ISSN 0331 – 9504

Journal of the Nigerian Statistical Association

Volume 25, 2013

TERTIARY EDUCATION TRUST FUND

Supported by the Tertiary Education Trust Fund (**TETFUND**)

Bayesian Estimation of an Over-identified Multi-equation Model in the Presence of Multicollinearity

Olusanya Elisa Olubusoye[1] and Dorcas Modupe Okewole[2]
[1]Department of Statistics, University of Ibadan (busoye2001@yahoo.com)
[2]Department of Mathematical Sciences, Redeemer's University,(dupedepo@yahoo.com

Abstract

Multi-equation systems have wide applications in modeling Economic issues. The Bayesian approach received very little attention in the past but is now gaining popularity with extensive application to areas hitherto handled by the classical method. The increasing interest is as a result of availability of numerical intensive software capable of solving intractable or complex numerical integration and other mathematical or computational difficulties. Violations of the assumptions underlying the models often arise in actual observed data. Multicollinearity is one of such violations which several researches have shown classical estimation approaches to be sensitive to. Studies on the performance of the Bayesian approach to such violations are however limited. This paper presents a Monte Carlo study of the Bayesian approach to multi-equation models estimation in the presence of multicollinearity. The mean, bias and MSE were used to compare the performance of the Bayesian approach to that of some classical approaches. A number of research scenarios were specified depicting presence and absence of multicollinearity. MSE from the scenario representing absence of multicollinearity was smaller than that from the scenario representing presence of multicollinearity. Results from the Bayesian approach in run 1 (representing presence of multicollinearity) showed that MSE for β_{22} (one of the correlated exogenous variables) are 0.2825, 0.1128, 0.1079 and 0.0649 for sample sizes 20, 40, 60 and 100 respectively, whereas, they were 0.2503, 0.0642, 0.0406 and 0.0414 in the absence of multicollinearity represented by run 2. MSE for β_{22} from the classical approach were 0.4230, 0.1583, 0.1498 and 0.0897 for sample sizes 20, 40, 60 and 100 respectively, whereas, they were 0.3639, 0.0837, 0.0517 and 0.0540 in run 2. MSE from the Bayesian approach were smaller than those from the classical approach. The results showed that the Bayesian approach is less sensitive to multicollinearity in estimating the coefficients of exogenous variables of over-identified model.

Key Words: Multi-equation model, Bayesian approach, Multicollinearity, Classical approach and Monte Carlo.

1.0 Introduction

Studies and analyses on relationships of variables often require the use of equations (models), some of which could be simple linear and nonlinear, while others are multiple, also linear and nonlinear, depending on the variables involved and the type of relationship connecting them. The linear multi-equation systems are of concern in this paper with particular focus on the Bayesian estimation approach.

One of the assumptions of regression models is that explanatory variables should not be highly correlated among themselves. The most commonly used estimation methods; least squares and maximum likelihood, are known to be sensitive to violation of this assumption, a condition generally referred to as Multicollinearity. The presence of highly correlated regressors in the model results in large variance and covariance which consequently leads to difficulty in making precise estimation. This article was motivated by the need for information on the performance of Bayesian estimation approach in this context which is hitherto limited.

The Bayesian approach is a statistical procedure that enables one to make probability statements about parameters of interest. It specifies prior information on the parameter in form of probability density function and represents observed data with a likelihood function. These two are then combined through the use of conditional probability in the Bayes' theorem to obtain the new information also represented by a probability density function. The results of the Bayesian analysis was compared in this work with those from some classical estimators: Ordinary least squares (OLS), Two stage least squares (2SLS), Three stage least squares (3SLS) and Limited information maximum likelihood (LIML) estimators.

Several literatures are available on studies about the classical approaches in estimation of multi-equation models in the presence of multicollinearity. However, although Bayesian approach to statistical inference has gained some attention of recent, literatures on its performance in the presence of multicollinearity are scarce. Some literatures on classical approach for estimating multi-equation models are discussed here.

Blomquist and Dahlberg (1999) studied the performance of 2SLS, LIML and four other new Jacknife IV estimators when the instruments are weak. The conclusion was that when instruments are weak, there was no easy way of obtaining reliable estimates in small samples; the only way to increase precision was to make use of better instruments and also use large samples. Olubusoye (2001) carried out a study on the extent of the effect of inadvertent use of non-mutually independent normal deviates in earlier studies on the performance of estimators of simultaneous equations models. Three arbitrarily fixed levels of correlation of the normal deviates were used; the negatively highly correlated, the feebly (positively or negatively) correlated, and the positively highly correlated. He suggested that in simultaneous-equation based comparison of the performance of estimators; only feebly correlated pairs of normal deviates should be used.

On the performance of classical estimators in the presence of multicollinearity, Oduntan (2004) presented a study considering a purely just-identified simultaneous equations model. His result showed that when multicollinearity exists, whether low or high, ILS and OLS had better performance, while others performed poorly. Another similar work was carried out by Agunbiade (2008), an extension of the research by Oduntan (2004). This work suggested that LIML, 2SLS and ILS were best for estimating parameters of a model having the relatively highly negative correlation level of multicollinearity, while OLS, unlike in Oduntan (2004), performed poorly under this scenario but performed best in the relatively highly positive correlation level of multicollinearity.

