Journal of Mathematics and Statistics 5 (4): 260-266, 2009
ISSN 1549-3644

Ranking of Simultaneous Equation Techniques to Small Sample Properties and Correlated Random Deviates

[1]A.A. Adepoju and [2]J.O. Olaomi
[1]Department of Statistics, University of Ibadan, Ibadan
[2]University for Development Studies, Tamale, Ghana

Abstract: Problem statement: All simultaneous equation estimation methods have some desirable asymptotic properties and these properties become effective in large samples. This study is relevant since samples available to researchers are mostly small in practice and are often plagued with the problem of mutual correlation between pairs of random deviates which is a violation of the assumption of mutual independence between pairs of such random deviates. The objective of this research was to study the small sample properties of these estimators when the errors are correlated to determine if the properties will still hold when available samples are relatively small and the errors were correlated. **Approach:** Most of the evidence on the small sample properties of the simultaneous equation estimators was studied from sampling (or Monte Carlo) experiments. It is important to rank estimators on the merit they have when applied to small samples. This study examined the performances of five simultaneous estimation techniques using some of the basic characteristics of the sampling distributions rather than their full description. The characteristics considered here are the mean, the total absolute bias and the root mean square error. **Results:** The result revealed that the ranking of the five estimators in respect of the Average Total Absolute Bias (ATAB) is invariant to the choice of the upper (P_1) or lower (P_2) triangular matrix. The result of the FIML using RMSE of estimates was outstandingly best in the open-ended intervals and outstandingly poor in the closed interval ($-0.05<r<0.05$) when P_1 and P_2 was re-combined. **Conclusion:** (i) The ranking of the various simultaneous estimation methods considered based on their small sample properties differs according to the correlation status of the error term, the identifiability status of the equation and the assumed triangular matrix. (ii) The nature of the relationship under study also determined which of the criteria for judging the performances of the estimators could be said to perform best when compared with others.

Key words: Monte Carlo, random deviates, mutual correlation, total absolute bias, root mean square error

INTRODUCTION

The simultaneous equations model is most important to econometricians both from a theoretical as well as applied perspective. It is unfortunate that the estimators employed have exact, finite-sample distribution that are difficult to derive. Thus, their properties are usually discussed only on the basis of large sample theory. Small sample properties have been studied using Monte Carlo techniques by many authors including Johnston[2,3,5,11]. However, these studies cannot sort out the possibly complex dependence of the distributions on unknown parameters, nor do they reveal the possibility that moments of the exact distribution do not exist, making comparisons of empirical mean square errors, biases and sampling variances meaningless.

The theoretical ranking of the various simultaneous estimation techniques on the basis of the asymptotic properties is important if the sample size is sufficiently large. However, given that in practice the researcher works usually with small samples, the asymptotic properties of the estimates are of little assistance in his choice of technique[8]. What is important is the ranking of the estimators on the merit they have when applied to small samples. Conventionally, the ranking has been based on some 'small-sample properties' which are considered as 'desirable' or 'optimal' for the estimate to possess. The properties considered in this study are:

Corresponding Author: A.A. Adepoju, Department of Statistics, University of Ibadan, Ibadan

- Average of estimates
- Absolute bias of estimates and
- Root mean square error

The question that is often asked is which of the criteria is the most important? Should we prefer an estimate with the smallest bias or minimum variance or if it has the smallest mean square error? There is no law that says that bias or efficiency should be ranked in some unique order. Much depends on the nature of the relationship being studied and the purpose it is going to serve. In some cases, the minimum variance may be more desirable than small bias, while in some cases the least bias may be the most desirable property to be possessed by an estimator Koutsoyiannis[8]. Obviously, the importance of each criterion is to a certain extent a matter of subjective decision of the econometrician. Cragg[2] noted that the standard errors of the consistent methods would lead to reliable inferences, but this was not always the case as the standard errors of the OLS are not useful for making inference about the true values of the parameters. Summers[10] reached a very similar conclusion that the OLS method is inferior to the consistent methods of estimation. The presence of autocorrelation in the structural disturbances leaves unchanged the ranking of the system estimators established under textbook assumptions and appears to have little effect on their bias Cragg[2]. The "typical" specification of serial independence of the errors in simultaneous equations models has been recently extended to include the possibility of auto correlated errors. The degree and type of auto correlation among the errors are reported by[6] as very vital.

