

BAYESIAN ESTIMATION OF SIMULTANEOUS EQUATION MODEL WITH LAGGED ENDOGENOUS VARIABLES AND FIRST ORDER SERIALY CORRELATED ERRORS

ADEPOJU ADEDAYO A., ALABA OLUWAYEMISI O. AND OGUNDUNMADETAYO P.

(Received 15 August 2018; Revision Accepted 21 November 2018)

ABSTRACT

Most simultaneous equation models involve the inclusion of lagged endogenous and/or exogenous variables and sometimes it may be misleading to assume that the errors are normally distributed when in reality they exhibit functional forms that are not normal especially in practical situations. The classical methods of estimating parameters of simultaneous equation models are usually affected by the presence of autocorrelation among the error terms. Unfortunately, in practice the form of correlation between the pairs of the random deviates is unknown. In this paper classical and Bayesian methods for the estimation of simultaneous equation model with lagged endogenous variables and first order serially correlated errors are considered. The small sample properties of the methods at different levels of correlation for $\rho = 0.2, 0.5$ and 0.8 are compared. Better parameter estimates were produced by the Bayesian estimator with smaller standard errors compared to the classical method. The standard deviations of the Bayesian estimator are consistently better than those of the OLS estimator for the sample sizes considered. For example, the standard deviations of the Bayesian for b_{14} (the coefficient of the lagged endogenous variable, y_{1t-1}) when $\rho = 0.2$ at $N = 10, 15, 20$ and 25 were $0.07712781, 0.05433923, 0.03230012$ and 0.03177252 respectively while those of OLS were $0.0784732, 0.4718914, 0.05701936$ and 1.31422868 . However, when $\rho = 0.8$, the standard deviations were $0.0548055, 0.03860254, 0.02572899$ and 0.02126175 for Bayesian and $0.0562190, 0.03882345, 0.053676$ and 0.0315632 for OLS. Interestingly, notice that even at high correlation level, the estimates produced by the Bayesian method are closer to the parameter values and the standard deviations decrease as the sample size increases. Hence, the Bayesian estimation method might be a better choice when lagged endogenous variables are included in a simultaneous equation model with auto-correlated disturbances since it appeared to give better results compared to the classical approach.

KEYWORDS: Bayesian estimation, Lagged endogenous variables, Simultaneous equations, Monte-Carlo Simulation, First-order autoregressive process.

INTRODUCTION

Simultaneous equations model (SEM) is a very important field of econometrics. Haavelmo (1943) presented some important statistical implications of a linear SEM such as estimation of the stochastic equations which should not be done separately; the restrictions imposed upon the same variables by other equations ought to be taken into consideration. Simultaneous equations model could be under-identified, just-identified or over-identified, depending on how each parameter of the model uniquely contributes to the endogenous variable. The just-identified model, where the equations are exactly identified is considered in this research work. The indirect least squares method, two-stage least squares method, k-class estimators,

three-stage least squares method, full information maximum likelihood method, Jackknife instrumental variable method due to Angrist, et. al. (1999), Lancaster (2004) and Blomquist and Dahlberg (1999) methods are the well-known classical inferential approaches that have been in use. They are majorly extensions of the two basic techniques of single-equation methods, the ordinary least squares and maximum likelihood estimators. The 'true' model structure is assumed unknown, and is being estimated. However, Dreze (1962) argued that such classical inference has a shortcoming in that, the available information on parameters is ignored; for instance, it is known that the marginal propensity to consume is in the unit interval, an information that could be made use of. The Bayesian inference however combines prior information on the

Adepoju, Adedayo A, Department of Statistics, University of Ibadan, Ibadan, Nigeria
Alaba, Oluwayemisi O, Department of Statistics, University of Ibadan, Ibadan, Nigeria
Ogundunmadetayo P, Department of Statistics, University of Ibadan, Ibadan, Nigeria

parameter of interest with the likelihood function to give the posterior value. The Posterior distribution thus provides updated information on the parameter(s) under study O'Hagan (1994), Press (1969) and Koop (2003). Fair (1970) worked on various classical methods for estimating simultaneous equation models with lagged endogenous variables and first order serially correlated errors. His methods differed in the number of the disturbance process, they allowed for both spatial correlation as well as correlation across equation. They suggested computationally simple limited and full information instrumental variable estimators for the parameters of the system and gave formal large sample results. They introduced both a limited information estimator, termed the FGS2SLS estimator, and a full information estimator, termed the FGS3SLS estimator, and derived their asymptotic properties. The estimators are based on an approximation of the optimal instruments and as a result they are computationally simple even in large samples.

Fingleton and LeGallo (2008) worked on estimation methods for models including an endogenous spatial lag, additional endogenous variables due to system feedback and an autoregressive or a moving average error process. They extended Kelejian, Prucha (1998) and Fingleton and Le Gallo's (2006) feasible generalized spatial two stage least squares estimators and also considered HAC estimation in a spatial framework as suggested by Kelejian and Prucha (1999). An empirical example using real estate data illustrating the different estimators was used. The finite sample properties of the estimators were finally investigated by means of Monte Carlo simulation.

Olubusoye and Okewole (2014) considered a two-equation model containing a combination of just-identified and over-identified equations. Bayesian analysis of multi-equation econometric model as well as a Monte Carlo study carried out using WinBUGS (windows version of the software: Bayesian analysis using Gibbs Sampling). Three different variances; 10, 100 and 1000 were specified to assess the sensitivity to the prior variance specification. The result of the Monte Carlo study showed that a prior variance of 10 gave the smallest mean squared error. The kernel density plots also showed that the distribution of the posterior estimates from the prior variance of 10 was the closest to the distribution obtained theoretically.

Adepoju and Idowu (2015) considered a case in which lagged endogenous variables were included among the predetermined variables to investigate the effects lag inclusion on three simultaneous equation estimators. Six sample sizes: 20, 30, 40, 100, 500 and 1000 each replicated 1000 times were simulated using Monte Carlo method. The estimation techniques considered were: Ordinary Least Squares (OLS); Two-Stage Least Squares (2SLS) and Three-Stage Least Squares (3SLS). The estimators were then evaluated using total absolute bias, standard error, variance and root mean square error of the estimates respectively. The result showed that OLS provided the best estimates for all the cases considered followed by 2SLS. The presence of lagged variables means OLS estimator does not give a linear unbiased estimator because of its inconsistency, and as a result an instrument variable estimator might instead be used. Hence, 2SLS will be the best estimator to be used. In assessing the robustness of the

instrumental variables used. The asymptotic and small sample properties of the estimators were derived to ensure consistent estimates.

