

A Comparative Study of the Classical and Bayesian Methods of Estimating a Just-Identified Simultaneous Equations Econometric Model

D. M. Okewole

Department of Mathematical Sciences, Redeemer's University, Ogun State, Nigeria

E-mail: dupedepo@yahoo.com

O. E. Olubusoye

Department of Statistics, University of Ibadan, Oyo State, Nigeria

E-mail: busoye2001@yahoo.com

D.K Shangodoyin

Department of Statistics, University of Botswana, Botswana

E-mail: shangodoyink@mopipi.ub.bw

Received: November 5, 2010 Accepted: December 8, 2010 doi:10.5539/ass.v7n5p222

Abstract

A just identified 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 softwares. 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.

Keywords: Bayesian procedures, Classical procedures, Simultaneous equations, Just-identified model, Monte-Carlo, Simulation, Estimates

1. 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. Simultaneous equations model could be under-identified, just-identified or over-identified, depending on how each parameter of the model uniquely contribute 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, 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.

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 1.1, 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).

In this paper, 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 $X = (\underline{x}_1, \dots, \underline{x}_T)$. Given the probability density function $f(X/\underline{\theta})$, the classical estimation often proceeds by making use of the likelihood function $L(\underline{\theta}) = f(X/\underline{\theta})$, 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 Eviewsc 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.

2. The model

The model analyzed here is

$$\begin{aligned} Y_{1t} &= \gamma Y_{2t} + U_{1t} \\ Y_{2t} &= \beta X_t + U_{2t} \end{aligned} \quad t = 1, 2, \dots, T \quad (2.1)$$

Which is a just identified model 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 (2.1) 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, bias and Mean Squared Error of the estimator.

The identification status of the model (2.1) is easily evident using the order condition, which specifies that for each endogenous variable in the model having a coefficient to be estimated; at least one exogenous variable must be excluded from the equation. Also by the rank condition which states that; in a system of m equations, any particular equation is identified if it is possible to construct at least one non-zero determinant of the order $(m-1)$ from the coefficients excluded from that particular equation but contained in other equations of the system. This is obvious in our model (2.1)

The matrix form of our model (2.1) is

$$Y\Gamma = XB + U \quad (2.2)$$

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 $n \times 1$ vector of observations on the predetermined variables. B is 1×2 vector of coefficients for the predetermined variables, and $U = (U_1, U_2)$ is an $n \times 2$ matrix of random disturbance terms. For easy analysis, we need the reduced form of (3.1) (and later carry out transformations as appropriate), given as;

$$\begin{aligned} Y_{1t} &= \pi_1 X_{1t} + V_{1t} \\ Y_{2t} &= \pi_2 X_{2t} + V_{2t} \end{aligned} \quad (2.3)$$

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 $n \times 2$ matrix of reduced form disturbance term, $V_1 = \gamma U_{2t} + U_{1t}$, $V_2 = U_{2t}$, $\pi_1 = \beta\gamma$, $\pi_2 = \beta$.

3. Methodology

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

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

RUN II

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

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, 1000 samples of size 20, 40, 60, and 100 were generated, making a total of 4,000 samples in one run, and 12,000 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 Eviews for the classical method and WinBUGS for the Bayesian method. (The same data sets for the two methods, the data was generated in Eviews). The following were used in comparing the two estimation methods: Performance in repeated sampling - frequency distribution; Mean Estimate; Bias and Mean Squared Error (MSE), for a parameter θ , is given as;

$$\text{M.S.E}(\theta) = E(\theta - \hat{\theta})^2, \text{ which is also the same as } = \text{Var}(\hat{\theta}) + (\text{Estimated bias})^2$$

$$\text{where } \text{Var}(\hat{\theta}) = \frac{1}{N_r} \sum_{i=1}^{N_r} \hat{\theta}^2 - \left(\frac{1}{N_r} \sum_{i=1}^{N_r} \hat{\theta} \right)^2$$

N_r is number of replications and therefore number of estimates $(\hat{\theta})$.

$$\text{Estimated bias} = \frac{1}{N_r} \sum_{i=1}^{N_r} \hat{\theta} - \theta$$

θ is the true value of the parameter, which in this case is the value used to generate the sample values. 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 for this just identified model is the same, which is given as;

$$\hat{\gamma} = \frac{\sum_{t=1}^T y_{2t} y_{1t}}{\sum_{t=1}^T y_{2t}^2} = \frac{\hat{\pi}_1}{\hat{\pi}_2} \quad (3.1)$$

Where

$$\hat{y}_{2t} = \hat{\pi}_2 x_t,$$

$$\hat{\pi}_2 = \frac{\sum_{t=1}^T x_t y_{2t}}{\sum_{t=1}^T x_t^2}$$

$$\hat{\pi}_1 = \frac{\sum_{t=1}^T x_t y_{1t}}{\sum_{t=1}^T x_t^2}$$

4. Bayesian Estimation method

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;

$$P(\pi, \Sigma^{-1}) \propto |\Sigma^{-1}|^{-3/2} \exp[-1/2 (\pi - \bar{\pi})' C^{-1} (\pi - \bar{\pi})] \quad (4.1)$$

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 $C = (C_{ul})$ is a 2 x 2 matrix of the prior covariance.

