

A Comparative Study of the Classical and Bayesian Methods of Simultaneous Equations Econometric Model Estimation

By

Dorcas Modupe Okewole

Department of Mathematical Sciences

Redeemer's University

Ogun State, Nigeria

Email: dupedepo@yahoo.com

Mobile: 08023312820

And

Olusanya E. Olubusoye

Department of Statistics

University of Ibadan,

Oyo State, Nigeria.

Email: busoye2001@yahoo.com

Mobile: 08058258883

A two-equation econometric model is simulated using both Classical and Bayesian procedures. The estimates of the parameters for both methods were compared under a wide range of scenarios; sample size, residual variance and variance of the data on the predetermined variable. The Monte Carlo experiment was performed using E-views and WinBUGS computer software. The median, being a robust estimator of average in terms of validity, was used as the posterior estimate. As indicated in similar research in the past where the posterior mode was used as estimate, the Bayesian procedure performed better in most cases, while some scenarios showed similar behavior for the two procedures.

JEL classification: C₁₁, C₁₃, C₁₅, C₃₀

Keywords: Bayesian procedures, Classical procedures, Simultaneous equations, Monte-Carlo, Simulation, Estimates.

[1.0] Introduction

Simultaneous equations model (SEM) is a very important field of Econometrics. Some important Statistical implications of a linear simultaneous equation model were presented by Haavelmo (1943), 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. 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, Imbens and Krueger (1999) and Blomquist and Dahlberg (1999) method 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. The 'true' model structure is assumed unknown, and is being estimated. However, Dreze (1962) argues 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 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. A comparative study of the classical and the Bayesian approaches is thus necessary so as to take advantage of their strength and research more on possible ways of improving on their weaknesses.

[2.0] Literature Review

The need to carry out valid, generally acceptable, appropriate and convenient estimation of the Simultaneous equations econometric models has brought about quite a number of researches on the classical and the Bayesian procedures. Arnold Zellner(1971) presented the comparison of these two procedures on model 1.2, the result showed that the Bayesian method performed better than the classical mostly for the small sample case. A research was also carried out by Gao and Lahiri(2001), focusing on weak instruments. 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 also 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 reflected some Bayesian estimation method of Simultaneous equations econometric models such as; Bayesian method of moment proposed by Zellner, the methods used by Chao and Phillips (1998), Geweke (1996) and Kleibergen and Van Dijk (1998).

Here in this research work, we compare the properties of the Bayesian estimators and the classical estimators in repeated trials as carried out by Arnold Zellner(1971) but making use of the median as Bayesian estimate and using additional comparison criteria; mean of the estimates, the bias, and Mean Squared Error.

Generally, Econometric models are often expressed in terms of an unknown vector of parameters $\underline{\theta} \in \Theta \subseteq \mathbb{R}^k$ which fully specifies the joint

probability distribution of the observations $\mathbf{X} = (x_1 \dots x_T)$. Given the probability density function $f(\mathbf{X}/\underline{\theta})$, the classical estimation often proceeds by making use of the likelihood function $L(\underline{\theta}) = f(\mathbf{X}/\underline{\theta})$, while the Bayesian estimation technique combines the likelihood function with the prior information which is usually expressed as probability density function of the parameters, $\pi(\underline{\theta})$. This gives the posterior distribution, which is proportional to $p(\underline{\theta}) = \pi(\underline{\theta})L(\underline{\theta})$. Most Bayesian inference problems, according to Geweke (1989), can be seen as evaluation of the expectation of a function $u(\underline{\theta})$ of interest under the posterior,

$$E[u(\underline{\theta})] = \int_{\mathfrak{R}} u(\underline{\theta}) p(\underline{\theta}) d(\underline{\theta}) / \int_{\mathfrak{R}} p(\underline{\theta}) d(\underline{\theta}) \quad (1.1)$$

Methods of solving this problem (1.1) are not as systematic, methodical or general as are those to classical inference problems: because classical inference is carried out routinely using likelihood functions for which the evaluation of (1.1) is difficult and for most practical purposes impossible. There is a problem with the analytical integration of (1.1) in that the range of likelihood functions that can be considered is small, and the class of priors and functions of interest that can be considered is severely restricted. Also, many numerical approaches like quadrature methods, require special adaptation for each u , π , or L , and become unworkable if k exceeds, say, three.