Model specification: Consider the following model:

$$Y_{1t} = \beta_{12} Y_{2t} + \gamma_{11} X_{1t} + U_{1t}$$

$$Y_{2t} = \beta_{2t} Y_{1t} + \gamma_{22} X_{2t} + \gamma_{23} X_{3t} + U_{2t}$$

Where:
Y's = The endogenous variables
X's = The predetermined variables
u's = The random disturbance terms
β's and γ's = The parameters

Three levels of assumed correlation between pairs of random deviates are considered as follows:

- Negatively correlated interval $\left(r_{\varepsilon_1,\varepsilon_2} < -0.05\right)$
- Feebly negatively or positively correlated interval $\left(-0.05 < r_{\varepsilon_1,\varepsilon_2} < +0.05\right)$
- Positively correlated interval $\left(r_{\varepsilon_1,\varepsilon_2} > +0.05\right)$

The pairs of random sequences $\{(\varepsilon_{1t}, \varepsilon_{2t}); t = 1,..., N\}$ are generated such that the disturbance terms are distributed N(0, Σ).

Data generation: In econometrics, while asymptotic properties of estimators obtained by various econometric techniques are deduced from postulates or self-evident assumptions, the small sample properties of the various econometric techniques have been studied from simulated data in what are known as Monte Carlo studies and not with direct application of the techniques to actual observations. This approach is due to the fact that actual observations on economic variables are often plagued with problems such as multi collinearity, autocorrelation, errors of measurements, non spherical disturbances and other economic "diseases" simultaneously. All the estimators whose small sample properties are studied here are based on the assumption that all these problems are absent, thus such studies cannot be successfully studied using a real life data. The Monte Carlo approach allows the experimenter to set up an artificial system where values are generated for the random disturbances for some specified sample size and using these values, values are calculated for the endogenous variables based on the assumptions of this artificial problem at each sample point[4]. Pretending that the parameters are unknown and using only the values of the endogenous and predetermined variables at each sample point, several estimating techniques are applied in turn to obtain associated estimates of the parameters. The process of generating values for the disturbances, obtaining values for the endogenous variables and calculating estimates of the parameters is repeated, or replicated, a large number of times. The set of estimates of each parameter by each estimator are then used to infer properties of the estimators for the given sample size and for the chosen values of the parameters[4]. This study uses this method with sample size, N, chosen to be N = 40 and replicated 100 times. The following values are arbitrarily assigned to the structural parameters; β_{12} = 1.5, β_{21} = 1.8, γ_{11} = 1.2, γ_{22} = 0.5, γ_{23} = 2.0[1] and

values are arbitrarily assigned to the covariance matrix of the disturbance terms as follows:

$$\Omega = \begin{bmatrix} 5.0 & 2.5 \\ 2.5 & 3.0 \end{bmatrix}$$

Fixed values are generated for the exogenous variables X_{1t}, X_{2t} and X_{3t} from the uniform (1, 0) distribution[7]. Furthermore, the pairs of random sequences $\{(\varepsilon_{1t}, \varepsilon_{2t})\}$ generated are then used to obtain values for the random disturbances U_{1t} and U_{2t} such that they are consistent with the covariance matrix Ω given above. A method presented by Nagar[9] for the transformation of W independent series of standard random deviates of length N into W series of random variables with zero means and a specified covariance matrix is used.