Literatures on the Bayesian approach to estimation of multi-equation Econometric models include its theories, justification for use, as well as issues on how to represent prior belief on parameters of interest. Zellner (1965) is one of the earliest papers advocating the use of Bayesian inference in Econometric models. Posterior distributions of parameters of general simultaneous equations model (SEM) were presented as well as results of some Monte Carlo experiments. The result suggested that in small samples, the Bayesian estimates are more highly concentrated about the true value of the coefficient being estimated. In Chao and Phillips (1998), the focus was on consequences of the use of Jeffreys prior in Bayesian limited information analysis of the SEM. Their results showed that Jeffreys prior brings Bayesian inference closer to classical inference because, the posterior obtained through the use of this prior exhibit Cauchy-like tail behavior just like the LIML estimator.

Kleibergen and van Dijk (1998) considered the problem of local non-identification leading to pathological behaviour of the posteriors when the diffuse prior is used in Bayesian analysis of SEM. Their suggestion was the specification of the reduced form of SEM as a multivariate linear model with nonlinear (reduced rank) restrictions on its parameters. The results in Gao and Lahiri (2001) showed that in models involving weak instruments, no estimator is superior to others in all cases. In the case of weak endogeneity, they showed that Zellner's MELO performs best. However, when endogeneity is not weak and $\rho w_{12} > 0$, where ρ is the correlation coefficient between the structural and reduced form errors, and w_{12} is the co-variance between the unrestricted reduced form errors, BMOM (Bayesian method of moments) performed far more than other estimators.

Kleibergen and Zivot (2003) focused on determining the form of priors that lead to posteriors for structural parameters having similar properties as classical 2SLS and LIML thus providing more insight into the small sample behaviour of the Bayesian and classical procedures. A study on the performance of the Bayesian and classical approaches in estimation of a just-identified model was presented in Okewole *et al.* (2011). The result of the Monte Carlo experiment was that the Bayesian approach performed better than the classical approaches mostly in small samples while the performance became similar in large samples. This literature search then led us to a gap to be filled; an evaluation of the performance of the Bayesian approach in the presence of multicollinearity.

2.0 Methodology

A two-equation model was considered as follows;

$$
\begin{aligned}
y_{1t} &= \gamma y_{2t} + \beta_{11} X_{1t} + u_{1t} \\
y_{2t} &= \beta_{21} X_{1t} + \beta_{22} X_{2t} + \beta_{23} X_{3t} + u_{2t}
\end{aligned}
\qquad (1)
$$

y_{1t} and y_{2t} are each $(T\text{x}1)$ vectors containing observations on the endogenous variables

X_{1t}, X_{2t}, X_{3t} are each (Tx1) vectors of observations on the exogenous variables

γ is the scalar coefficient of the endogenous explanatory variable

$\beta_{11}, \beta_{21}, \beta_{22}, \beta_{23}$ are the scalar coefficients of the predetermined explanatory variables.

u_{1t} and u_{2t} are each (Tx1) random disturbance terms.

Bayesian analysis was carried out by working directly on the structural model (1) which we write in matrix form as

$$Y = X\delta + U \tag{2}$$

Where $Y = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}$, $X = \begin{pmatrix} y_{2t} & X_{1t} & 0 & 0 \\ 0 & X_{1t} & X_{2t} & X_{3t} \end{pmatrix}$, $\delta = \begin{pmatrix} \gamma & 0 \\ \beta_{11} & \beta_{21} \\ 0 & \beta_{22} \\ 0 & \beta_{23} \end{pmatrix}$ and $U = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$

The following assumptions were made about the model (2)

(1) The assumption about the exogenous regressors or predetermined variable: it is assumed that the matrix of the exogenous regressors is of full rank, ,i.e., the number of independent columns of the matrix X, which in the case of model (2) is three.
(2). The second assumption is the major one which is about the residual term U, the summary of this assumption is;

$$U \sim NIID(0, \Sigma) \tag{3}$$

2.1 **Prior Probability density function**

The flat or locally-uniform prior is assumed for the parameters of our models. The idea was to make inferences that are not greatly affected by external information or when external information is not available. Two rules were suggested by Jeffrey (1961) to serve as guide in choosing a prior distribution. The first one states that "If the parameter may have any value in a finite range, or from $-\infty$ to $+\infty$, its prior probability should be taken as uniformly distributed". While the second is that if the parameter, by nature, can take any value from 0 to ∞, the prior probability of its logarithm should be taken as uniformly distributed.

For the purpose of this study, we assume that little is known, a priori, about the elements of the parameter δ, and the three distinct elements of Σ. As the prior probability density function (pdf), we assume that the elements of δ and those of Σ are independently distributed; that is,

$$P(\delta, \Sigma) = P(\delta)P(\Sigma) \tag{4}$$

$$P(\delta) = \text{constant} \tag{5}$$

$$P(\Sigma) \propto |\Sigma|^{-3/2} \tag{6}$$

By denoting $\sigma^{\mu\mu'}$ as the $(\mu,\mu')th$ element of the inverse of Σ, and the Jacobian of transformation of the three variances, $(\sigma_{11},\sigma_{12},\sigma_{22})$ to $(\sigma^{11},\sigma^{12},\sigma^{22})$ as

$$J = \left|\frac{\partial(\sigma_{11},\sigma_{12},\sigma_{22})}{\partial(\sigma^{11},\sigma^{12},\sigma^{22})}\right| = |\Sigma|^{3} \tag{7}$$