However, with the advent of powerful and cheap computing, numerically intensive methods for solving (1.1) have become more interesting. This is where Monte Carlo integration comes in (among others) as a way out, particularly the Markov Chain Monte Carlo (MCMC), which involves Gibbs sampling. It provides a systematic approach that, in principle, can be applied in any situation in which $E[u(\underline{\theta})]$ exists, and is practical for large values of

k. In this regard, the works of Gao and Lahiri (2001) are of note, also Geweke (1989), Gilks, Richardson and Spiegelhalter (1996), and others. The analysis was carried out electronically with the use of Eviews as well as WinBUGS (Bayesian Analysis using Gibbs sampling), a Computer software developed by Biostatistics Unit at MRC(Medical Research council) Cambridge, and Imperial College School of medicine, London.

[3.0]: The model

The model analysed here is

$$\begin{aligned}
 Y_{1t} &= \gamma Y_{2t} + U_{1t} \\
 & \qquad \qquad \qquad t = 1, 2, \dots, T \qquad \qquad (1.2) \\
 Y_{2t} &= \beta X_t + U_{2t}
 \end{aligned}$$

Where Y_{1t} and Y_{2t} are observations on two endogenous variables, X_t is an observation on an exogenous variable, U_{1t} and U_{2t} are disturbance terms, and γ and β are scalar parameters. Zellner (1971) analyzed the model (1.2) using diffuse prior on the parameters and the posterior modal value as the Bayesian estimate. Here, we use the median of the posterior distribution as the Bayesian estimate and also make use of other comparison criteria; mean of the estimates and Mean Squared Error of the estimator.

The matrix form of our model (1.2) is

$$Y\Gamma = XB + U \dots\dots\dots (2.1)$$

Where $Y = (Y_1, Y_2)$, an $n \times 2$ matrix of observations on two endogenous variables. Γ is a unit matrix I_2 of coefficients for the endogenous variables.

X is nx1 vector of observations on the predetermined variables. B is 1x2 vector of coefficients for the predetermined variables, and U = (U₁ U₂) is an nx2 matrix of random disturbance terms. For easy analysis, we need the reduced form of (1.2) (and later carry out transformations as appropriate), given as;

$$Y_{1t} = \pi_1 X_{1t} + V_{1t} \dots\dots\dots (2.2)$$

$$Y_{2t} = \pi_2 X_{2t} + V_{2t}$$

Which in matrix form, is

$$Y = \pi X + V$$

Where $\pi = (\pi_1 \pi_2)$, the matrix of the reduced form parameter, $V = (V_1 V_2)$ an n x 2 matrix of reduced form disturbance term, $V_1 = \gamma U_{2t} + U_{1t}$, $V_2 = U_{2t}$, $\pi_1 = \beta\gamma$, $\pi_2 = \beta$.

[4.0] Research Method

The model was simulated using the two methods under study (classical and Bayesian). The first step was to generate the data. The following are the conditions under which the data was generated;

RUN I

$\gamma = 2.0$, $\beta = 0.5$, $X_t : NID(0,1)$, $(U_{1t} U_{2t}) : NID(0,0; \sigma_{11}, \sigma_{12} \sigma_{22})$, $\sigma_{11} = 1.0$, $\sigma_{12} = -1.0$, $\sigma_{22} = 4.0$

RUN II

$\gamma = 2.0$, $\beta = 0.5$, $X_t : NID(0,2)$, $(U_{1t} U_{2t}) : NID(0,0; \sigma_{11}, \sigma_{12} \sigma_{22})$, $\sigma_{11} = 1.0$, $\sigma_{12} = 1.0$, $\sigma_{22} = 4.0$

RUN III

$\gamma = 2.0, \beta = 0.5, X_t: \text{NID}(0,9), (U_{1t} U_{2t}): \text{NID}(0,0;\sigma_{11}, \sigma_{12} \sigma_{22}), \sigma_{11}=1.0, \sigma_{12}=1.0, \sigma_{22}= 4.0$

In each of these runs, 50 samples of size 20, 40, 60, and 100 were generated, making a total of 200 samples in one run, and 600 samples altogether. To obtain random disturbance terms that behave as stated in the three runs, we made use of the method presented in Nagar (1969)

Making use of the data generated the parameter estimates were obtained with the aid of E views for the classical method and WinBUGS for the Bayesian method. (The same data sets for the two methods).