Σ is therefore decomposed by a non- singular upper triangular matrix P such that:

$$\Omega = PP'$$

So that:

$$P_1 = \begin{pmatrix} D_{11} & D_{12} \\ 0 & D_{22} \end{pmatrix}$$

Then:

$$\begin{pmatrix} D_{11} & D_{12} \\ 0 & D_{22} \end{pmatrix}\begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$$

$$D_{22} = +\sqrt{\sigma_{22}}$$

$$D_{12} = \sigma_{12} / D_{22}$$

$$D_{11} = +\sqrt{\sigma_{11} - D_{12}^2}$$

The random disturbance series are obtained as follows:

$$U_t = P_1\begin{pmatrix} D_{11} & D_{12} \\ 0 & D_{22} \end{pmatrix}\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

$$\begin{pmatrix} U_{1t} \\ U_{2t} \end{pmatrix} = \begin{pmatrix} +\sqrt{\sigma_{11} - D_{12}^2} & \sigma_{12} / D_{22} \\ 0 & +\sqrt{\sigma_{22}} \end{pmatrix}\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

Using the values of the covariance matrix, we have:

$$U_{1t} = 1.707825128\varepsilon_{1t} + 1.443375673\varepsilon_{2t}$$

$$U_{2t} = 1.732050808\varepsilon_{2t}$$

The above procedure is repeated for the lower triangular matrix, P_2, such that:

$$P_2 = \begin{pmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{pmatrix}$$

MATERIALS AND METHODS

The main task in the present context is the generation of stochastic dependent (endogenous) variables, Y_{it} (i = 1,2, t = 1,…,T) which are subsequently used in estimating the parameters of the model.

To achieve this, the following have to be assumed:

- Values of the predetermined variables X_{1t}, X_{2t} and X_{3t} (t = 1,…, T)
- Values of the parameters $\beta_{12}, \beta_{21}, \gamma_{11}, \gamma_{22}$ and γ_{23}
- Values of the elements of Ω, the variance-covariance matrix of U, i.e., σ_{11}, σ_{12} and σ_{22}

Essentially, attention is focused on computing numerical values for Y_{it} (i =1, 2, t = 1,…,T) using the reduced form of equations:

$$y_{1t} = \Pi_{11}X_{1t} + \Pi_{12}X_{2t} + \Pi_{13}C_{3t} + V_{1t}$$

$$y_{2t} = \Pi_{21}X_{1t} + \Pi_{22}X_{2t} + \Pi_{23}C_{3t} + V_{2t}$$

Where:

$$\Pi = \begin{bmatrix} \gamma_{11}\beta^* & \beta_{21}\gamma_{22}\beta^* & \beta_{21}\gamma_{23}\beta^* \\ \beta_{12}\gamma_{11}\beta^* & \gamma_{22}\beta^* & \gamma_{23}\beta^* \end{bmatrix}$$

$$V_{1t} = (u_{1t} + \beta_{21}u_{2t})\beta^*$$

$$V_{2t} = (\beta_{12}u_{1t} + u_{2t})\beta^*$$

$$\beta^* = (1 - \beta_{12}\beta_{21})^{-1}$$

The most complex step in generating stochastic dependent variables is the simulation of the error terms U_{it} (i = 1, 2; t = 1,…, T) where selection of only pairs of ε_{it} which fall into one of these three categories above is made.

RESULTS

One of the objectives of this study is to identify which of the three levels of correlation coefficient between the error terms accommodates best estimates of the parameters produced by each of the five estimators. It is also of interest to compare the distribution of 'best' estimates for the two equations and for P_1 and P_2. To achieve this objective 'best' estimates are identified and presented by summarizing the ranking of the various techniques when there is mutual correlation between the disturbance terms in the model for cases of upper and lower triangular matrices (P_1 and P_2). Table 1-11 are generated from the results of the Monte Carlo studies carried out as outlined above in section 3.0. The following criteria are considered for judging the performances of the estimators; average of estimates, absolute bias of estimates and root mean squared error of parameter estimates.

On the criterion of Average, here 'best' estimate implies that estimate that is closest to the true parameter value.