Kelejian and Prucha (2004) developed estimation theory for a simultaneous system of spatially interrelated cross-sectional equation, an extension of the widely used single equation model of Cli and Ord (1973, 1981). In modeling

estimators, two major criteria were employed, namely Akaike Information Criterion (AIC) and Bayesian Information Criterion (or Schwartz Bayesian Information Criterion (SBC)). 2SLS performed better with these criteria compared to 3SLS, making it a more robust estimator. Finally, all the estimators revealed a remarkable asymptotic pattern.

A comparative study of the classical and the Bayesian approaches is thus necessary so as to take advantage of their strengths and investigates more on possible ways of improving on their weaknesses. The need to carry out valid, generally acceptable, appropriate and convenient estimation of the SEM has brought about quite a number of researches on the classical and the Bayesian procedures.

A research carried out by Gao and Lahiri (2001) focused on weak instruments where in cases with very weak instruments, there was no estimator that was superior to another, while in the case of weak endogeneity, Zellner's MELO (Minimum expected loss), a Bayesian procedure, was the best. Their result showed that under certain scenario (See Gao and Lahiri, 2001), of all the estimators, the BMOM (Bayesian method of moments) performed best. However, Jackknife instrumental variable estimator, a classical procedure due to Angrist, Imbens and Krueger (1999) and Blomquist and Dahlberg (1999), had a poor performance throughout.

These studies reflect some Bayesian estimation methods of SEM; BMOM proposed by Zellner, the methods used by Chao and Phillips (1998), Geweke (1989, 1996), Geisser (1965), Kleibergen and Zivot (2003). Others related works on simultaneous equation problems are Zellner (1971, 1979, 1997a, 1997b), Nagar (1959 and 1969), Li and Poirier (2003), Poirier (1995) and Kleibergen and Van Dijk (1998, 2002) and Kloek and Van Dijk (1978). The major objective of this study is to compare the asymptotic behaviours of classical and Bayesian estimators at different levels of first order serially correlated errors. The rest of the paper is structured as follows: section 2 specifies the model considered in this study, section 3 discusses the classical and the Bayesian Methodologies, section 4 describes the data generation process using the Monte Carlo method and section 5 presents the results of the experiment while section 6 concludes the paper.

THE MODEL

The model is

$$y_{1t} = -a_{12}y_{2t} - b_{11}x_{1t} - b_{12}x_{2t} - b_{14}y_{1t-1} + u_{1t}$$

$$y_{2t} = -a_{21}y_{1t} - b_{22}x_{2t} - b_{23}x_{3t} - b_{25}y_{2t-1} + u_{2t} \quad \dots (1)$$

Where

$$u_{1t} = \rho_1 u_{1t-1} + V_{1t}$$

$$u_{2t} = \rho_2 u_{2t-1} + V_{2t}$$

The model in (1) is a just identified model where y_{1t} and y_{2t} are observations on two endogenous variables, y_{1t-1} and y_{2t-1} are lagged variables of y_{1t} and y_{2t} respectively. x_{1t} , x_{2t} , x_{3t} are observations on exogenous variables, v_{1t} and v_{2t} are the disturbance terms, a_{12} , b_{11} , b_{12} , b_{14} , a_{21} , b_{22} , b_{23} and b_{25} are scalar parameters.

The matrix form of our model (1) is

$$BY = X\Gamma + U \quad \dots (2)$$

$$Y = B^{-1}X\Gamma + B^{-1}U_t$$

Given that

$$\pi = B\Gamma$$

$$W_t = B^{-1}U_t$$

$$\text{Then, } Y_t = \pi X_t + W_t$$

Where

$$B = \begin{pmatrix} 1 & a_{12} \\ a_{21} & 1 \end{pmatrix}, \Gamma = \begin{pmatrix} -b_{11} & -b_{12} & 0 & -b_{14} & 0 \\ 0 & -b_{22} & -b_{23} & 0 & -b_{25} \end{pmatrix}, U = \begin{pmatrix} \rho_1 u_{1t-1} + V_{1t} \\ \rho_2 u_{2t-1} + V_{2t} \end{pmatrix}$$

Where $Y = (Y_1, Y_2)$, a $n \times 2$ matrix of observations on two endogenous variables, Γ is a matrix 2×5 of coefficients for the endogenous variables. X is $n \times 1$ vector of observations on the predetermined variables, B is a 2×2 matrix of coefficients for the predetermined variables, and $U = (U_1, U_2)$ is an $n \times 2$ matrix of random disturbance terms where n is the number of observations.

To obtain the reduced form of (1), consider the following equations;

$$y_{1t} = -a_{12}y_{2t} - b_{11}x_{1t} - b_{12}x_{2t} - b_{14}y_{1t-1} + u_{1t} \quad \dots (3)$$

$$y_{2t} = -a_{21}y_{1t} - b_{22}x_{2t} - b_{23}x_{3t} - b_{25}y_{2t-1} + u_{2t} \quad \dots (4)$$