The prior pdf in (6) implies the following prior pdf on the three distinct elements of Σ^{-1}

$$P(\Sigma^{-1}) \propto |\Sigma^{-1}|^{-\frac{3}{2}} \tag{8}$$

This could also be seen as taking an informative prior pdf on Σ^{-1} in the Wishart pdf form and allowing the "degrees of freedom" in the prior pdf to be zero. With zero degrees of freedom, there is a "spread out" Wishart pdf which then serve as a diffuse prior pdf since it is diffuse enough to be substantially modified by a small number of observations. The Wishart distribution is the conjugate for the multivariate normal distribution, which is the distribution of the variance-covariance matrix (Σ)

Hence, our prior p.d.f 's are (5), (6), and (8)

Zellner (1971), Geisser (1965) and others obtained these prior pdf's or parameters of similar models as in (2).

2.2 Likelihood Function

Based on the assumption from our model that rows of U are normally and independently distributed, each with zero mean vector and 2x2 covariance matrix Σ, the likelihood function for δ and Σ is;

$$L(\delta,\Sigma / Y,X) \propto |\Sigma|^{-n/2} \exp[-\tfrac{1}{2} tr(Y-X\delta)'(Y-X\delta)\Sigma^{-1}] \tag{9}$$

This is the same as;

$$L(\delta,\Sigma / Y,X) \propto |\Sigma|^{-n/2} \exp[-\tfrac{1}{2} trS\Sigma^{-1} - \tfrac{1}{2} tr(\delta-\hat{\delta})'X'X(\delta-\hat{\delta})\Sigma^{-1}] \tag{10}$$

Where $(Y-X\delta)'(Y-X\delta) = (Y-X\hat{\delta})'(Y-X\hat{\delta}) + (\delta-\hat{\delta})'X'X(\delta-\hat{\delta})$,

$$= S + (\delta-\hat{\delta})'X'X(\delta-\hat{\delta})$$

$S = (Y-X\hat{\delta})'(Y-X\hat{\delta})$ and $\hat{\delta}$ is the estimate of δ

Thus, the likelihood function for the parameters is as given in (10)

2.3 The Posterior Pdf

Combining the Prior pdf (5) and (8) with the likelihood function (10), we have the joint posterior distribution for δ and Σ^{-1} given as;

$$P(\delta,\Sigma^{-1}/Y,X) \propto |\Sigma|^{-(n+3/2)} EXP\{-\tfrac{1}{2} tr\left[S+(\delta-\hat{\delta})' Z'Z(\delta-\hat{\delta})\right]\Sigma^{-1}\} \quad (11)$$

Integrating (11) with respect to Σ^{-1}, we have the marginal posterior pdf for δ given as:

$$P(\delta/Y,X) \propto \left[S+(\delta-\hat{\delta})'Z'Z(\delta-\hat{\delta})\right]^{-T/2} \quad (12)$$

Which is a pdf in the generalized student-t form.

3.0 The Design and methodology of the experiment

3.1 Generating data for the experiment

Monte Carlo simulation approach was used in this research work. The data was generated by arbitrarily fixing values for the parameters of the model and stating specific distributions for the predetermined variables and the error terms. These values are put together in three runs to depict some possible research scenario involving multicollinearity. They are stated as follows.

$\gamma = 3.0, \beta_{11} = 1.0, \beta_{21} = 2.0, \beta_{22} = 0.5, \beta_{23} = 1.5$

$$X_{1t} : NID(0,1), X_{2t} : NID(0,1), X_{3t} : NID(0,1), (u_{1t}u_{2t}) : NID(0,0;\sigma_{11},\sigma_{12},\sigma_{22}),$$
$$\sigma_{11} = 1.0, \sigma_{12} = 1.0, \sigma_{22} = 4.0$$

RUN ONE: Depicting a high level of multicollinearity of two exogenous terms in the second equation. $\rho(X_{2t}X_{3t}) = 0.8$,

RUN TWO: Depicting a low level of multicollinearity of two exogenous terms in the second equation $\rho(X_{2t}X_{3t}) = 0.2$

RUN THREE: Depicting a very high level of multicollinearity of two exogenous terms in the second equation $\rho(X_{2t}X_{3t}) = 0.9$

In each of these runs, $N = 5000$ samples of size $T = 20, 40, 60,$ and 100 were generated, that is, the number of replicates is 5000, making a total of 20,000 samples in one run, and 60,000 samples altogether. $N =$ number of replicates $T =$ and sample size. Simulation of the endogenous variables as well as generating values for the exogenous variables was done in STATA

3.2 Analysis of the data

The analysis of the classical estimators was carried out using Mathlab while WinBUGS was used for the Bayesian approach. WinBUGS is windows version of the software referred to as Bayesian Analysis using Gibbs sampling. A sub-program is usually written by the researcher specifying the model and the prior distributions, and then WinBUGS uses Markov Chain Monte Carlo simulation to draw samples from the posterior distribution. Here, we first carried out 1000 iterations after which we observed sign of convergence, then a further 5000 iterations were carried out and the first 1000 taken as 'burn in'. History plots for a few samples are displayed below in fig 3.1 to show the convergence.