(2) Diffuse prior. The idea behind the use of diffuse (otherwise known as 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,

$$P(\pi, \Sigma) = p(\pi) p(\Sigma) \quad (4.2)$$

Using the Jeffrey's invariance theory, we take

$$P(\pi) = \text{constant} \quad (4.3)$$

And

$$P(\Sigma) \propto |\Sigma|^{-3/2} \quad (4.4)$$

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

$$J = \left| \frac{\partial(\sigma_{11}, \sigma_{12}, \sigma_{22})}{\partial(\sigma^{11}, \sigma^{12}, \sigma^{22})} \right| = |\Sigma|^3 \quad (4.5)$$

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

$$P(\Sigma^{-1}) \propto |\Sigma^{-1}|^{-3/2} \quad (4.6)$$

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 (5.3), (5.4), and (5.6). These diffuse prior pdf ‘s were also arrived at by Zellner(1971), Geisser(1965) and others.

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;

$$I(\pi, \Sigma / Y, X) \propto \exp[-1/2 \text{tr}(Y-X\pi)'(Y-X\pi)\Sigma^{-1}] \quad (4.7)$$

This is the same as;

$$I(\pi, \Sigma / Y, X) \propto |\Sigma|^{-N/2} \exp[-1/2 \text{tr} \Sigma^{-1} - 1/2 \text{tr}(\pi - \hat{\pi})'X'X(\pi - \hat{\pi}) \Sigma^{-1}] \quad (4.8)$$

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

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

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

Thus, the likelihood function for the parameters is as given in (4.8).

Combining the diffuse Prior p.d.f (4.3) and (4.6) with the likelihood function (4.8), we have a posterior distribution that is in the bivariate student-t form. This is given as:

$$P(\pi_1, \pi_2 | y_1, y_2) \propto \left[1 + (\pi_1 - \hat{\pi}_1 : \pi_2 - \hat{\pi}_2) \begin{pmatrix} s^{11} & s^{12} \\ s^{21} & s^{22} \end{pmatrix} (\pi_1 - \hat{\pi}_1 : \pi_2 - \hat{\pi}_2)' \right]^{-N/2} \quad (4.9)$$

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,j th element of

$$\left[\sum_{n=1}^N (y_{in} - \hat{\pi}_i x_n)(y_{jn} - \hat{\pi}_j x_n) \right]_{i,j=1,2} \quad i, j=1,2. \text{ To obtain the posterior distribution in terms of } \gamma \text{ and } \beta, \text{ we}$$

carry out the following transformation:

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

$$P(\gamma, \beta | y_1, y_2) \propto |\beta| \left[1 + (\beta\gamma - \hat{\pi}_1)^2 s^{11} + (\beta - \hat{\pi}_2)^2 s^{22} + 2(\beta\gamma - \hat{\pi}_1)(\beta - \hat{\pi}_2) s^{12} \right]^{-N/2} \quad (4.10)$$

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 study. The WinBUGS mentioned earlier is used to obtain the Bayesian estimate by drawing samples from the Posterior distribution and obtaining the mean after ensuring 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 paper

Details of the Bayesian Monte Carlo Simulations

*Two parallel chains were run

*Convergence was checked by monitoring the trace and history plots of the chains for the point of mixture of the two. (Gilks et al, 1996)

*There were 1000 Burn-in after ascertaining convergence of the simulations, then additional 5000 samples were obtained for each chain. The Monte Carlo Error is between 0.01 and 0.9 for the estimates. Each of these MC errors is less than 5% of the corresponding standard deviation, which is another way of confirming the accuracy of the estimates

*The WinBUGS code is given in the appendix

5. Results Discussion

The point estimates' summary presented in Tables 1-3 reflects some properties of the two estimation methods under discussion. For the classical method, the various estimators (least squares, maximum likelihood) were not separately considered because they give the same estimate for this model being a just-identified model. 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.