Substituting (3) in (4)

$$y_{1t} = -a_{12}[-a_{21}y_{1t} - b_{22}x_{2t} - b_{23}x_{3t} - b_{25}y_{2t-1} + u_{2t}] - b_{11}x_{1t} - b_{12}x_{2t} - b_{14}y_{1t-1} + u_{1t}$$

$$y_{1t} = a_{12}a_{21}y_{1t} + a_{12}b_{22}x_{2t} + a_{12}b_{23}x_{3t} + a_{12}b_{25}y_{2t-1} - a_{12}u_{2t} - b_{11}x_{1t} - b_{12}x_{2t} - b_{14}y_{1t-1} + u_{1t}$$

$$y_{1t} - a_{12}a_{21}y_{1t} = [a_{12}b_{22}x_{2t} - b_{12}x_{2t}] - b_{11}x_{1t} + a_{12}b_{23}x_{3t} + a_{12}b_{25}y_{2t-1} - b_{14}y_{1t-1} + u_{1t} - a_{12}u_{2t}$$

$$y_{1t}[1 - a_{12}a_{21}] = -b_{11}x_{1t} + x_{2t}[a_{12}b_{22} - b_{12}] + a_{12}b_{23}x_{3t} + a_{12}b_{25}y_{2t-1} - b_{14}y_{1t-1} + u_{1t} - a_{12}u_{2t}$$

$$y_{1t} = \frac{1}{1 - a_{12}a_{21}} [-b_{11}x_{1t} + x_{2t}[a_{12}b_{22} - b_{12}] + a_{12}b_{23}x_{3t} + a_{12}b_{25}y_{2t-1} - b_{14}y_{1t-1} + u_{1t} - a_{12}u_{2t}] \quad \dots (5)$$

Substituting (4) in (3)

$$y_{2t} = -a_{21}[-a_{12}y_{2t} - b_{11}x_{1t} - b_{12}x_{2t} - b_{14}y_{1t-1} + u_{1t}] - b_{22}x_{2t} - b_{23}x_{3t} - b_{25}y_{2t-1} + u_{2t}$$

$$y_{2t} = a_{21}a_{12}y_{2t} + a_{21}b_{11}x_{1t} + a_{21}b_{12}x_{2t} + a_{21}b_{14}y_{1t-1} - a_{21}y_{1t} - b_{22}x_{2t} - b_{23}x_{3t} - b_{25}y_{2t-1} + u_{2t}$$

$$= a_{21}a_{12}y_{2t} + a_{21}b_{11}x_{1t} + (a_{21}b_{12} - b_{22})x_{2t} - b_{23}x_{3t} - b_{25}y_{2t-1} + a_{21}b_{14}y_{1t-1} - a_{21}y_{1t} + u_{2t}$$

$$y_{2t} - a_{21}a_{12}y_{2t} = a_{21}b_{11}x_{1t} + (a_{21}b_{12} - b_{22})x_{2t} - b_{23}x_{3t} - b_{25}y_{2t-1} + a_{21}b_{14}y_{1t-1} - a_{21}y_{1t} + u_{2t}$$

$$y_{2t}(1 - a_{21}a_{12}) = a_{21}b_{11}x_{1t} + (a_{21}b_{12} - b_{22})x_{2t} - b_{23}x_{3t} - b_{25}y_{2t-1} + a_{21}b_{14}y_{1t-1} - a_{21}y_{1t} + u_{2t}$$

$$y_{2t} = \frac{1}{1 - a_{21}a_{12}} [a_{21}b_{11}x_{1t} + (a_{21}b_{12} - b_{22})x_{2t} - b_{23}x_{3t} - b_{25}y_{2t-1} + a_{21}b_{14}y_{1t-1} - a_{21}y_{1t} + u_{2t}] \quad \dots (6)$$

(5) and (6) can be written in a linear form as

$$y_{1t} = A_1x_{1t} + B_1x_{2t} + C_1x_{3t} + D_1y_{1t-1} + E_1y_{2t-1} + \gamma_1u_{1t} + \alpha_1u_{2t} \quad \dots (7)$$

Where

$$\gamma_1 = \frac{-a_{12}}{1 - a_{12}a_{21}}$$

$$\alpha_1 = \frac{1}{1 - a_{12}a_{21}}$$

$$A_1 = \frac{-b_{11}}{1 - a_{12}a_{21}}$$

$$B_1 = \frac{a_{12}b_{22} - b_{12}}{1 - a_{12}a_{21}}$$

$$C_1 = \frac{a_{12}b_{23}}{1 - a_{12}a_{21}}$$

$$D_1 = \frac{-b_{14}}{1 - a_{12}a_{21}}$$

$$E_1 = \frac{a_{12}b_{25}}{1 - a_{12}a_{21}}$$

And

$$y_{2t} = A_2x_{1t} + B_2x_{2t} + C_2x_{3t} + D_2y_{1t-1} + E_2y_{2t-1} + \gamma_2u_{1t} + \alpha_2u_{2t} \quad \dots (8)$$

Where

$$\begin{aligned}\gamma_2 &= \frac{-a_{21}}{1 - a_{21}a_{12}} \\ \alpha_2 &= \frac{1}{1 - a_{12}a_{21}} \\ A_2 &= \frac{a_{21}b_{11}}{1 - a_{21}a_{12}} \\ B_2 &= \frac{a_{21}b_{12} - b_{22}}{1 - a_{21}a_{12}} \\ C_2 &= \frac{-b_{23}}{1 - a_{21}a_{12}} \\ D_2 &= \frac{a_{21}b_{14}}{1 - a_{21}a_{12}} \\ E_2 &= \frac{-b_{25}}{1 - a_{21}a_{12}}\end{aligned}$$

(7) and (8) are:

$$y_{1t} = A_1x_{1t} + B_1x_{2t} + C_1x_{3t} + D_1y_{1t-1} + E_1y_{2t-1} + \alpha_1u_{2t} + \gamma_1u_{1t} \quad \dots (7)$$

$$y_{2t} = A_2x_{1t} + B_2x_{2t} + C_2x_{3t} + D_2y_{1t-1} + E_2y_{2t-1} + \alpha_2u_{1t} + \gamma_2u_{2t} \quad \dots (8)$$

The two equations (7) and (8) are used to determine the values of the endogenous variables at each point in terms of the predetermined variables and structural disturbance terms.