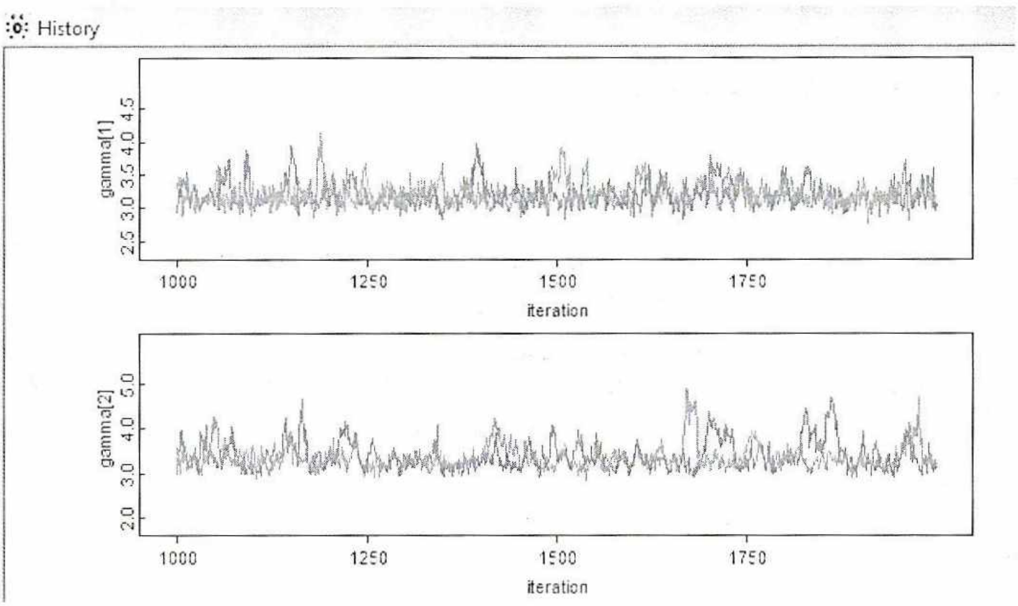

Fig. 3.1: History plots to check convergence

3.3 Criteria for assessing the performance of the Estimators

There are a number of comparison criteria used in literature, however, the mean, bias, and Mean Squared Error (MSE) were used

There are 5000 replicates so the mean of the estimates= $\frac{1}{5000}\sum_{i=1}^{N=5000}\hat{\theta}_i$

Estimated bias $=\frac{1}{5000}\sum_{i=1}^{N=5000}\hat{\theta}_i-\theta$

The mean squared error, for an estimator of a parameter θ, is given as;

$MSE(\hat{\theta})=E(\theta-\hat{\theta})^2$

$= \mathrm{Var}(\hat{\theta}) + (\text{Estimated bias})^2$

where $Var(\hat{\theta}) = \frac{1}{N_r}\sum_{1}^{N_r}\hat{\theta}^2 - (\frac{1}{N_r}\sum_{1}^{N_r}\hat{\theta})^2$ and N_r is number of replications and therefore number of estimates ($\hat{\theta}$). θ is the true value of the parameter, which in this case is the value used to generate the sample values.

It is important to note that in Bayesian estimation using Monte Carlo integration method, the convergence process of the estimates to the expected value, takes time. As a result of this, posterior estimates obtained early in the convergence process might make the estimator appear to be biased. Also, the effect of multicollinearity is increased in variance of estimates as mentioned in the introduction. Hence, we focus more on MSE of the estimates which also has a component that represents variance of the estimates; this enables appropriate comparison of the Bayesian estimates with those from the classical approaches. The relative efficiency of the estimators were also computed from the MSE and used for easier comparison

4.0 Discussion of Results

The results are presented in Tables 4.1a to 4.12. **Note**: The value in the parenthesis gives the relative efficiency of the Bayesian approach over the classical approach
In run one where multicollinearity is high between two exogenous variables of the second equation, the MSE from the Bayesian approach was smaller than those from OLS and LIML for γ and β_{11} (the regression coefficients in the first equation) but higher than 2SLS as shown in tables 4.1a to 4.4. OLS is known to be very sensitive to over-identification constraint which is the reason for the poor performance; it was also biased as expected.

The Bayesian approach could also get better than the 2SLS after further iterations, an experiment with fewer replications (10) and larger number of iterations (100,000) showed the Bayesian to be better than all the classical estimators both in terms of MSE and bias (table 4.1(b). We could not carry out more than 6000 iterations in the WinBUGS because the size of the data increased the volume of data thereby making the updating process in WinBUGS very slow.

Results from run one further showed that as the sample size (T) increases, the difference between the Bayesian approach and the classical approaches for all parameters of the model reduces as reflected in the reduction in relative efficiency. Also, increase in sample size brought LIML closer to the Bayesian approach for γ and β_{11}, estimates from the two approaches were similar in the larger sample sizes $T = 60$ and 100 than the lower sample sizes. This is consistent with result in literature; see Chao and Phillips (1998).

It is important to note that β_{21}, β_{22}, and β_{23} (regression coefficients in the second equation of the model) were the same for OLS, 2SLS/3SLS and LIML because there was no endogenous variable among the explanatory variables of that equation. There was a marked difference between the MSE of β_{22} and β_{23} (the coefficient of the variables with multicollinearity issue) from the Bayesian method and those from the classical method: the Bayesian method had smaller MSE for all the sample sizes considered. The relative efficiency as presented in parenthesis (close to the MSE of β_{22} and β_{23}) clearly reflected this difference. This result suggests that the Bayesian approach is less affected by the presence of multicollinearity than the classical approaches, particularly the least squares. Further results to support this reflected in the comparison of results from run one and two.