In Table 1, the best method is the 3SLS closely followed by OLS at the negatively correlated interval while 2SLS performed poorly. Whereas OLS maintained its position at the positively correlated region, the 3SLS is the poorest. At the feebly correlated region, 2SLS method ranks highest while FIML appears at the bottom of the list. The FIML appears to be the best method when the errors are positively correlated.

From Table 2, it was observed that for the lower triangular matrix, FIML retained its position as the best method when dealing with positively correlated errors. OLS however, moved to the top position while 3SLS appears to be the least important at the negatively correlated interval. At closed interval, 3SLS is best while OLS is the poorest.

Collapsing the Table 1 and 2 and looking at the general performance of these methods when the triangular matrices are unimportant, we have the Table 3.

Table 1: Rank of estimators using Average (P_1)

r<-0.05	-0.05<r<0.05	r>0.05
3SLS	2SLS	FIML
OLS	3SLS	OLS
FIML	LIML	LIML
LIML	OLS	2SLS
2SLS	FIML	3SLS

Table 2: Rank of estimators using average (P_2)

r<-0.05	-0.05<r<0.05	r>0.05
OLS	3SLS	FIML
LIML	LIML	2SLS
FIML	2SLS	OLS
2SLS	FIML	3SLS
3SLS	OLS	LIML

In Table 3, the ranking also shows that while OLS ranks high as best estimator of error terms with large negative or positive correlation, 3SLS is best with feebly correlated error terms. The ranking of estimators in which P_1 and P_2 are combined is dominated by the ranking obtained under P_2. In that Table 3, OLS ranks high in the two open intervals while 3SLS ranks high in the closed interval where the error terms are feebly correlated.

The next Table 4 and 5 contain summaries of the performance of estimators using total absolute bias of estimates. The criterion is to consider an estimator as 'best' if it produces the smallest total absolute bias out of the three levels of correlation coefficient.

Here in Table 4, OLS is the poorest method on the criterion of bias at the negatively correlated area. Note that we make use of the 'estimated bias' as our criterion, which is the difference between the mean of the estimates and the true value of the parameters i.e. Absolute Bias $(\theta) = \left|\hat{\bar{\theta}} - \theta\right|$. OLS ranked best in the other two intervals while it changed positions with the 3SLS at the open-ended intervals. The performances of all the estimators are not different in the middle interval where the ranks are the same.

In Table 5, When the errors are not strongly correlated 3SLS puts up the best performance and FIML seems to be inferior in this group. Nonetheless, in the other open-ended intervals FIML is outstandingly best.

In order to know which of the three intervals of the correlation coefficient house the 'best' estimates of each parameter produced by each estimator, the following method is adopted; the interval that produces the minimum estimate of RMSE (smallest root mean square error) is counted as the one that accommodates the 'best' estimate, this is shown in Table 6.

Table 3: Rank of estimators using average (P_1 and P_2 combined)

r<-0.05	-0.05<r<0.05	r>0.05
OLS	3SLS	FIML
FIML	2SLS	OLS
LIML	LIML	2SLS
3SLS	OLS	LIML
2SLS	FIML	3SLS

Table 4: Relative importance of estimators using absolute bias (P_1)

r<-0.05	-0.05<r<0.05	r>0.05
3SLS	OLS	OLS
2SLS/LIML	2SLS/LIML/3SLS/FIML	FIML
FIML		LIML/2SLS
OLS		3SLS

Table 5: Relative importance of estimators using bias (P_2)

r<-0.05	-0.05<r<0.05	r>0.05
FIML	3SLS	FIML/OLS
LIML/3SLS	OLS/2SLS 2SLS/LIML	
2SLS	LIML	3SLS
OLS	FIML	

Table 6: Ranking of estimators using RMSE (P_1)

r<-0.05	-0.05<r<0.05	r>0.05
FIML	3SLS	FIML
OLS	2SLS	OLS/2SLS
3SLS	LIML	LIML
2SLS	OLS	3SLS
LIML	FIML	

Table 7: Ranking of estimators using RMSE (P_2)