2.0 Classical and Bayesian Methodology

3.1 Classical Methodology

Given a model

$$y^* = X^*\beta + \varepsilon^* \quad \dots(9)$$

Where $y^* = Py$, $X^* = PX$ and $\varepsilon^* = P\varepsilon$. Since Σ is a positive definite matrix, it follows that there exists an $N \times N$ matrix P such that $P\Sigma P^I = I_N$

The Ordinary Least Squares method estimates are obtained as follows,

$$\begin{aligned}\varepsilon &= y^* - X^*\beta \\ \varepsilon^I\varepsilon &= (y^* - X^*\beta)^I(y^* - X^*\beta)\end{aligned}$$

By differentiating with respect to β , we have

$$\frac{d(\varepsilon^I\varepsilon)}{d\beta} = -X^{*I}X^*\beta + X^{*I}y^*$$

Equating to zero, the estimate of β gives

$$\begin{aligned}\hat{\beta} &= (X^{*I}X^*)^{-1}X^{*I}y^* \quad \dots (10) \\ \hat{\beta} &= (x^I\Sigma^{-1}x)^{-1}x^I\Sigma^{-1}y\end{aligned}$$

3.2 Bayesian Methodology

Given a model

$$y^* = X^*\beta + \varepsilon^* \quad \dots(11)$$

Where $y^* = Py$, $X^* = PX$ and $\varepsilon^* = P\varepsilon$. Since Σ is a positive definite matrix, it follows that there exists an $N \times N$ matrix P such that $P\Sigma P^I = I_N$

The vector of disturbance terms U , is assumed to be a normal distribution with mean 0 and covariance σ^2 , $N(0, h^{-1})$ where $h = \sigma^{-2}$, the error precision.

3.2.1 The Likelihood Function

Using the definition of the multivariate normal density, the likelihood function for the model can be written as:

$$P(y|\beta, h) = \frac{h^{1/2}}{(2\pi)^{N/2}} \left\{ \Sigma^{1/2} \exp \left[-\frac{h}{2} (y^* - X^*\hat{\beta})^I \Sigma^{-1} (y^* - X^*\hat{\beta}) \right] \right\} \left\{ h^{v/2} \exp \left[-\frac{hv}{2s^2} \right] \right\} \quad \dots (12)$$

where, $v = N - k$, $\hat{\beta}_j = (X^I X^*)^{-1} X^{I*} y$ is the ordinary least squares estimator and

$$s^2 = \frac{(y^* - X^*\hat{\beta}^I)^I (y^* - X^*\hat{\beta}^I)}{v}$$

is the variance for the error (Mean Square Error).

3.2.2 The Priors for Normal Linear Regression Models (NLRM)

Since the likelihood function of models determines the structure or distribution of the prior especially for easy interpretations and computations therefore, the natural conjugate prior Normal-Gamma density is used. Thus, if we elicit a prior for β conditional on h of the form:

$$\beta|h \sim N(\beta_*, V_*) \quad \dots (13)$$

and a prior for h of the form:

$$h \sim G(s_*^{-2}, v_*) \quad \dots (14)$$

then, the joint prior for the two parameters is given as:

$$\beta h = (\beta|h \times h) \sim NG(\beta_*, V_*, s_*^{-2}, v_*) \quad \dots (12)$$

3.3 The Posterior for (NLRM)

The posterior will also have Normal-Gamma density of the form:

$$\beta, h|D \sim NG(\beta^*, V^*, s^{*-2}, v^*) \quad \dots (16)$$

Where:

$$V^* = (V_*^{-1} + hX^{*'}X^*)^{-1} \quad \dots (17)$$

$$\beta^* = V^*(V_*^{-1}\beta_* + X^{*'}X^*\hat{\beta}) \quad \dots (18)$$

$$v^* = v_* + N$$

and s^{*-2} is defined implicitly through:

$$v^*s^{*2} = v_*s_*^2 + [(y^* - X^*\hat{\beta})\Sigma^{-1}(y^* - X^*\hat{\beta})]^{-1} \quad \dots (19)$$

3.4 The Posterior Means and Variances of Parameters in NLRM

By integrating out from (16), it gives the marginal posterior distribution for β which is a multivariate t distribution given as:

$$\beta|h \sim N(\beta^*, s_*^2, V^*, v^*) \quad \dots (20)$$

The mean,

$$E(\beta|h) = \beta^* = V^*X^{*'}y^* \quad \dots (21)$$

and the variance,

$$Var(\beta|h) = \frac{v^*s^{*2}}{v^*-2}V^* \quad \dots (22)$$

while posterior for the precision given the data follows the gamma distribution:

$$h/\beta \sim G(s^{*-2}, v^*) \quad \dots (23)$$

then the mean,

$$E(h/\beta) = s^{*-2} \quad \dots (24)$$

and the variance,

$$var(h/\beta) = \frac{2s^{*-2}}{v^*} \quad \dots (25)$$

3.0 Generation of Monte Carlo Data

The main task is the generation of stochastic dependent (endogenous variables) Y_{1t} and Y_{2t} which are subsequently used in estimating the parameters of the model.

To achieve this, the following have to be assumed

- (i) Values of the predetermined variables X_{1t}, X_{2t}, X_{3t} ($t = 1, 2, 3, \dots, T$)
- (ii) Values of the parameters $a_{12}, a_{21}, b_{11}, b_{12}, b_{14}, b_{22}, b_{23}, b_{25}$.

The simulation of the error term U_{1t} ($t = 1, 2, 3, \dots, T$) is the important in generating stochastic dependent variables. To set up the Monte Carlo experiment, we proceed as follows:

- (i) The sample size N is specified as $N = 10, 15, 20, 25$.
- (ii) Numerical values are assigned to each of the structural parameters as follows;

$$a_{12} = 0.5, b_{11} = 1.5, b_{12} = 1.5, b_{14} = 1.5, a_{21} = 0.5, b_{22} = 1.5, b_{23} = 0.5, b_{25} = 0.5$$

$\rho_1 = \rho_2 = 0.2, 0.5$ and 0.8 for all cases

- (iii) X_{1t}, X_{2t}, X_{3t} are generated from a uniform(0,1), Kmenta (1971)

(3) Hyper parameters assumed for prior distribution of the Bayesian approach are:

$$h = 20, \beta_* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, V_* = 0_{4 \times 4}, s_*^{-2} = \frac{1}{h}, v_* = 1$$

4.1 Generation of Random Disturbance Terms, U

This is the method of generating a random disturbance terms U_{1t} , and U_{2t} .

Recall that

$$u_{1t} = \rho_1 u_{1t-1} + V_{1t}$$

$$u_{2t} = \rho_2 u_{2t-1} + V_{2t}$$

4.1.1 Generation of Random Disturbance Terms, V

This is the method of generating a random disturbance terms V_{1t} , and V_{2t} .