6. 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, at least for the just identified model, since in this case, all the classical estimation methods give the same estimate of the parameter gamma. 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. The result of this Monte Carlo experiment also emphasized the more on the importance of large samples in Statistical inference

References

Angrist,J.D., Guido W. Imbens and Alan B. Krueger. (1999). Jackknife Instruemental 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. (1997b). The Bayesian method of moments. *Advances in Econometrics*, Vol. 12, pp.85- 105, Greenwich, CT: JAI Press.
- Arnold Zellner. (1997a). *Bayesian analysis in Econometrics and Statistics: The Zellner view and papers*. Edward elgar Publishing, pp.250-305.
- 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.
- Chao, J.C. & 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.
- 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., Zivot Eric. (2003). Bayesian and classical approaches to instrumental variable regression. *Journal of Econometrics*, Vol. 114, pp.29-72.
- Kleibergen, F.& van Dijk, H.K. (1998). Bayesian Simultaneous equation analysis using reduced rank structures. *Econometric Theory*, 14, pp.701-743.
- 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.
- Nagar, A.L. (1959). A Monte Carlo study of Alternative Simultaneous equations Estimators. *Econometrica*, 28, pp.573-590.
- S. Geisser. (1965). A Bayes Approach for Combining Correlated Estimates. *J.Am.Statist. Assoc.*, 60, pp.602 – 607.
- S. James Press. (1969). The t- Ratio Distribution. *Journal of the American Statistical Association*, Vol. 64. pp.242-252.
- 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.
- W. R. Gilks, S.Richardson and D.j. Spiegelhalter. (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hall/CRC.

Appendix

(A). WinBUGS code for the Bayesian analysis

```
model {
  for(j in 1:N) {
    for (i in 1: T) {
      y[j, i,1:2 ] ~ dmnorm(mu[ j, i,1:2], tau[,])
      mu[ j, i,1] <- gamma[j]*mu[ j, i, 2]
      mu[j, i, 2] <- beta[j]*x[i]
    }
    #Priors
    gamma[j] ~ dnorm(0, 0.001)
    beta[j] ~ dnorm(0, 0.001)
  }
  tau[1:2,1:2] ~dwish(R[ ,], 2)
  R[1,1] <- 0.01
  R[1,2] <- 0
  R[2,1] <- 0
  R[2,2] <- 0.01
  sigma[1:2,1:2] <- inverse(tau[1:2,1:2 ])
}
```

(B). Eviews code for the classical analysis

```
'store monte carlo results in matrix
'create workfile
create simult1 u 20
'create data series for xt
series xt=@nrnd(1)
'set seed for random number generator
rndseed 123456
'assign number of replications to a control variable
!reps = 1000
'declare storage matrix
matrix (!reps, 2) m
'begin loop
for !i =1 to !reps
  'create data series for et
  series e1t=@nrnd(1)
  series e2t=@nrnd(1)
  'create data series for ut
  series u1t=.866025*e1t - 0.5*e2t
  series u2t = 2*e2t
  'simulate y data
  model
```

```
series yt!i1 = xt + u1t + 2*u2t
series yt!i2 = 0.5*xt + u2t
solve
'regress y on x
system statis1!i
statis1!i.append yt!i1=c(1)* xt
statis1!i.append yt!i2=c(2)* xt
statis1!i.append inst xt
statis1!i.ls
'obtain mse
'store coefficients and variance estimate in matrix
m(!i, 1) = statis1!i.@coefs(1)
m(!i, 2) = statis1!i.@coefs(2)

next
'end of loop
'expand workfile
expand
smpl 1 1000
'change matrix m to series
mtos(m,group1)
rename ser01 gamma
rename ser02 beta
gamma.freq(b=10,n,obs,count,dropna)
series summary
summary.fill @mean(beta), @mean(gamma), @cov(e1t, e2t), @cov(u1t, u2t),@mean(e1t), @mean(e2t),
@var(u1t), @var(u2t)
```