Similar results were obtained in run 2 as shown in tables 4.5 to 4.8 where the multicollinearity is low. However, as expected, the MSE reduced in this run as compared with the first run. The size of this change (SC) in MSE from run 1 to run 2 is greater in the classical method than the Bayesian method. For instance, for β_{22} and β_{23}, SC of MSE of the classical approach from run 1 to 2 are 0.0591 and 0.1171 for T=20, 0.0746 and 0.1294 for T=40, 0.0981 and 0.0615 for T=60, 0.0357 and 0.0694 for T=100 whereas it was 0.0322 and 0.0141 for T=20, 0.0486 and 0.0814 for T=40, 0.0673 and 0.0212 for T=60 and 0.0235 and 0.0352 for T=100 in the Bayesian approach. The differences were bigger in the classical approach than the Bayesian suggesting that the Bayesian approach is less affected by multicollinearity than the least squares of the classical approaches.

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	3.0906	3.0000	2.9903	2.9867
	ABS Bias	0.0906	0.0000	0.0097	0.0133
	MSE	0.0127	0.0089	0.0103	0.0103
$\beta_{11}(1.0)$	Mean	0.8315	1.0000	1.0181	1.0110
	ABS Bias	0.1685	0.0000	0.0181	0.0110
	MSE	0.0819	0.0746	0.0803	0.0768
$\beta_{21}(2.0)$	Mean	2.0062	2.0062	2.0062	1.9630
	ABS Bias	0.0062	0.0062	0.0062	0.0370
	MSE	0.1845	0.1845	0.1845	0.1762
$\beta_{22}(0.5)$	Mean	0.5118	0.5118	0.5118	0.4992
	ABS Bias	0.0118	0.0118	0.0118	0.0008
	MSE	0.4230	0.4230	0.4230	0.2825(1.4973)
$\beta_{23}(1.5)$	Mean	1.4989	1.4989	1.4989	1.4411
	ABS Bias	0.0011	0.0011	0.0011	0.0589
	MSE	0.3502	0.3502	0.3502	0.2502(1.3997)

Table 4.1a: Results from run 1, Sample Size T= 20

In run three involving a higher multicollinearity than run one, the MSE reflected this high level of multicollinearity; they were much higher than other runs. The relative efficiency of the Bayesian over the classical was much higher than the first two runs. For instance, as presented in table 4.9 for sample size 20, the relative efficiency of the Bayesian over the classical estimator for β_{22} and β_{23} was 1.9905 and 1.9135 respectively whereas it was 1.4973 and 0.3997 in run one and 1.4539 and 0.9873 in run two. This result supported earlier results from runs one and two that Bayesian approach is less sensitive to multicollinearity. Furthermore, it also shows that the higher the degree of multicollinearity, the wider the gap between the Bayesian approach and the classical approach

Results from tables 4.10 to 4.12 for sample sizes 40, 60 and 100 also follows the same pattern as in runs one and two, that is, as the sample size increased, the gap between the Bayesian approach and the classical approach reduced as reflected in the relative efficiency. The relative efficiency of the Bayesian over the classical were 1.7073 and 1.5451 for $T = 40$, 1.5168 and 1.2731 for $T = 60$ and 1.4072 and 1.3077 for $T = 100$.

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	2.8756	3.0988	3.0634	3.0949
	ABS Bias	0.1244	0.0988	0.0634	0.0949
	MSE	0.0241	0.0638	0.0558	0.0529
$\beta_{11}(1.0)$	Mean	1.0857	0.8395	0.8734	0.854
	ABS Bias	0.0857	0.1605	0.1266	0.1461
	MSE	0.0579	0.1174	0.124	0.0976
$\beta_{21}(2.0)$	Mean	2.1383	2.1383	2.1383	2.035
	ABS Bias	0.1383	0.1383	0.1383	0.035
	MSE	0.3002	0.3002	0.3002	0.2419
$\beta_{22}(0.5)$	Mean	0.2739	0.2739	0.2739	0.3223
	ABS Bias	0.2261	0.2261	0.2261	0.1777
	MSE	0.6278	0.6278	0.6278	0.2463(2.5489)
$\beta_{23}(1.5)$	Mean	1.6063	1.6063	1.6063	1.4091
	ABS Bias	0.1063	0.1063	0.1063	0.0909
	MSE	0.7427	0.7427	0.7427	0.3301(2.2499)

Table 4.1b:Results from 10 replicates with 100,000 WinBUGS iterations sample size T= 20