Comparison criteria

The following were used in comparing the two estimation methods;

* Performance in repeated sampling, frequency distribution.

*Mean Estimate

*Bias

*Mean Squared Error (MSE), for a parameter θ , is given as;

$$\text{M.S.E}(\theta) = E(\theta - \hat{\theta})^2$$

Which is also the same as $= \text{Var}(\hat{\theta}) + (\text{Estimated bias})^2$

where $\text{Var}(\hat{\theta}) = \frac{1}{N_r} \sum_1^{N_r} \hat{\theta}^2 - (\frac{1}{N_r} \sum_1^{N_r} \hat{\theta})^2$ and N_r is number of replications and

therefore number of estimates ($\hat{\theta}$). Estimated bias $= \frac{1}{N_r} \sum_1^{N_r} \hat{\theta} - \theta$

θ is the true value of the parameter, which in this case is the value used to generate the sample values.

[5.0] Classical Estimation method

Routes to deriving point estimates through the classical method are;

- ❖ Maximum Likelihood (ML)
- ❖ Method of moments, generalized method of moments
- ❖ Minimum mean square error (MMSE)
- ❖ Minimum variance unbiased estimator (MVUE)
- ❖ Best linear unbiased estimator (BLUE)

The estimator of γ , by most of these principles of classical (sampling-theory) estimation, is given as;

$$\hat{y} = \frac{\sum_1^T y_{2t} y_{1t}}{\sum_1^T y_{2t}^2} = \frac{\hat{\pi}_1}{\hat{\pi}_2} \dots\dots\dots(5.1)$$

Where $\hat{y}_{2t} = \hat{\pi}_2 x_t, \hat{\pi}_2 = \frac{\sum_1^T x_t y_{2t}}{\sum_1^T x_t^2}$, and $\hat{\pi}_1 = \frac{\sum_1^T x_t y_{1t}}{\sum_1^T x_t^2}$

[6.0] Bayesian Estimation method

[6.1] Prior information

The Prior Probability density function could be informative or diffuse (non-informative).

(1) Informative prior . This is a situation where information is available about the Prior pdf. The informative prior applied here is;

$$P(\pi, \Sigma^{-1}) \propto |\Sigma^{-1}|^{-3/2} \exp[-1/2(\pi - \bar{\pi})'C^{-1}(\pi - \bar{\pi})] \dots\dots\dots(6.1.1)$$

Where $\bar{\pi}$, a 2 x 1 vector, is the mean of the prior pdf, Σ^{-1} is the inverse of the variance-covariance matrix and $C = (C_{ol})$ is a 2 x 2 matrix of the prior covariance.

(2) Diffuse prior

The idea behind the use of diffuse (a.k.a non-informative, vague) prior distributions is to make inferences that are not greatly affected by external information or when external information is not available.

Here, we assume little is known, a priori, about the parameters, π , and the three distinct elements of Σ . As our diffuse prior pdf, we assume that the elements of π and those of Σ are independently distributed; that is,

$$P(\pi, \Sigma) = p(\pi) p(\Sigma) \dots\dots\dots(6.1.2)$$

Using the Jeffrey's invariance theory, we take

$$P(\pi) = \text{constant} \dots\dots\dots(6.1.3)$$

And

$$P(\Sigma) \propto |\Sigma|^{-3/2} \dots\dots\dots(6.1.4)$$

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

$$J = \left| \frac{\partial(\sigma_{11}, \sigma_{12}, \sigma_{22})}{\partial(\sigma^{11}, \sigma^{12}, \sigma^{22})} \right| = |\Sigma|^3 \dots\dots\dots (6.1.5)$$

So that the prior pdf in (1.7) implies the following prior pdf on the three distinct elements of Σ^{-1}

$$P(\Sigma^{-1}) \propto |\Sigma^{-1}|^{-3/2} \dots\dots\dots(6.1.6)$$

This is as a result of 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.