Σ is the variance-covariance matrix ($V_t V_t'$) decomposed into the non-singular upper triangular matrix P_1 and lower triangular matrix P_2 .

$$\Sigma = P_1 P_2 = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 4.0 & 1.5 \\ 1.5 & 3.0 \end{pmatrix}$$

Where $Var(V_{1t}) = \sigma_{11}$, $Var(V_{2t}) = \sigma_{22}$ and $cov(V_{1t}V_{2t}) = \sigma_{12} = \sigma_{21}$

Let $P_1 = \begin{pmatrix} \rho_{11} & \rho_{12} \\ 0 & \rho_{22} \end{pmatrix}$ be upper triangular matrix and

$P_2 = \begin{pmatrix} \rho_{11} & 0 \\ \rho_{21} & \rho_{22} \end{pmatrix}$ be lower triangular matrix.

Now

$$\Sigma = P_1 P_2 = \begin{pmatrix} \rho_{11} & \rho_{12} \\ 0 & \rho_{22} \end{pmatrix} \begin{pmatrix} \rho_{11} & 0 \\ \rho_{21} & \rho_{22} \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 4.0 & 1.5 \\ 1.5 & 3.0 \end{pmatrix}$$

$$\rho_{11}^2 + \rho_{12}\rho_{21} = 4.0$$

$$\rho_{12}\rho_{22} = 1.5$$

$$\rho_{22}\rho_{21} = 1.5$$

$$\rho_{22}^2 = 3.0$$

Solving the above four equations;

$$\rho_{11} = 1.802776$$

$$\rho_{12} = 0.8660254$$

$$\rho_{21} = 0.8660254$$

$$\rho_{22} = 1.732051$$

Thus, the pair of standard deviates can be transformed into pair of random normal variables using

$$\begin{bmatrix} V_{1t} \\ V_{2t} \end{bmatrix} = P_1 \varepsilon_{1t} = \begin{pmatrix} \rho_{11} & \rho_{12} \\ 0 & \rho_{22} \end{pmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} = \begin{pmatrix} 1.802776 & 0.8660254 \\ 0 & 1.732051 \end{pmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

$$V_{1t} = 1.802776\varepsilon_{1t} + 0.8660254\varepsilon_{2t}$$

$$V_{2t} = 1.732051\varepsilon_{2t}$$

Making use of the data generated, the parameter estimates are obtained using the classical and the Bayesian methods. The following criteria are used in comparing the two estimation methods: Mean and Mean Squared Error (MSE) of the estimates.

4.2 Generation of Endogenous Variables

With the numerical values already assigned to the structural parameters, we have all the values required for the generation of the endogenous variables. Considering the AR(1) for the disturbance terms U_{1t} , and U_{2t} given below:

$$u_{1t} = \rho_1 u_{1t-1} + 1.802776\varepsilon_{1t} + 0.8660254\varepsilon_{2t}$$

$$u_{2t} = \rho_2 u_{2t-1} + 1.732051\varepsilon_{2t}$$

Solving y_{1t} and y_{2t} given in equations (7) and (8) above after estimating the resolved parameters with the assumed structural parameters, then we have

$$y_{1t} = A_1 x_{1t} + B_1 x_{2t} + C_1 x_{3t} + D_1 y_{1t-1} + E_1 y_{2t-1} + \alpha_1 u_{2t} + \gamma_1 (\rho_1 u_{1t-1} + 1.802776\varepsilon_{1t} + 0.8660254\varepsilon_{2t}) \dots (26)$$

$$y_{2t} = A_2 x_{1t} + B_2 x_{2t} + C_2 x_{3t} + D_2 y_{1t-1} + E_2 y_{2t-1} + \alpha_2 u_{1t} + \gamma_2 (\rho_2 u_{2t-1} + 1.732051\varepsilon_{2t}) \dots (27)$$

Where e_{1t} and e_{2t} are generated from $N(0,1)$. The estimation of the disturbance terms U_{1t} , and U_{2t} are generated recursively using the given model above.

Estimation of the structural parameters

Having generated the values of exogenous and endogenous variables as well as disturbance terms, the next step is the estimation of the structural parameters for 0.2, 0.5 and 0.8 correlation levels. These structural parameters are obtained using statistical software called R (version 3.3.3).

Criteria for evaluating the performance of the estimators

For easy comparison and evaluation of the performance of the estimators, the following criteria are used.

(i) Mean values of the parameter estimates.

(ii) Standard deviation

The following classical estimation methods used in this study are: OLS and Bayesian approach.

Result and Discussion

The point estimates' summary presented in Tables 1-3 reflect some properties of the two estimation methods under discussion. Interestingly, the results of the OLS and Bayesian inference are not too different. Table 1 gives the estimates of the two estimators for $\rho = 0.2$ and sample sizes 10, 15, 20 and 25. In all cases, the standard deviations of the parameter estimates decrease as the sample size increases. However, the Bayesian method performs better than the OLS since it produces smaller standard deviations. The observations in Table 1 are similar to those recorded in Table 2 when $\rho = 0.5$, the standard deviations also are smaller for both estimators. Interestingly, the parameter estimates

and the standard deviations are closer for the two estimators. In Table 3 for $\rho = 0.8$, the estimates are much closer to the parameter values than for the other cases considered. The standard deviations produced are also better for Bayesian than for the OLS estimator but closer.

Clearly, as the sample size increases, standard errors of the estimators decrease. In Table 3, where the rho of the error term (e_t) was raised to 0.8, the estimates from the two methods were more concentrated around the class containing the true value than in the first and second Tables where the rho(s) were 0.2 and 0.5 respectively. This is an indication that the distribution of the exogenous variables also affects the properties of the estimators. We noticed that

in Tables 1 and 2, the standard deviation was questionably large for the classical method when N=20, this is as a result of the characteristic of the Bayesian method.