r<-0.05	-0.05<r<0.05	r>0.05
OLS	OLS	3SLS
LIML/FIML	LIML	2SLS
2SLS	2SLS	FIML
3SLS	3SLS/ FIML	LIML
		OLS

Table 8: Ranking of estimators using RMSE (P_1 and P_2 combined)

r<-0.05	-0.05<r<0.05	r>0.05
FIML	OLS/2SLS/3SLS	FIML
OLS	LIML	
	2SLS/3SLS	
LIML	FIML	LIML
2SLS/3SLS		OLS

The performance of FIML on the criterion of RMSE is similar to its performance on the criterion of bias when upper triangular matrix (P_2) is assumed as shown in Table 6. In Table 6, FIML put up the best performance when the error terms are either negatively or positively correlated, closely followed by OLS in both regions. They both performed poorly in the middle interval. 3SLS ranks first in the middle interval and ranks least at the positively correlated area. LIML ranks last in the interval when the errors are negatively correlated.

As shown in Table 7, when the upper triangular matrix is assumed and the error terms are either negatively or feebly correlated, OLS has an outstanding performance. It however, performs badly at the positively correlated region. In the regions where OLS ranks best, 3SLS ranks last while it performs best where OLS seems to be least. Again, FIML is non existent in the closed interval which implies that its performance at this interval is regardless of the triangular matrix assumed.

Table 9: Summary of the ranking of estimators of average and RMSE P_1

r<-0.05		-0.05<r<0.05		r>0.05	
AVERAGE	RMSE	AVERAGE	RMSE	AVERAGE	RMSE
3SLS	FIML	2SLS	3SLS	FIML	FIML
OLS	OLS	3SLS	2SLS	OLS	OLS/2SLS
FIML	3SLS	LIML	LIML	LIML	LIML
LIML	2SLS	OLS	OLS	2SLS	3SLS
2SLS	LIML	FIML	FIML	3SLS	

Table 10: Summary of the ranking of estimators of average and RMSE P_2

r<-0.05		-0.05<r<0.05		r>0.05	
AVERAGE	RMSE	AVERAGE	RMSE	AVERAGE	RMSE
OLS	OLS	3SLS	OLS	FIML	3SLS
LIML	LIML/FIML	LIML	LIML	2SLS	2SLS
FIML	2SLS	2SLS	2SLS	OLS	FIML
2SLS	3SLS	FIML	3SLS/FIML	3SLS	LIML
3SLS		OLS		LIML	OLS

Table 11: Ranking of estimators under P_1 and P_2 on ATAB and CV

ATAB		CV	
P_1 P_2	P_1	P_2	
3SLS	3SLS	OLS	2SLS
LIML	LIML	2SLS	LIML
2SLS	2SLS	LIML	OLS
OLS	OLS	3SLS	3SLS
FIML	FIML	FIML	FIML

A comparison of the results for combined P_1 and P_2 in Table 8 shows that:

- FIML is outstandingly best in the open-ended intervals and outstandingly poor in the closed interval
- OLS is reasonably good at all intervals except the third interval (r>0.05)
- 3SLS is the most ubiquitous of the five estimators in term of its position in the three intervals under P_1 and P_2
- Entries in this table show that P_2 dominates results of combined P_1 and P_2

Comparable results of the ranking of estimators: It is of interest to study the extent to which the ranking of estimator performance on several criteria are in agreement or otherwise.

Comparisons of the entries in the following Table 9 and 10 reveal some agreement in the ranking of these estimators in the middle (closed) and third interval

(-0.05<r<0.05 and r>0.05) for P_1 and in the first interval (r<-0.05) for P_2.

Using the Average Total Absolute Bias (ATAB) and its Coefficient of Variation (CV) the five estimators are ranked as follows in increasing order of bias and coefficient of variation under P_1 and P_2 as these can be seen in Table 11.