Hence, our diffuse prior pdf ‘s are (6.1.3), (6.1.4), and (6.1.6)

These diffuse prior pdf ‘s were also arrived at by Zellner(1971), Savage, Geisser(1965) and others.

[6.2] Likelihood function

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

$$I(\pi, \Sigma / Y, X) \propto \exp[-1/2 \text{tr}(Y-X\pi)'(Y-X\pi)\Sigma^{-1}] \dots\dots\dots(6.2.1)$$

This is the same as;

$$I(\pi, \Sigma / Y, X) \propto |\Sigma|^{-N/2} \exp[-1/2 \text{tr} S \Sigma^{-1} - 1/2 \text{tr}(\pi - \hat{\pi})'X'X(\pi - \hat{\pi}) \Sigma^{-1}] \dots (6.2.2)$$

$$\begin{aligned} \text{Where } (Y-X\pi)'(Y-X\pi) &= (Y-X\hat{\pi})'(Y-X\hat{\pi}) + (\pi - \hat{\pi})'X'X(\pi - \hat{\pi}), \\ &= S + (\pi - \hat{\pi})'X'X(\pi - \hat{\pi}) \end{aligned}$$

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

Thus, the likelihood function for the parameters is as given in (6.2.2)

[6.3]The Posterior pdf

Combining the diffuse Prior pdf (6.1.3) and (6.1.6) with the likelihood function(6.2.2), we have a posterior distribution that is in the bivariate student-t form. This is given as:

$$p(\pi_1, \pi_2 | y_1, y_2) \propto \left[1 + (\pi_1 - \hat{\pi}_1; \pi_2 - \hat{\pi}_2) \begin{pmatrix} s^{11} & s^{12} \\ s^{21} & s^{22} \end{pmatrix} (\pi_1 - \hat{\pi}_1; \pi_2 - \hat{\pi}_2)' \right]^{-N/2} \dots (6.3.1)$$

Where $\hat{\pi}_1$ and $\hat{\pi}_2$ are the least squares estimates, $s^{ij} = \hat{w}^{ij} \sum_{n=1}^N x_n^2$, and \hat{w}^{ij} the

i,jth element of $\left[\sum_{n=1}^N (y_{in} - \hat{\pi}_i x_n)(y_{jn} - \hat{\pi}_j x_n) \right]$ i, j=1,2. To obtain the posterior

distribution in terms of γ and β , we carry out the following transformation:

$$\gamma = \frac{\pi_1}{\pi_2}, \quad \beta = \pi_2 \text{ with the Jacobian of transformation } |\beta|. \text{ This gives us:}$$

$$p(\gamma, \beta | y_1 y_2) \propto |\beta| \left[1 + (\beta\gamma - \hat{\pi}_1)^2 s^{11} + (\beta - \hat{\pi}_2)^2 s^{22} + 2(\beta\gamma - \hat{\pi}_1)(\beta - \hat{\pi}_2) s^{12} \right]^{-N/2} \dots (6.3.2)$$

The Baye's estimate is the mean of the Posterior distribution, if it can be identified, this is solving for the posterior distribution analytically. If the solution of the Posterior function cannot be obtained analytically, then numerical integration is employed in obtaining the normalizing constant which will be the antiderivative of the function, as the situation is in this research. The WinBUGS mentioned earlier is used to obtain the Bayesian estimate by drawing samples from the Posterior distribution and obtaining the mean after ensured there is convergence. Other measures of central tendencies like median and mode) could also be used as the Bayesian estimate (see Zellner, 1971). Hence, in this case, the median is used.