Table 1: Estimates and the standard deviations (in parenthesis) of the estimators for $\rho = 0.2$ at different sample sizes

$\rho = 0.2$ OLS								
	$\alpha_{12} = 0.5$	$b_{14} = 1.5$	$b_{11} = 1.5$	$b_{12} = 1.5$	$\alpha_{21} = 0.5$	$b_{25} = 0.5$	$b_{22} = 1.5$	$b_{23} = 0.5$
10	1.1421583 (0.0572104)	1.5029686 (0.0784732)	1.6644831 (0.3278941)	2.219356 (0.5451088)	0.3208768 (0.01090927)	0.8830466 (0.05444762)	2.0694503 (0.90473288)	0.9273410 (1.9018776)
15	0.3160969 (0.2292887)	1.7611213 (0.4718914)	1.4866595 (0.5300868)	1.5144441 (0.1529465)	0.7042131 (0.06525014)	0.1963875 (0.75866131)	1.1175546 (0.9176295)	0.7226590 (0.6401652)
20	0.41211335 (0.027861580)	1.6257235 (0.05701936)	1.8311398 (2.002614059)	2.3247171 (1.75783801)	0.56933820 (0.007979349)	0.26224965 (0.09389295)	1.33332337 (1.26843815)	0.4862227 (1.39211352)
25	0.7180436 (0.6298178)	1.1882123 (1.31422868)	2.0203616 (3.4520269)	1.2963440 (0.1791496)	0.51686675 (0.03141017)	0.44212803 (0.3696435)	0.85574878 (0.92431707)	0.6532876 (0.95541019)

$\rho = 0.2$ Bayesian								
	$\alpha_{12} = 0.5$	$b_{14} = 1.5$	$b_{11} = 1.5$	$b_{12} = 1.5$	$\alpha_{21} = 0.5$	$b_{25} = 0.5$	$b_{22} = 1.5$	$b_{23} = 0.5$
10	1.1434892 (0.05451088)	0.5014273 (0.07712781)	1.6648094 (0.35056475)	2.2120594 (0.41962164)	0.3205546 (0.0365487)	0.8822516 (0.08139063)	2.0776748 (0.33400470)	0.9137493 (0.47789115)
15	0.3152768 (0.03788361)	1.76422847 (0.05433923)	1.4900361 (0.20219549)	1.5124157 (0.16645973)	0.7047816 (0.01747366)	0.1983340 (0.05958626)	1.1316252 (0.29448303)	0.7074965 (0.33775360)
20	0.4124684 (0.02257865)	1.6252441 (0.03230012)	1.8359105 (0.19841870)	2.3157619 (0.1873965)	0.5691561 (0.01091812)	0.2628744 (0.03745233)	1.3290222 (0.13934922)	0.4475114 (0.14427824)
25	0.7167002 (0.02221076)	1.190095 (0.03177252)	2.017005 (0.1348537)	1.298164 (0.13951188)	0.61202877 (0.009173145)	0.11566743 (0.03332095)	1.9126947 (0.1627291)	0.8835209 (0.2253054)

Table 2: Estimates and the standard deviations (in parenthesis) of the estimators for $\rho = 0.5$ at different sample sizes

$\rho = 0.5$ OLS								
	$\alpha_{12} = 0.5$	$b_{14} = 1.5$	$b_{11} = 1.5$	$b_{12} = 1.5$	$\alpha_{21} = 0.5$	$b_{25} = 0.5$	$b_{22} = 1.5$	$b_{23} = 0.5$
10	1.1761813 (0.0441453)	1.4410236 (0.0615117)	0.9147995 (0.29726951)	2.5023112 (0.36889843)	0.4511492 (0.0003524065)	0.5383666 (0.000352406)	0.5383666 (0.002048201)	2.0958359 (0.02560162)
15	0.09102833 (0.0365743)	2.08982007 (0.051209)	1.4082559 (0.263256)	1.27491456 (0.24102345)	0.8319344 (0.05624101)	0.6324357 (0.6524169)	1.2125793 (0.18471546)	0.7570141 (0.2183754)
20	0.7318313 (0.02206759)	1.1682133 (0.0315613)	1.4259474 (0.14660461)	0.6944928 (0.1711234)	0.7746194 (0.009730889)	0.4421004 (0.11482429)	1.7489464 (3.576563002)	0.1959167 (3.11130968)
25	0.5103722 (0.01912345)	1.4851723 (0.0281001)	1.2524993 (0.18352134)	0.3118128 (0.14102231)	0.5017188 (0.006005658)	0.4941017 (0.07067673)	2.1125059 (2.4343532)	0.3621733 (2.67647333)

		$\rho = 0.5$ BAYESIAN						
	$\alpha_{12} = 0.5$	$b_{14} = 1.5$	$b_{11} = 1.5$	$b_{12} = 1.5$	$\alpha_{21} = 0.5$	$b_{25} = 0.5$	$b_{22} = 1.5$	$b_{23} = 0.5$
10	1.17714935 (0.0453265 3)	1.4403528 1 (0.061511 71)	0.91902185 (0.29726951)	2.49827389 (0.3799050 2)	0.4502320 0 (0.034131 34)	0.53957013 (0.0825657 0)	2.00089610 (0.2950760 9)	0.88764138 (0.42661995)
15	0.09056807 (0.0302833 0)	2.0904888 8 (0.043568 33)	1.41928099 (0.26928842)	1.27990678 (0.2378773 8)	0.8321559 (0.015672 61)	0.6331881 (0.0533835 2)	1.2019080 (0.2924038 0)	0.7706201 (0.31745358)
20	0.7326059 (0.0220675 9)	1.1671041 (0.031561 30)	1.4292734 (0.14660461)	1.6983903 (0.1750753 0)	0.7750681 (0.017800 7)	0.4436400 (0.0369822 7)	1.7559164 (0.2094849 2)	1.1877835 (0.2017413)
25	0.5112107 (0.0193439 6)	1.4839758 (0.027671 56)	1.2555160 (0.18387260)	1.3102501 (0.1398226 9)	0.5017090 (0.007256 89)	0.4941637 (0.0072565 89)	0.4941637 (0.0248937 60)	2.1148511 (0.14645878 3)

Table 3: Estimates and the standard deviations (in parenthesis) of the estimators for $\rho = 0.8$ at different sample sizes

