Trends in

Applied Sciences Research

ISSN 1819-3579

Academic Journals Inc.

www.academicjournals.com

Trends in Applied Sciences Research 2 (4): 354-359, 2007
ISSN 1819-3579

A Comparative Study of the OLS and Some GLS Estimators When Normally Distributed Regressors Are Stochastic

[1]Kayode Ayinde and [2]R.A. Ipinyomi
[1]Department of Pure and Applied Mathematics,
Ladoke Akintola University of Technology, P.M.B. 4000, Ogbomoso,
Oyo State, Nigeria
[2]Department of Statistics, University of Ilorin, P.M.B. 1515, Ilorin,
Kwara State, Nigeria

Abstract: Regressors are assumed fixed (non-stochastic) in repeated samples in the Classical Linear Regression Model. Situations where this assumption is not tenable are often found economics and other social sciences. In this study, we made a comparative study of the Ordinary Least Squares (OLS) and some Feasible Generalized Least Squares (GLS) estimators when normally distributed regressors are stochastic using Monte Carlo methods under both low and high replications. Comparison was done by examining the small sample performances of the estimators via bias, absolute bias, variance and more importantly the mean squared error of the estimated model parameters. Results show that the performances of the estimators improve with increased replication. The Maximum Likelihood (ML) and the Maximum Likelihood Grid (MLGD) estimators only compete with the OLS estimator when replication is low. However with increased replication, the OLS estimator is most efficient among the estimators in estimating all the parameters of the model.

Key words: Stochastic regressors, OLS estimator, feasible GLS estimators

INTRODUCTION

One of the fundamental assumptions about the independent variables (regressors) of the Classical Linear Regression Model (CLRM) is that the regressors are non-stochastic. By this assumption, the regressors are fixed or selected in advance by the experimenter at predetermined levels. However, this assumption is often violated by economist and social scientist. This is because the regressors are often generated by stochastic process beyond their control. For instance, consider regressing daily bathing suit sales by a departmental store on the mean daily temperature. Certainly, the departmental store can not control daily temperature, so it would not be meaningful to think of repeated samples when temperature levels are the same from sample to sample (Fomby *et al*., 1984).

With non-stochastic regressors, Markov (1900) proved that the Ordinary Least Squares (OLS) estimator $\hat{\beta}$ of β given as:

$$\hat{\beta} = \left(X^1X\right)^{-1} X^1Y \tag{1}$$

is the best linear unbiased estimator (BLUE) with variance-covariance matrix of $\hat{\beta}$ given as:

$$V\left(\hat{\beta}\right) = \sigma^2 \left(X^1X\right)^{-1} \tag{2}$$

Corresponding Author: Kayode Ayinde, Department of Pure and Applied Mathematics,
Ladoke Akintola University of Technology, P.M.B. 4000, Ogbomoso, Oyo State, Nigeria

Neter and Wasserman (1974), Fomby *et al*. (1984), Chartterjee *et al*. (2000), Maddala (2002) and many others demonstrated that the essential results of the CLRM remains intact even with stochastic regressors provided the regressors are not correlated with the error terms. They also pointed out that all results on estimations, testing and prediction obtained using the CLRM still apply if the conditional distribution of the dependent variable given the regressors are normal and independent with conditional means $X\beta$ and conditional variance σ^2; and if the probability distribution of the regressor does not involve the parameter of the CLRM and the conditional variance σ^2. Moreover, they pointed out that modification would occur in the area of confidence interval calculated for each sample and the power of the test.

When all the assumptions of the CLRM hold except that the error terms are not homoscedastic (i.e., $E(UU^1)$ # $\sigma^2 I_n$) but are heteroscedastic (i.e., $E(UU^1)$ # $\sigma^2 \Omega$), the resulting model is the Generalized Least Squares (GLS) Model. Aitken (1935) has shown that the GLS estimator $\hat{\beta}$ of β given as:

$$\hat{\beta} = \left(X^1\Omega^{-1}X\right)^{-1} X^1\Omega^{-1}Y \tag{3}$$

is efficient among the class of linear unbiased estimators of β with variance-covariance matrix of $\hat{\beta}$ given as:

$$V\left(\hat{\beta}\right) = \sigma^2\left(X^1\Omega^{-1}X\right)^{-1} \tag{4}$$

Where, Ω is assumed to be known. However in practice, Ω is not always known; it is often estimated by $\hat{\Omega}$ to have what is known as Feasible GLS estimator. Many consistent estimates of $\hat{\Omega}$ can be obtained (Fomby *et al*., 1984).