Routes to deriving point estimates via Bayesian analysis are;

- * Maximum a posteriori
- * Particle filter
- * Markov chain Monte Carlo (MCMC)
- * Kalman filter
- * Wiener filter

In particular, the MCMC method was used in this research

[7.0] Results Discussion

The point estimates' summary presented in Tables 1-3 reflects some properties of the two estimation methods under discussion. For all the 3 runs, the Bayesian estimates in the class containing the true value (2.0) are more than the classical estimates, mostly for the small sample case. This performance was the same in all the comparison criteria considered (i.e, frequency distribution of estimates, mean, mean squared error and bias). The distribution of these estimates was closer to Normal in the large samples (see fig.1 and fig.2) for the two estimation methods. As expected, as the sample size increases, the mean squared error of the estimators reduces. In run III, where the variance of the exogenous variable (x_t) was raised to 9, the estimates from the two methods were more concentrated around the class containing the true value than in the first and second runs where the variances were 1 and 2 respectively (fig.3). This is an indication that the distribution of the exogenous variable also affects the properties of the estimates. We noticed that in run I and II, the mean squared error was questionably large for the classical method when $N=20$, this is as a result of outliers that are uncharacteristic of the Bayesian method. An explanation for the outliers was presented by Zellner(1971) that, "the distribution of the estimator $\hat{\gamma}$, given in (5.1), is such that under the conditions underlying run I and II, extreme values can be encountered with a non-negligible probability". These results suggest that the features of the underlying model also influence the bias and consistency of the estimator.

Table 1: Summary of Monte Carlo Experiment, RUN I

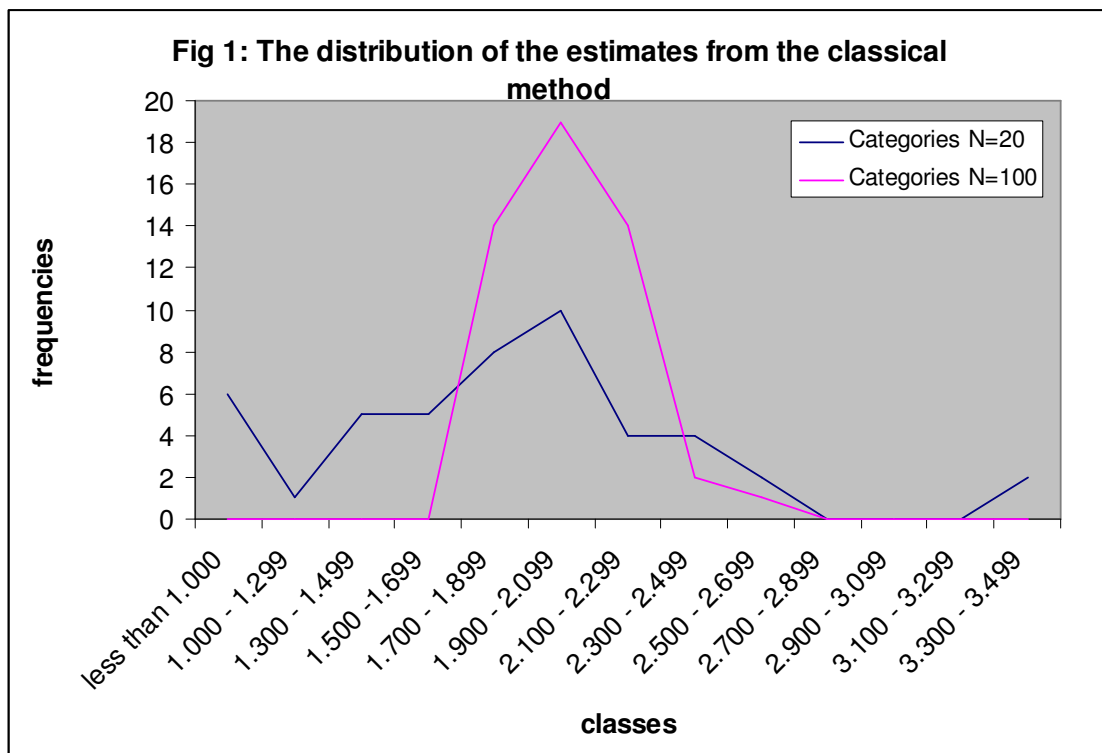
Categories	Frequencies							
	N=20		N=40		N=60		N=100	
	Classical	Bayesian	Classical	Bayesian	Classical	Bayesian	Classical	Bayesian
less than 1.000	6	4	6	1	2	0	0	0
1.000 - 1.299	1	1	2	0	1	0	0	0
1.300 - 1.499	5	4	1	3	1	3	0	0
1.500 - 1.699	5	10	3	7	5	6	0	0
1.700 - 1.899	8	11	12	12	10	10	14	11
1.900 - 2.099	10	11	11	14	16	15	19	20
2.100 - 2.299	4	3	6	5	7	9	14	11
2.300 - 2.499	4	3	6	3	6	4	2	6
2.500 - 2.699	2	2	0	2	0	3	1	1
2.700 - 2.899	0	0	2	3	0	0	0	1
2.900 - 3.099	0	1	1	0	1	0	0	0
3.100 - 3.299	0	0	0	0	0	0	0	0
3.300 - 3.499	2	0	0	0	0	0	0	0
Mean	1.161	1.766	1.39	1.949	2.023	1.977	2.022	2.0698
Mean Squared Error	29.833	0.231	7.032	0.083	1.222	0.112	0.032	0.059
Bias	0.839	0.234	0.61	0.051	0.023	0.023	0.022	0.07