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	3.1366	3.0000	2.9905	2.9891
	ABS Bias	0.1366	0.0000	0.0095	0.0109
	MSE	0.0216	0.0087	0.0099	0.0096
$\beta_{11}(1.0)$	Mean	0.6773	1.0011	1.0236	1.0206
	ABS Bias	0.3227	0.0011	0.0236	0.0206
	MSE	0.1357	0.0676	0.0751	0.0712
$\beta_{21}(2.0)$	Mean	2.0018	2.0018	2.0018	1.9955
	ABS Bias	0.0018	0.0018	0.0018	0.0045
	MSE	0.0840	0.0840	0.0840	0.0823
$\beta_{22}(0.5)$	Mean	0.5060	0.5060	0.5060	0.5019
	ABS Bias	0.0060	0.0060	0.0060	0.0019
	MSE	0.1583	0.1583	0.1583	0.1128(1.4034)
$\beta_{23}(1.5)$	Mean	1.4953	1.4953	1.4953	1.4383
	ABS Bias	0.0047	0.0047	0.0047	0.0617
	MSE	0.2284	0.2284	0.2284	0.1799(1.2696)

Table 4.2: Results from run 1, sample size T= 40

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	3.1244	2.9984	2.9937	2.9927
	ABS Bias	0.1244	0.0016	0.0063	0.0073
	MSE	0.0173	0.0045	0.0048	0.0048
$\beta_{11}(1.0)$	Mean	0.7033	1.0046	1.0157	1.0145
	ABS Bias	0.2967	0.0046	0.0157	0.0145
	MSE	0.1078	0.0379	0.0397	0.0391
$\beta_{21}(2.0)$	Mean	2.0019	2.0019	2.0019	1.9973
	ABS Bias	0.0019	0.0019	0.0019	0.0027
	MSE	0.0480	0.0480	0.0480	0.0474
$\beta_{22}(0.5)$	Mean	0.5050	0.5050	0.5050	0.5074
	ABS Bias	0.0050	0.0050	0.0050	0.0074
	MSE	0.1498	0.1498	0.1498	0.1079(1.3883)
$\beta_{23}(1.5)$	Mean	1.4939	1.4939	1.4939	1.4573
	ABS Bias	0.0061	0.0061	0.0061	0.0427
	MSE	0.1614	0.1614	0.1614	0.1280(1.2609)

Table 4.3: Results from run 1 sample size T= 60

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	3.1168	2.9995	2.9972	2.9963
	ABS Bias	0.1168	0.0005	0.0028	0.0037
	MSE	0.0146	0.0023	0.0023	0.0024
$\beta_{11}(1.0)$	Mean	0.7765	1.0013	1.0058	1.0049
	ABS Bias	0.2235	0.003	0.0058	0.0049
	MSE	0.0607	0.0171	0.0174	0.0173
$\beta_{21}(2.0)$	Mean	1.9987	1.9987	1.9987	1.9895
	ABS Bias	0.0013	0.0013	0.0013	0.0105
	MSE	0.0354	0.0354	0.0354	0.0352
$\beta_{22}(0.5)$	Mean	0.4920	0.4920	0.4920	0.4967
	ABS Bias	0.0080	0.0080	0.0080	0.0033
	MSE	0.0897	0.0897	0.0897	0.0649(1.3821)
$\beta_{23}(1.5)$	Mean	1.5054	1.5054	1.5054	1.4820
	ABS Bias	0.0054	0.0054	0.0054	0.0180
	MSE	0.1019	0.1019	0.1019	0.0797(1.2785)

Table 4.4: Results from run 1 sample size T= 100

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	3.1552	3.0013	2.9706	2.9781
	ABS Bias	0.1552	0.0013	0.0294	0.0219
	MSE	0.0315	0.0292	0.2249	0.0293
$\beta_{11}(1.0)$	Mean	0.6959	0.9967	1.0499	1.0122
	ABS Bias	0.3041	0.0033	0.0499	0.0122
	MSE	0.1756	0.1870	1.0188	0.1623
$\beta_{21}(2.0)$	Mean	1.9959	1.9959	1.9959	1.9264
	ABS Bias	0.0041	0.0041	0.0041	0.0736
	MSE	0.2581	0.2581	0.2581	0.2497
$\beta_{22}(0.5)$	Mean	0.5033	0.5033	0.5033	0.4570
	ABS Bias	0.0033	0.0033	0.0033	0.0430
	MSE	0.3639	0.3639	0.3639	0.2503(1.4539)
$\beta_{23}(1.5)$	Mean	1.5099	1.5099	1.5099	1.3720
	ABS Bias	0.0099	0.0099	0.0099	0.1280
	MSE	0.2331	0.2331	0.2331	0.2361(0.9873)

Table 4.5: Results from run 2 sample size T= 20

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	2.8656	2.9983	3.0071	3.0034
	ABS Bias	0.1344	0.0016	0.0071	0.0034
	MSE	0.0210	0.0080	0.0089	0.0091
$\beta_{11}(1.0)$	Mean	1.2778	1.0023	0.9840	0.9991
	ABS Bias	0.2778	0.0023	0.0160	0.0009
	MSE	0.1139	0.0650	0.0694	0.1240
$\beta_{21}(2.0)$	Mean	2.0011	2.0011	2.0011	1.9807
	ABS Bias	0.0011	0.0011	0.0011	0.0193
	MSE	0.1202	0.1202	0.1202	0.1175
$\beta_{22}(0.5)$	Mean	0.5016	0.5016	0.5016	0.4868
	ABS Bias	0.0016	0.0016	0.0016	0.0132
	MSE	0.0837	0.0837	0.0837	0.0642(1.3037)
$\beta_{23}(1.5)$	Mean	1.5054	1.5054	1.5054	1.4541
	ABS Bias	0.0054	0.0054	0.0054	0.0459
	MSE	0.0990	0.0990	0.0990	0.0985(1.0051)