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 0	43	0	32	0	9	0	5	0
0 - 1.099.	47	0	28	0	11	0	2	0
1.100 - 1.299	33	5	9	28	5	26	0	2
1.300 - 1.499	50	31	36	52	13	34	2	1
1.500 -1.699	132	61	96	140	56	120	38	2
1.700 - 1.899	179	120	232	240	222	223	256	215
1.900 - 2.099	186	138	223	280	313	300	383	392
2.100 - 2.299	109	133	150	100	175	157	191	210
2.300 - 2.499	69	134	62	75	70	65	60	121
2.500 - 2.699	28	97	34	25	44	60	30	28
2.700 - 2.899	33	79	23	21	19	9	8	12
2.900 - 3.099	16	53	16	14	16	3	8	5
3.100 - 3.299	11	45	10	7	8	2	6	1
3.300 - 3.499	8	38	9	5	8	1	1	1
3.500 - 3.699	8	27	3	4	1	0	1	3
3.700 - 3.899	8	14	5	4	2	0	2	1
3.900 - 4.099	3	10	4	1	2	0	2	1
4.100 - 4.299	3	7	3	2	2	0	0	2
4.300 - 5.001	4	7	7	2	7	0	1	1
5.001 - 10.000	22	1	14	0	12	0	3	1
10.001-50.000	7	0	4	0	4	0	1	1
50.001-100.000	1	0	0	0	1	0	0	0
Mean	1.955	2.417	1.9703	1.842	1.9549	2.0311	2.0329	2.0257
Mean Squared Error	26.5562	0.58	5.03585	1.02745	4.978	0.5283	0.51057	0.3781
Bias	0.455	0.417	0.02967	0.158	0.045	0.0311	0.0329	0.0257

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 0	38	0	34	0	22	0	22	0
0 - 1.099.	44	0	56	0	51	0	51	0
1.100 - 1.299	17	6	19	11	20	3	20	0
1.300 - 1.499	29	6	44	12	36	17	36	20
1.500 -1.699	78	22	87	17	73	20	73	51
1.700 - 1.899	134	60	121	140	140	280	140	220
1.900 - 2.099	217	326	264	351	259	511	259	569
2.100 - 2.299	252	278	213	380	239	100	239	71
2.300 - 2.499	87	218	78	44	88	36	88	65
2.500 - 2.699	33	49	23	21	27	19	27	2
2.700 - 2.899	18	33	10	13	5	8	5	2
2.900 - 3.099	9	1	3	6	8	5	8	0
3.100 - 3.299	6	1	6	2	6	1	6	0
3.300 - 3.499	6	0	7	1	2	0	2	0
3.500 - 3.699	1	0	5	1	2	0	2	0
3.700 - 3.899	4	0	3	1	2	0	2	0
3.900 - 4.099	2	0	2	0	0	0	0	0
4.100 - 4.299	1	0	1	0	2	0	2	0
4.300 - 5.001	4	0	3	0	4	0	4	0
5.001 - 10.000	10	0	13	0	9	0	9	0
10.001-50.000	7	0	7	0	3	0	3	0
50.001-100.000	2	0	1	0	1	0	1	0
Mean	2.1786	2.1253	1.86501	1.8922	1.9328	1.9201	1.9019	1.9012
Mean Squared Error	8.492	1.8621	7.8594	0.9432	0.61705	0.60511	1.5869	0.8435
Bias	0.1786	0.1253	0.135	0.1078	0.06715	0.0799	0.0981	0.0948

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 0	37	0	0	0	0	0	0	0
0 - 1.099.	37	10	1	0	0	0	0	0
1.100 - 1.299	35	14	1	9	1	1	0	0
1.300 - 1.499	35	26	2	12	2	1	0	0
1.500 - 1.699	62	65	14	31	11	21	1	0
1.700 - 1.899	120	305	212	138	181	167	85	101
1.900 - 2.099	157	318	639	645	668	671	869	871
2.100 - 2.299	196	182	129	146	137	120	45	28
2.300 - 2.499	141	40	2	16	0	19	0	0
2.500 - 2.699	45	40	0	3	0	0	0	0
2.700 - 2.899	32	0	0	0	0	0	0	0
2.900 - 3.099	20	0	0	0	0	0	0	0
3.100 - 3.299	10	0	0	0	0	0	0	0
3.300 - 3.499	14	0	0	0	0	0	0	0
3.500 - 3.699	7	0	0	0	0	0	0	0
3.700 - 3.899	2	0	0	0	0	0	0	0
3.900 - 4.099	3	0	0	0	0	0	0	0
4.100 - 4.299	7	0	0	0	0	0	0	0
4.300 - 5.001	13	0	0	0	0	0	0	0
5.001 - 10.000	14	0	0	0	0	0	0	0
10.001-50.000	12	0	0	0	0	0	0	0
50.001-100.000	0	0	0	0	0	0	0	0
Mean	1.9724	1.974	1.9757	1.9771	1.9861	1.987	1.9922	1.9924
Mean Squared Error	0.05582	0.0452	0.01561	0.01384	0.01275	0.0126	0.00464	0.00461
Bias	0.0276	0.026	0.0243	0.0229	0.0139	0.013	0.0078	0.0076

