values obtained for the parameters has a higher number of positive significant values of the posterior means for $\beta's$.

In table 3, the summary of the ABIAS for the posterior means for $\beta's$, it can be observed that the larger the sample size, the closer the posterior mean of the parameters gets to its true population. Table 4 and 5 which shows the summary of the standard deviations for $\beta's$ and $\Sigma's$, also affirmed that as the various sample size N = 50, 100, 200, 500 and 1000 increases, the standard deviation values obtained decreased consistently for the parameters.

CHAPTER FIVE

SUMMARY AND CONCLUSION

5.1. SUMMARY OF THE STUDY.

This study was carried out to employ the Gibbs sampling algorithm of MCMC simulation techniques in estimating the parameters and its distribution on equal observations in a three-equation SUR model, the objectives are to estimate the marginal posterior probability density function and the conditional posterior probability density function. Data used in this work was simulated from the R package using Gibbs sampling, to generate the explanatory variables from a Uniform (-3, 3) distribution for various samples sizes N = 50, 100, 200, 500 and 1000 using 10000 replicates and taking 10% of each sample size as the burn-in period. Specific values were arbitrarily assumed for the structural parameters and the true values of the variance-covariance (Σ) used as $\varepsilon \sim (0,1)$ was also assumed, as normally used in Monte Carlo simulation experiments.

5.2. CONCLUSION

It was observed that as the sample size increases, the standard deviation values obtained decreased consistently for the parameters. It can also be observed that the larger the sample size, the closer the posterior mean of the parameters get to its true population parameters.

REFERENCES

Adepoju, A. A., & Akinwumi, A. O. (2017). Effects of Atypical Observations on the Estimation of Seemingly Unrelated Regression Model, *5*(2), 30–35. https://doi.org/10.12691/jmsa-5-2-1

Ando, T. (2011). Bayesian variable selection for the seemingly unrelated regression models with a large number of predictors. *Journal of The Japan Statistical Society*, *41*(2), 187–203. Retrieved from http://www.terrapub.co.jp/journals/jjss/pdf/4102/41020187.pdf

Chib, S., & Greenberg, E. (1994). Bayes inference in regression models with ARMA (p, q) errors. *Journal of Econometrics*, *64*(1–2), 183–206. https://doi.org/10.1016/0304-4076(94)90063-9

Chib, S., & Greenberg, E. (1996). Markov Chain Monte Carlo Simulation Methods in Econometrics. *Econometric Theory*, *12*(3), 409–431. https://doi.org/10.2307/3532527

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-Based Approaches to Calculating Marginal Densities, *85*(410), 398–409.

Griffiths, W. E., & Valenzuela, M. R. (2006). Gibbs samplers for a set of seemingly unrelated regressions. *Australian and New Zealand Journal of Statistics*, *48*(3), 335–351. https://doi.org/10.1111/j.1467-842X.2006.00444.x

Huang, H. R. (2017). Bayesian analysis of the SUR Tobit model Bayesian analysis of the SUR Tobit model, *4851*(March). https://doi.org/10.1080/13504850010026069

Kmenta, J., & Gilbert, R. F. (1968). Small Sample Properties of Alternative Estimators of Seemingly Unrelated Regressions. *Journal of the American Statistical Association*, *63*(324), 1180–1200. https://doi.org/10.1080/01621459.1968.10480919

Liu, A. (2002). Efficient Estimation of Two Seemingly Unrelated Regression Equations, *456*, 445–456. https://doi.org/10.1006/jmva.2001.2026

Revankar, N. S. (1974). Some Finite Sample Results in the Context of Two Seemingly Unrelated Regression Equations. *Journal of American Statistical Association*, *69*(345), 187–190.

Richard .W.Parks. (1967). Efficient Estimation of a System of Regression Equations when Disturbances are Both Serially and Contemporaneously Correlated Author (s): Richard W . Parks Source : Journal of the American Statistical Association , Vol . 62 , No . 318 (Jun ., 1967), p, *62*(318), 500–509.

Smith, M., & Kohn, R. (2000). Nonparametric seemingly unrelated regression, *98*, 257–281.

Wang, H. (2010). Sparse seemingly unrelated regression modelling: Applications in finance and econometrics. *Computational Statistics and Data Analysis*, *54*(11), 2866–2877. https://doi.org/10.1016/j.csda.2010.03.028

Wang, J., Huda, A., Lunyak, V. V., & Jordan, I. K. (2010). A gibbs sampling strategy applied to the mapping of ambiguous short-sequence tags. *Bioinformatics*, *26*(20), 2501–2508. https://doi.org/10.1093/bioinformatics/btq460

Yahya, W. B., Adebayo, S. B., Jolayemi, E. T., Oyejola, B. A., & Sanni, O. O. M. (2008). Effects of non-orthogonality on the efficiency of seemingly unrelated regression (SUR) models. *InterStat Journal*, (May), 1–29.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias Author (s): Arnold Zellner Source : Journal of the American

Statistical Association , Vol . 57 , No . 298 (Jun ., 1962), pp . 348- Published by : American. *Journal of the American Statistical Association*, *57*(298), 348–368. https://doi.org/10.1080/01621459.1962.10480664

Zellner, A. (2010). Sampler Convergence Criteria, *90*(431), 921–927.

Zellner, A., & Ando, T. (2010). Bayesian and non-Bayesian analysis of the seemingly unrelated regression model with Student- t errors , and its application for forecasting. *International Journal of Forecasting*, *26*(2), 413–434. https://doi.org/10.1016/j.ijforecast.2009.12.012