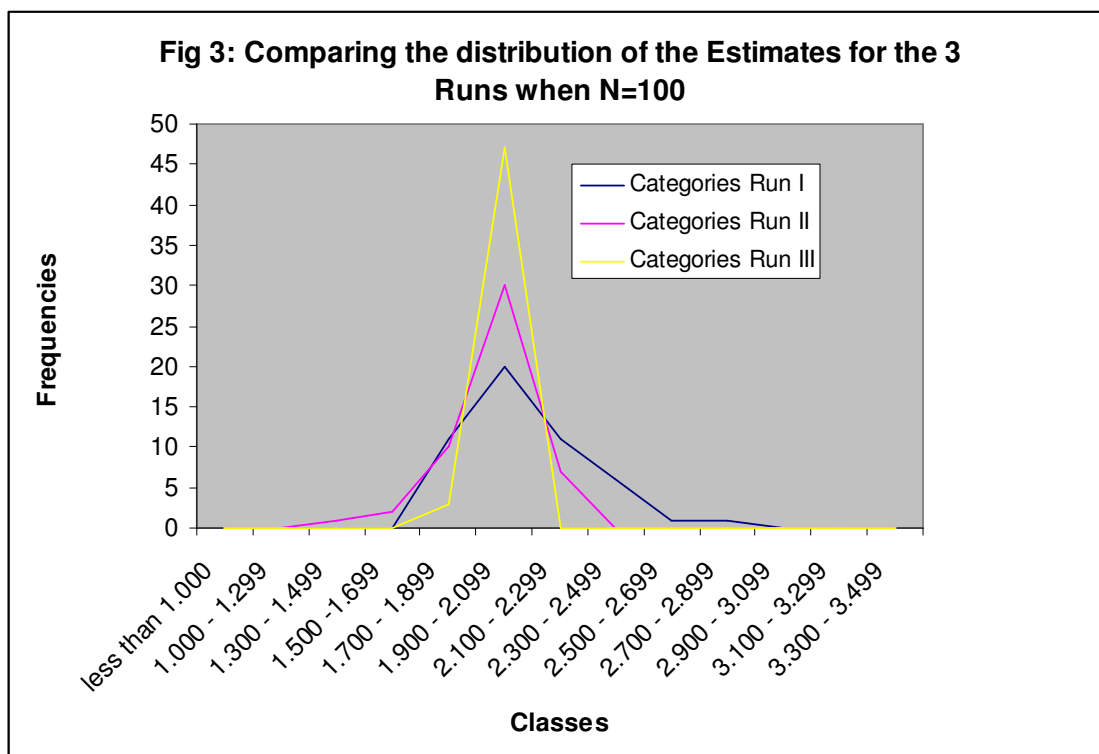
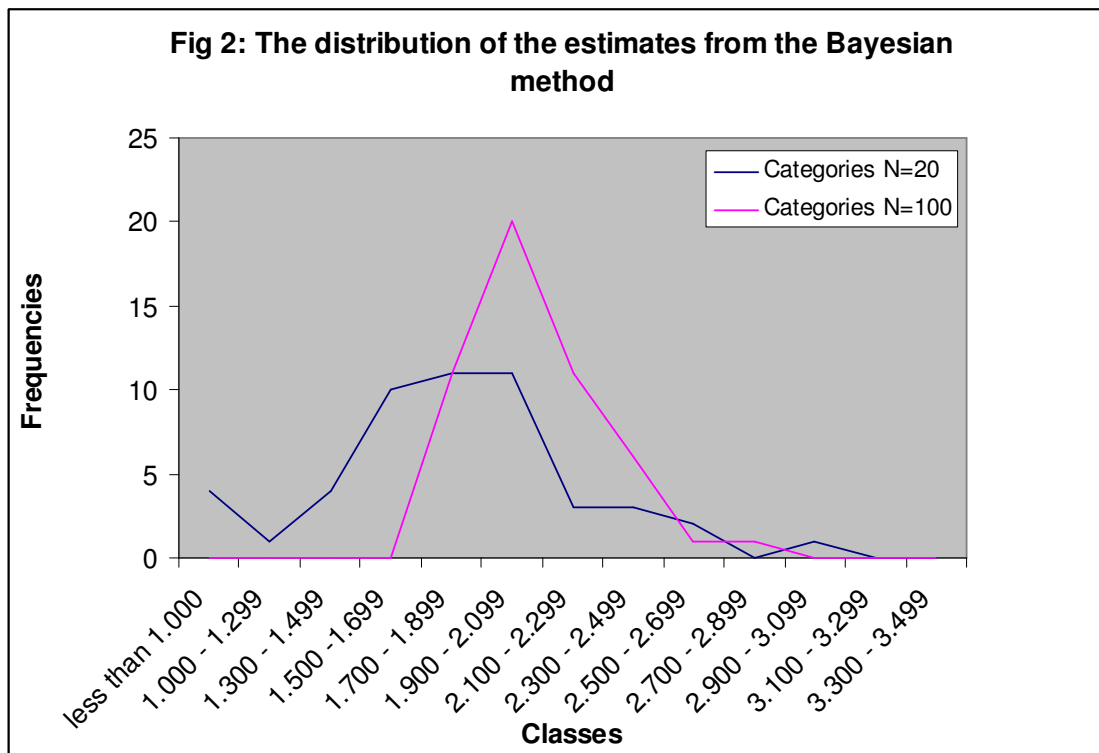
Table 2: Summary of Monte Carlo Experiment, RUN II