		$\rho = 0.8$ OLS						
	$\alpha_{12} = 0.5$	$b_{14} = 1.5$	$b_{11} = 1.5$	$b_{12} = 1.5$	$\alpha_{21} = 0.5$	$b_{25} = 0.5$	$b_{22} = 1.5$	$b_{23} = 0.5$
1 0	1.1281097 (0.052104 2)	1.5208754 (0.0562190)	1.8051547 (0.2711123)	1.197716 (0.3432230 1)	0.5057598 (0.1134199)	0.1273185 (0.9435939)	2.9259913 (1.8186495)	1.2880730 (0.7568514)
1 5	0.0963155 8 (0.027123)	2.0811050 6 (0.0388234 5)	0.5888471 2 (0.3412390)	1.4897914 3 (0.2621345)	0.4793415 (0.01632354 6)	0.5697210 (0.1936433)	2.2842028 (3.7876815)	0.8204916 (2.04389767 7)
2 0	0.602517 (0.126762)	1.362004 (0.053676)	1.379061 (0.1823565)	1.5232258 (0.1712343 6)	0.62130628 (0.01220953)	0.8387021 (0.1436817 6)	2.1950165 0 (3.7071736 1)	1.43993070 (4.26144596)
2 5	0.5198012 (0.033672)	1.4716723 (0.0315632)	1.4049989 (0.1999123 5)	1.9616000 (0.1523100 1)	0.7959777 (0.09455589)	0.5153297 (1.1127633)	1.5191538 (2.3500002 7)	0.5472089 (3.50871302)

		$\rho = 0.8$ BAYESIAN						
	$\alpha_{12} = 0.5$	$b_{14} = 1.5$	$b_{11} = 1.5$	$b_{12} = 1.5$	$\alpha_{21} = 0.5$	$b_{25} = 0.5$	$b_{22} = 1.5$	$b_{23} = 0.5$
1 0	1.1284892 (0.0433065 3)	1.5214703 (0.0548055)	1.7979201 (0.2622692 3)	1.1956277 (0.3528980 6)	0.5050173 (0.0436710 2)	0.1279974 (0.1274456)	2.9187457 (0.4787587 5)	1.2932701 (0.4159199 7)
1 5	0.09528057 (0.0268096 9)	2.08260447 (0.0386025 4)	1.61869281 (0.3377917 0)	1.51750440 (0.2544058 8)	0.4790739 (0.0127669 4)	0.5706361 (0.0439680 7)	2.2899869 (0.1957923 2)	0.8177730 (0.1445121 4)
2 0	0.6028221 (0.0179524 9)	1.3615206 (0.0257289 9)	1.3817440 (0.1528073 7)	1.5190742 (0.1702161 0)	0.62114229 (0.1097389)	0.8443285 (0.0376459 6)	2.19989414 (0.1953736 4)	1.44829506 (0.2106177 5)
2 5	0.520191 (0.0148631 5)	1.471090 (0.0212617 5)	1.404523 (0.1895278 1)	1.963403 (0.1422138 0)	0.7956012 (0.0072245 4)	0.5140112 (0.0247837 9)	1.5248470 (0.1139094)	0.5464498 (0.1381035)

DISCUSSION

Tables 1-3 presented above reflect some properties of the two estimation methods under discussion. For the classical method, the OLS was considered because they give the same estimate for this model being a just-identified model. For all the sample sizes considered under different correlation levels, the Bayesian estimates performed better than the classical estimates mostly for the small sample cases.

The performance of the two estimators was considered for a two-equation simultaneous model with lagged endogenous variables and auto-correlated errors. This study is similar to the work of Fair (1970), though Fair used classical approaches with different instrumental variables and merely derived asymptotic and small sample properties of the estimators without estimating the parameters of the model but in this work, Bayesian approach is used to estimate the parameters of the simultaneous equation with lagged endogenous variables and first serially correlated error terms and the results compared with the classical method.

The parameters of equations 1 and 2 are assumed to be $\alpha_{12} = 0.5, b_{14} = 1.5, b_{11} = 1.5, b_{12} = 1.5$ and $\alpha_{21} = 0.5, b_{25} = 0.5, b_{22} = 1.5, b_{23} = 0.5$ respectively. Auto-correlated error was set at $\rho = 0.2, 0.5$ and 0.8 (low, moderate and high correlation levels). The distribution of these estimates was closer for the two estimation methods. As expected, as the sample size increases, the standard error of the estimators reduces. At $\rho = 0.2$, the estimates from the Bayesian method was more concentrated around the true value than in the classical method considered, this is an indication that the lag of the exogenous variables also affects the properties of the estimators. It is noticed that at $\rho = 0.5$, which implies moderated auto-correlated level, the standard deviations are questionably large for the classical method when $N=20$ and $N=25$, this is as a result of lagged endogenous variables that are uncharacteristic of the Bayesian method.

CONCLUSION

Estimation of simultaneous equation model in econometric research should be approached with care. The choice of estimation method as observed in this research work affects the estimates in terms of bias and consistency especially when dealing with small samples. The Bayesian estimation method has gained a lot of attention recently which makes practical statistical inference more interesting (Lahiri et. al. (2000), Herman and Dijk (2002)). The present study showed that the Bayesian estimation method performs better than the classical for smaller sample sizes but it becomes simply impossible to estimate the parameters when the sample size is large. The model in this study is a just identified model, and the classical estimation method give relatively close estimates to the assumed parameter values. The Bayesian estimation method being more easily applied might be a better choice since it appears to give better results compared to the classical approach.