It is noteworthy in Table 11 that in respect of average absolute bias that the five estimators rank uniformly under P_1 and P_2. This finding clearly shows that the ranking of the estimators in terms of the magnitude of the average total absolute bias is invariant to the choice of the upper (P_1) or lower (P_2) triangular matrix.

It is also remarkable that whereas the average absolute biases of the other four estimators take the first four positions, those of FIML maintain a very distant fifth position. The poor ranking of FIML in this situation of correlated disturbances and over-identified equation may be attributed to the fact that it uses more information as an estimator than any of the other four estimators as this is clearly shown in Table 11.

DISCUSSION

Evidently, the results presented from Table 1-11 show that the only remarkable uniformity in the ranking of estimators on the dispersion of the total absolute bias is the fact that the 3SLS and FIML are in the fourth and fifth positions respectively under P_1 and P_2 as presented in Table 11.

Finally, a decision on the best estimator for this model cannot be taken on the basis of our findings on total absolute bias alone. This is because the yardstick is the total absolute bias of two equations, which differ in their identifiability status. In estimating multi-equation models, the choice of estimator is equation specific. Hence, the findings here will have to be reconciled with findings elsewhere before a prescription of best estimator of each equation can be suggested.

CONCLUSION

None of the criteria considered here for judging the performances of the five estimation techniques can be proved to rank better than the others. For example, no law says that bias or efficiency should be ranked in some unique order. Much depends on the nature of the relationship being studied and the purpose it is going to serve. While a minimum variance may be more desirable in some cases than small bias, the least bias may be a more desirable property to be possessed by our estimators in some other cases. It is important to note that the importance of each criterion is to a certain extent a matter of subjective decision of the researcher.

The ranking of the various simultaneous estimation techniques considered on the basis of their small sample properties differs according to the correlation status of the error term (whether the errors are positively correlated, negatively correlated or feebly correlated), the identifiability status of the equation and the assumed triangular matrix (P_1 or P_2).

REFERENCES

1. Adepoju, A.A., 2009. Comparative performance of the limited information technique in a two-equation structural model. Eur. J. Sci. Res., 28: 253-265.
2. Cragg, J.G., 1967. On the relative small-sample properties of several structural-equation estimators. Econometrica, 35: 89-109. http://www.jstor.org/stable/1909385
3. Dhrymes, P., 1974. Econometrics: Statistical Foundations and Applications. Springer-Verlag, New York, ISBN: 0387900950, pp: 592.
4. Gujarati, D.N., 1995. Basic Econometrics. 3rd Edn., McGraw-Hill, Inc., New York. ISBN: 0-07-025214-9, pp: 85-86.
5. Johnston, J., 1972. Econometric Methods. 2nd Edn., McGraw-Hill, New York, ISBN: 0070326851, pp: 437.
6. Formby, T.B., R.C. Hill and S.R. Johnson, 1988. Advanced Econometrics Methods. Springer-Verlag, New York, ISBN: 0387968687, pp: 624.
7. Kmenta, J., 1971. Elements of Econometrics. Macmillian, New York, ISBN: 0023650605, pp: 655.
8. Koutsoyiannis, A., 2003. Theory of Econometrics: An introductory Exposition of Econometric Methods. 2nd Edn., Palgrave Publishers Ltd., New York, USA., ISBN: 0-333-77822-7, pp: 498-511.

9. Nagar, A.L., 1960. A monte carlo study of alternative simultaneous equation estimators. Econometrica, 28: 573-590. http://www.jstor.org/pss/1910132
10. Summers, R., 1965. A capital intensive approach to the small sample properties of various simultaneous equation estimators. Econometrica, 33: 1-41. http://ideas.repec.org/p/cwl/cwldpp/64.html
11. Wagner, H.M., 1958. A Monte Carlo study of estimates of simultaneous linear structural equations. Econometrica, 26: 117-133. http://www.jstor.org/stable/1907386