Categories	Frequencies							
	N=20		N=40		N=60		N=100	
	Classical	Bayesian	Classical	Bayesian	Classical	Bayesian	Classical	Bayesian
less than 1.099	3	0	5	0	2	0	0	0
1.099 - 1.299	2	0	0	0	0	0	0	0
1.300 - 1.499	0	0	1	0	1	1	1	1
1.500 - 1.699	4	2	1	2	5	1	2	2
1.700 - 1.899	3	3	7	7	11	14	9	10
1.900 - 2.099	12	16	16	17	23	26	29	30
2.100 - 2.299	9	15	17	21	6	6	9	7
2.300 - 2.499	7	10	2	3	2	2	0	0
2.500 - 2.699	2	2	0	0	0	0	0	0
2.700 - 2.899	2	2	0	0	0	0	0	0
2.900 - 3.099	3	0	0	0	0	0	0	0
3.100 - 3.299	1	0	0	0	0	0	0	0
3.300 - 3.499	0	0	0	0	0	0	0	0
Mean	1.505	2.162	1.88	2.058	1.855	1.957	1.97	1.962
Mean Squared Error	15.08	0.094	0.199	0.036	0.245	0.034	0.024	0.024
Bias	0.495	0.162	0.12	0.058	0.145	0.043	0.03	0.038