Fig 1: the Distribution of the Estimates from the classical method

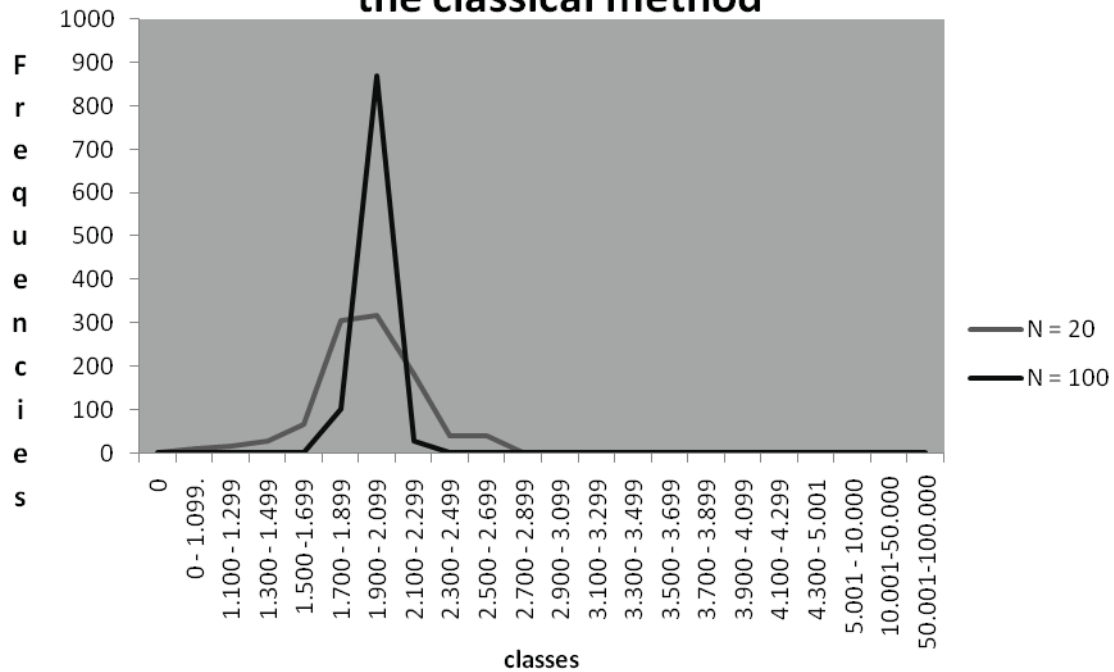


Fig 2: The Distribution of the Estimates from the Bayesian method

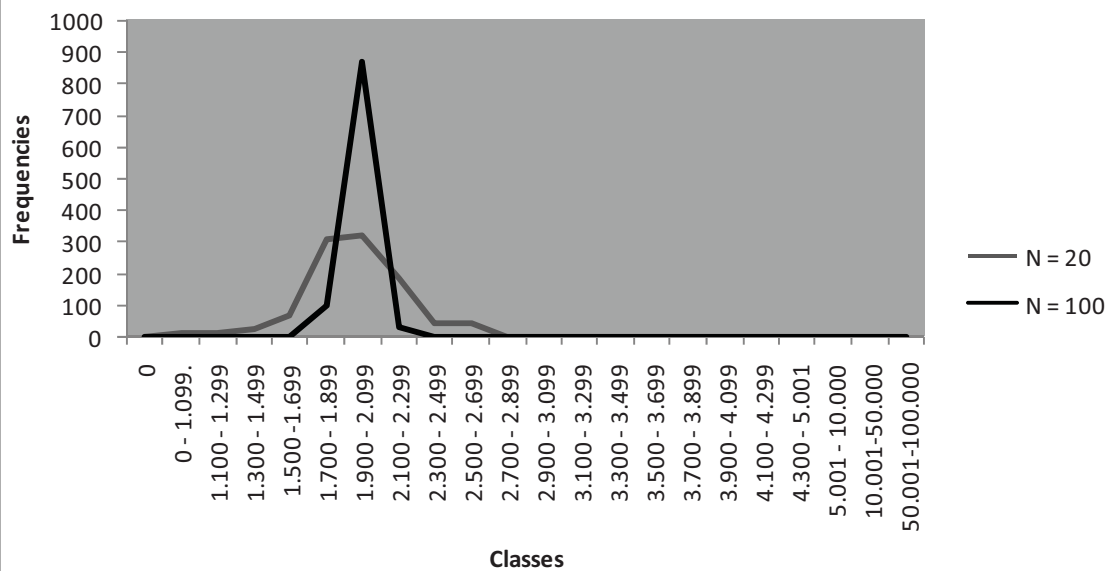


Fig 3: Comparing the distribution of the estimates for the 3 runs when N=100