Table 4.6: Results from run 2 sample size T= 40

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	2.8365	3.0014	3.0111	3.0017
	ABS Bias	0.16349	0.0014	0.0111	0.0017
	MSE	0.0291	0.0094	0.0106	0.0163
$\beta_{11}(1.0)$	Mean	1.3348	0.9964	0.9764	1.0198
	ABS Bias	0.3348	0.0036	0.0236	0.0198
	MSE	0.1373	0.0594	0.0654	0.1306
$\beta_{21}(2.0)$	Mean	2.0028	2.0028	2.0028	1.9899
	ABS Bias	0.0028	0.0028	0.0028	0.0101
	MSE	0.0750	0.0750	0.0750	0.0740
$\beta_{22}(0.5)$	Mean	0.5083	0.5083	0.5083	0.4877
	ABS Bias	0.0083	0.0083	0.0083	0.0123
	MSE	0.0517	0.0517	0.0517	0.0406(1.2734)
$\beta_{23}(1.5)$	Mean	1.5013	1.5013	1.5013	1.4437
	ABS Bias	0.0013	0.0013	0.0013	0.0563
	MSE	0.0999	0.0999	0.0999	0.1068(0.9354)

Table 4.7: Results from run 2 sample size T= 60

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	2.8628	3.0012	3.0044	3.0012
	ABS Bias	0.1372	0.0012	0.0044	0.0012
	MSE	0.0110	0.0034	0.0035	0.0036
$\beta_{11}(1.0)$	Mean	1.2648	0.9960	0.9897	1.0405
	ABS Bias	0.2648	0.0040	0.0103	0.0405
	MSE	0.0816	0.0222	0.0228	0.1644
$\beta_{21}(2.0)$	Mean	2.0010	2.0010	2.0010	1.9925
	ABS Bias	0.0010	0.0010	0.0010	0.0075
	MSE	0.0382	0.0382	0.0382	0.0380
$\beta_{22}(0.5)$	Mean	0.4977	0.4977	0.4977	0.4874
	ABS Bias	0.0023	0.0023	0.0023	0.0126
	MSE	0.0540	0.0540	0.0540	0.0414(1.3043)
$\beta_{23}(1.5)$	Mean	1.4977	1.4977	1.4977	1.4677
	ABS Bias	0.0023	0.0023	0.0023	0.0323
	MSE	0.0325	0.0325	0.0325	0.0445(0.7303)

Table 4.8: Results from run 2 sample size T= 100

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	2.9161	3.0015	3.0096	3.0080
	ABS Bias	0.0839	0.0015	0.0096	0.0080
	MSE	0.0113	0.0079	0.0088	0.0114
$\beta_{11}(1.0)$	Mean	1.2145	0.9969	0.9762	0.9941
	ABS Bias	0.2145	0.0031	0.0238	0.0059
	MSE	0.1213	0.1084	0.1150	0.1373
$\beta_{21}(2.0)$	Mean	2.0032	2.0032	2.0032	1.9798
	ABS Bias	0.0032	0.0032	0.0032	0.0202
	MSE	0.2188	0.2188	0.2188	0.2087
$\beta_{22}(0.5)$	Mean	0.4957	0.4957	0.4957	0.5591
	ABS Bias	0.0043	0.0043	0.0043	0.0591
	MSE	1.5317	1.5317	1.5317	0.7695(1.9905)
$\beta_{23}(1.5)$	Mean	1.5081	1.5081	1.5081	1.3718
	ABS Bias	0.0081	0.0081	0.0081	0.1282
	MSE	1.3442	1.3442	1.3442	0.7025(1.9135)

Table 4.9: Results from run 3 sample size T= 20

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	2.8767	2.9997	3.0068	3.0028
	ABS Bias	0.1233	0.0003	0.0068	0.0028
	MSE	0.0180	0.0067	0.0073	0.0084
$\beta_{11}(1.0)$	Mean	1.2618	1.0044	0.9897	1.0251
	ABS Bias	0.2618	0.0044	0.0103	0.0251
	MSE	0.0967	0.0487	0.0514	0.1147
$\beta_{21}(2.0)$	Mean	1.9961	1.9961	1.9961	1.9856
	ABS Bias	0.0039	0.0039	0.0039	0.0144
	MSE	0.0764	0.0764	0.0764	0.0753
$\beta_{22}(0.5)$	Mean	0.4999	0.4999	0.4999	0.5239
	ABS Bias	0.0001	0.0001	0.0001	0.0239
	MSE	0.7150	0.7150	0.7150	0.4188(1.7073)
$\beta_{23}(1.5)$	Mean	1.4967	1.4967	1.4967	1.3822
	ABS Bias	0.0033	0.0033	0.0033	0.1178
	MSE	0.7928	0.7928	0.7928	0.5131(1.5451)