REFERENCES

- Adepoju A.A. and Idowu A.O., 2015. Estimation of Simultaneous Equation Model Involving Lagged Endogenous Variables. *Journal of the Nigerian Statistical Association* Vol. 27:89-105.
- Angrist, J.D., Imbens, G. W. and Krueger, A. B., 1999. Jackknife Instrumental Variable Estimation. *Journal of Applied Econometrics*, Vol. 14, No. 1, January/February, pp.57 – 67.
- Blomquist, S. and Dahlberg, M., 1999. Small Sample Properties of LIML and Jackknife IV Estimators: Experiments with weak Instruments. *Journal of Applied Econometrics*. Vol. 14, No. 1, pp.69– 88.
- Chao, J. C. and Phillips, P. C. B., 1998. Posterior Distributions in Limited Information Analysis of the Simultaneous Equations Model using the Jeffrey's Prior. *Journal of Econometrics*, Elsevier, Vol. 87(1), pp.49 – 86, August.
- Fair, R. C., 1970. Estimations of Simultaneous Equation Models with Lagged Endogenous Variables and First Order Serially Correlated Errors. *Econometrica*, Vol. 38, No. 2. Pp.507 – 516.
- Fingleton, B. and Le Gallo, J., 2006. Finite Sample Properties of Estimators of Spatial Models with Autoregressive, or Moving Average Disturbances and System Feedback, paper presented at the ADRES International Conference, CREUSET, Universite de Saint-Etienne (France).
- Fingleton, B. and Le Gallo, J., 2008. Estimating Spatial Models with Endogenous Variables, a Spatial Lag and Spatially Dependent Disturbances: Finite Sample Properties. *Papers in Regional Science*. Wiley Blackwell, Vol. 87 (3), 319 – 339.
- Gao, C. and Lahiri, K., 2001. A comparison of some recent Bayesian and Classical procedures for simultaneous equation models with weak instrument. Discussion paper, State University of New York at Albany.
- Geisser, S., 1965. A Bayes Approach for Combining Correlated Estimates. *J. Am. Statist. Assoc.*, 60, pp.602 –607.
- Geweke, J., 1989. Bayesian Inference in Econometric Models using Monte Carlo Integration. *Econometrica*, Vol. 57, No. 6, pp.1317-1339.
- Geweke, J., 1996. Bayesian Reduced Rank Regression in Econometrics. *Journal of Econometrics*, Vol. 75, issue 1, pp.121-146.
- Gilks, W. R., Richardson, S. and Spiegelhalter, D. J., 1996. *Markov Chain Monte Carlo in Practice*. Chapman and Hall/CRC.

- Haavelmo, T., 1943. The Statistical Implications of a System of Simultaneous Equations. *Econometrica*, Vol. 11, No. 1, pp.1-12.
- Herman K. and Dijk, V., 2002. On Bayesian Structural Inference in a Simultaneous Equation Model, *Econometric Institute Report EI 2002-10*
- Kelejian, H. H., and Prucha, I. R., 1998: A Generalized Spatial Two Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances. *Journal of Real Estate Finance and Economics*, Vol. 17, pp. 99 – 121.
- Kelejian, H. H., and Prucha, I. R., 1999: A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review*, Vol. 40, pp. 509 – 533.
- Kelejian, H. H., and Prucha, I. R., 2004: Estimation of Simultaneous Systems of Spatially Interrelated Cross Sectional Equations. *Journal of Econometrics*, Vol. 118, pp. 27 – 50.
- Kleibergen, F. and Zivot E., 2003. Bayesian and Classical Approaches to Instrumental Variable Regression. *Journal of Econometrics*, Vol. 114, pp.29-72.
- Kleibergen, F. and Van Dijk, H. K., 1998. Bayesian Simultaneous Equation Analysis using Reduced Rank Structures. *Econometric Theory*, 14, pp.701-743.
- Kloek, T. and Van Dijk, H. K., 1978. Bayesian Estimates of Equation System Parameters, An Application of Integration by Monte Carlo. *Econometrica*, Vol. 46, No. 1, pp.1-19.
- Kmenta, J., 1971: *Elements of Econometrics*. Macmillan publishing Co., Inc. New York.
- Koop, G., 2003 *Bayesian Econometrics*. New York: John Wiley & Sons.
- Lahiri, K. and Gao, C., 2000. A Comparison of Some Recent Bayesian and Classical Procedures for Simultaneous Equation Models with Weak Instruments, Department of Economics, State University of New York at Albany, Albany, NY 12222.
- Lancaster, T., 2004. *An Introduction to Modern Bayesian Econometrics*. Blackwell Publishing Ltd.
- Li, K. and Poirier, D. J., 2003. Bayesian Analysis of an Econometric model of birth inputs and outputs. *Journal of Population Economics*, Vol. 16, pp.597-625.
- Nagar, A. L., 1959. A Monte Carlo Study of Alternative Simultaneous equations Estimators. *Econometrica*, 28, pp.573-590.
- Nagar, A. L., 1969. Stochastic Simulation of d-brookings Econometric Model, in *The Brookings Model: some further results*, ed. J.S. Duesenberry, G.Fromm. L.R.Klein and E. Kuh. Chicago: Rand McNally.
- O'Hagan, T., 1994. *Kendall's Advanced Theory of Statistics. Volume 2B- Bayesian Inference*. Edward Arnold, London.
- Olubusoye O.E. and Okewole D.M., 2014. The Bayesian Approach to Multi-equation Econometric Model Estimation. *Journal of Statistical and Econometric Methods*, Vol.3, No.1: Scienpress Ltd.
- Poirier, D. J., 1995. *Intermediate Statistics and Econometrics: A comparative Approach*. MIT Press.
- Dreze J., 1962. *The Bayesian Approach to Simultaneous Equations Estimation*. Research Memorandum No. 67, Technological Institute, Northwestern University.
- Press, S. J., 1969. The t- Ratio Distribution. *Journal of the American Statistical Association*, Vol. 64. pp. 242-252.
- Van Dijk, H. K., 2002. On Bayesian Structural Inference in a Simultaneous Equation Models. *Economic institute report, EI 2002-10*.
- Zellner, A., 1971. *An Introduction to Bayesian Inference in Econometrics*. John Wiley & sons, Inc. pp.225-287.
- Zellner, A., 1979. Statistical Analysis of Econometrics models. *Journal of the American Statistical Association*. Vol. 74, No. 367, Invited Paper, pp.628 – 643.
- Zellner, A., 1997a. *Bayesian Analysis in Econometrics and Statistics: The Zellner view and papers*. Edward Elgar Publishing, pp.250-305.
- Zellner, A., 1997b. The Bayesian Method of Moments. *Advances in Econometrics*, Vol. 12, pp.85- 105, Greenwich, CT: JAI Press.