Table 3: Summary of Monte Carlo Experiment, RUN III

Categories	Frequencies							
	N=20		N=40		N=60		N=100	
	Classical	Bayesian	Classical	Bayesian	Classical	Bayesian	Classical	Bayesian
less than 1.000	0	0	0	0	0	0	0	0
1.000 - 1.299	0	0	0	0	0	0	0	0
1.300 - 1.499	0	0	0	0	0	0	0	0
1.500 - 1.699	0	2	2	2	0	0	0	0
1.700 - 1.899	9	7	9	10	8	10	3	3
1.900 - 2.099	18	23	29	30	35	33	47	47
2.100 - 2.299	10	12	10	8	7	0	0	0
2.300 - 2.499	5	6	2	0	0	0	0	0
2.500 - 2.699	1	0	0	0	0	0	0	0
2.700 - 2.899	3	0	0	0	0	0	0	0
2.900 - 3.099	2	0	0	0	0	0	0	0
3.100 - 3.299	1	0	0	0	0	0	0	0
3.300 - 3.499	0	0	0	0	0	0	0	0
Mean	2.192	2.05	1.978	1.973	2.003	1.996	1.993	1.989
Mean Squared Error	0.199	0.042	0.022	0.019	0.011	0.012	0.003	0.004
Bias	0.192	0.05	0.022	0.027	0.003	0.004	0.007	0.021





[8.0] Conclusion

Estimation of Simultaneous equations 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. Our study here has shown, as expected, that the Bayesian estimation method performs better than the classical for small samples. The results also suggested that the median as a measure of central tendency, gives the same result as the mode as the point estimate of the posterior distribution. However, it is important to put the loss function into consideration, an issue that should be given more research. The classical estimation method being more easily applied, might be a better choice when handling large samples, since it appears to give the same result with the Bayesian approach.

References

Angrist, J.D., Guido W. Imbens and Alan B. Krueger (1999) "Jackknife Instrumental variable Estimation" *Journal of Applied Econometrics*, Vol. 14, No. 1, January/February, pp 57 – 67.

Arnold Zellner (1971) "An Introduction to Bayesian Inference in Econometrics" John Wiley & sons, Inc. pp. 225 -287.

Arnold Zellner (1979) "Statistical Analysis of Econometrics models" *Journal of the American Statistical Association*. Vol. 74, No. 367. Invited Paper. Pp. 628 – 643

Arnold Zellner (1997a) "Bayesian analysis in Econometrics and Statistics: The Zellner view and papers" Edward Elgar Publishing. pp. 250-305

Arnold Zellner (1997b) "The Bayesian method of moments" *Advances in Econometrics*. Vol. 12, pp. 85- 105, Greenwich, CT: JAI Press.

Blomquist, Soren, and Matz Dahlberg (1999) "Small Sample Properties of LIML and Jackknife IV Estimators: Experiments with weak Instruments." *Journal of Applied Econometrics*. Vol. 14, No. 1, January/February, pp. 69 – 88.

Dale J. Poirier (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.

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

Geweke John (1989) "Bayesian inference in Econometric models using Monte Carlo Integration" *Econometrica*, Vol. 57, No.6 pp. 1317-1339

Geweke John (1996) "Bayesian Reduced Rank Regression in Econometrics" *Journal of Econometrics*, Vol. 75, issue 1, pp 121-146

Haavelmo T (1943) "The Statistical implications of a System of Simultaneous Equations" *Econometrica*, Vol. 11, No.1 pp.1-12

Herman K. Van Dijk (2002) "On Bayesian structural inference in a simultaneous equation models" *Economic institute report*, EI 2002-10

Kai Li & dale J. Poirier (2003) "Bayesian analysis of an Econometric model of birth inputs and outputs" *Journal of Population Economics*, Vol 16, pp. 597-625

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

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

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.

T.Kloek & H.K. Van Dijk (1978) "Bayesian Estimates of equation System Parameters, An Application of integration by Monte Carlo". *Econometrica*, Vol 46, No.1, pp. 1-19

Tony Lancaster (2004) "An Introduction to modern Bayesian Econometrics". Blackwell Publishing Ltd

Tony O'Hagan (1994) "Kendall's Advanced Theory of Statistics" Volume 2B- Bayesian Inference. Edward Arnold, London

S. James Press (1969) "The t- Ratio Distribution" *Journal of the American Statistical Association*, Vol. 64. pp. 242-252