Table 4.10: Results from run 3 sample size T= 40

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	2.8726	2.9994	3.0041	2.9997
	ABS Bias	0.1274	0.0006	0.0041	0.0003
	MSE	0.0180	0.0047	0.0050	0.0049
$\beta_{11}(1.0)$	Mean	1.2753	1.0006	0.9904	1.0462
	ABS Bias	0.2753	0.0006	0.0096	0.0462
	MSE	0.0956	0.0359	0.0372	0.1722
$\beta_{21}(2.0)$	Mean	2.0037	2.0037	2.0037	1.9977
	ABS Bias	0.0037	0.0037	0.0037	0.0023
	MSE	0.0560	0.0560	0.0560	0.0551
$\beta_{22}(0.5)$	Mean	0.4881	0.4881	0.4881	0.4979
	ABS Bias	0.0119	0.0119	0.0119	0.0021
	MSE	0.3936	0.3936	0.3936	0.2595(1.5168)
$\beta_{23}(1.5)$	Mean	1.5059	1.5059	1.5059	1.4259
	ABS Bias	0.0059	0.0059	0.0059	0.0741
	MSE	0.3422	0.3422	0.3422	0.2688(1.2731)

Table 4.11: Results from run 3 sample size T= 60

		OLS	2SLS/3SLS	LIML	BAYESIAN
$\gamma(3.0)$	Mean	2.8741	2.9993	3.0020	3.0004
	ABS Bias	0.1259	0.0007	0.0020	0.0004
	MSE	0.0169	0.0028	0.0029	0.0029
$\beta_{11}(1.0)$	Mean	1.2759	1.0009	0.9950	1.0275
	ABS Bias	0.2759	0.0009	0.0050	0.0275
	MSE	0.0882	0.0226	0.0232	0.1550
$\beta_{21}(2.0)$	Mean	2.0062	2.0062	2.0062	2.0010
	ABS Bias	0.0062	0.0062	0.0062	0.0010
	MSE	0.0369	0.0369	0.0369	0.0366
$\beta_{22}(0.5)$	Mean	0.4872	0.4872	0.4872	0.4962
	ABS Bias	0.0128	0.0128	0.0128	0.0038
	MSE	0.2160	0.2160	0.2160	0.1535(1.4072)
$\beta_{23}(1.5)$	Mean	1.5136	1.5136	1.5136	1.4766
	ABS Bias	0.0136	0.0136	0.0136	0.0234
	MSE	0.1976	0.1976	0.1976	0.1511(1.3077)

Table 4.12: Results from run 3 sample size T= 100

5.0 Conclusion

The importance of studies on multicollinearity and other violations of basic assumptions of regression models cannot be over emphasized. Bayesian approach to the estimation of multi-equation econometric models has been shown in this study to be less sensitive to the presence of non-orthogonal regressors than the classical estimators. Although there are various remedies for multicollinearity, in some cases, the data are often analyzed without checking for consistency with basic assumptions. It is therefore important to be aware of estimation approach that is not seriously affected by violations of these assumptions. With the availability of numerical intensive software, recently, the Bayesian approach has become more widely used such that there is a vast literature on its theories and application.

References

Adejumobi , A. A. (2005). "Robustness of Simultaneous estimation techniques to over-identified and correlated random deviates". Unpublished Ph.D Thesis

Agunbiade, D. A. (2008). "A Monte Carlo approach to the estimates of a just identified simultaneous three-equation model with three multicollinear exogenous variates" Unpublished Ph.D Thesis

Blomquist, S. and M. Dahlberg .(1999). "Small Sample Properties of LIML and Jacknife IV Estimators: Experiments with weak Instruments". *Journal of Applied Econometrics.* Vol. 14. no. 1 pg 69 – 88. JSTOR

Chao J.C. and Phillips P.C.B. (1998). "Posterior distributions in limited information analysis of the simultaneous equations model using the Jeffreys prior. *Journal of Econometrics.* Vol. 87 pg. 49-86. JSTOR

Gao C. and Lahiri K. (2001). "A comparison of some recent Bayesian and Classical procedures for simultaneous equations models with weak instrument" Discussion paper, SUNY at Albany.

Geisser, S. (1965). "A Bayes Approach for Combining Correlated Estimates". *Journal of the American Stat. Assoc.* Vol. 60 no. 310 pg. 602-607

Jeffreys, H. (1961). "Theory of Probability". 3rd edition, London: Oxford University Press.

Kleibergen, F. and van Dijk, H.K.(1998). "Bayesian Simultaneous equation analysis using reduced rank structures". *Econometric Theory* Vol.14 pg. 701-743.

Kleibergen F., Zivot Eric. (2003). "Bayesian and classical approaches to instrumental variable regression". *Journal of Econometrics.* Vol. 114 pg. 29-72. JSTOR

Oduntan, E. A. (2004). "A Monte Carlo study of the problem of Multicollinearity in a Simultaneous equations model". Unpublished PhD Thesis, University of Ibadan

Olubusoye, O.E. (2001). "Relative sensitivity of Estimators of simultaneous equations to violations of mutual independence of random variates". Unpublished PhD Thesis, University of Ibadan.

Okewole, D.M., Olubusoye O.E. and Shangodoyin D.K. (2011)." A Comparative Study of the Classical and Bayesian Methods of Estimating a Just-identified Simultaneous Equations Econometric Mode"l. Asian Social Science. Vol. 7. no. 5 pg. 222 – 235

Zellner, A. (1965). "Bayesian inference and Simultaneous Equation Econometric Models". paper presented to the first world Conference of the Econometric Society, Rome, Italy. Abstract published in Econometrica Vol. 34 no. 5 pg. 47. JSTOR

Zellner, A. (1971). "An Introduction to Bayesian Inference in Econometrics" John Wiley & sons, Inc.