With first order autocorrelated error terms (AR (1)), among the Feasible GLS estimators in literatures are the Cochrane a nd Orcutt (1949) estimator, Hildreth and Lu (1960) estimator, Prais-Winsten (1954) estimator, Thornton (1982) estimator, Durbin (1960) estimator, Theil's (1971) estimator, the Maximum Likelihood estimator and the Maximum Likelihood Grid estimator (Beach and Mackinnon, 1978). Some of these estimators have now been incorporated into White's SHAZAM program (White, 1978) and the new version of the time series processor (TSP, 2005). However, all of these estimators are known to be asymptotically equivalent but the question on which is to be preferred in small samples is the worry of researchers (Fomby *et al*., 1984).

Consequently in the absence of autocorrelated error terms (AR (1)), this paper attempts to examine and compare the performances of the OLS estimator with some Feasible GLS estimators when normally distributed regressors are stochastic. It also identifies the estimator that is most efficient in estimating the parameters of the model.

MATERIALS AND METHODS

Consider the CLRM with stochastic regressors of the form:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + e_t \tag{5}$$

Where, $t = 1,2,\ldots, n$ $e_t \sim N(0, \sigma^2 I_n)$. One of the methods that can be used for its parameter estimation is the OLS method.

Also, consider the GLS model with stochastic regressors and AR (1) of the form:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t \tag{6}$$

where

$$u_t = \rho u_{t-1} + e_t \qquad |\rho| < 1 \quad t = 1, 2, ..., n \quad e_t \sim N(0, \sigma^2 I_n)$$

Its parameter estimations can be done using the (feasible) GLS methods. However for the purpose of comparison, Eq. 6 was made equivalent to Eq. 5 by setting $\rho = 0$. Thus, the performances of the OLS estimator and the following feasible GLS estimators were studied under model (5): Cochrane Orcutt (CORC), Hildreth-Lu (HILU), Maximum Likelihood (ML) and the Maximum Likelihood Grid (MLGD) estimators. The CORC and HILU estimators do not retain the first observation in their methods of estimation while the ML and MLGD estimators do. While this difference may be negligible in large sample investigation, this is unlikely to be so especially in small sample study such as in this work (Formby *et al*., 1984).

Monte Carlo experiments were performed for n = 20, a small sample size representative of many time series study (Park and Mitchell, 1980) with four replication (R) levels (R = 10, 40, 80, 120). At a particular choice of R (a scenario), each replication was obtained by generating $e_t \sim N(0, 1)$ and $x_{it} \sim N(0,1)$ i = 1,2. The values of y_t in Eq. 5 were also calculated by setting the true regression coefficients as $\beta_0 = \beta_1 = \beta_2 = 1$. This process continued until all replications in this scenario were obtained. Another scenario then started until all the scenarios were completed.

Evaluation and comparison of estimators were examined using the criteria of finite sampling properties of estimators which include Bias (B), Absolute Bias (AB) and variance (Var) and the more importantly the Mean Squared Error (MSE). Mathematically, for any estimator $\hat{\beta}_i$ of β_i of model (5):

$$\bar{\hat{\beta}}_i = \frac{1}{R}\sum_{j=1}^{R}\hat{\beta}_{ij} \tag{7}$$

$$B\left(\hat{\beta}_i\right) = \frac{1}{R}\sum_{j=1}^{R}\left(\hat{\beta}_{ij} - \beta_i\right) = \bar{\hat{\beta}}_i - \beta_i \tag{8}$$

$$AB\left(\hat{\beta}_i\right) = \frac{1}{R}\sum_{j=1}^{R}\left|\hat{\beta}_{ij} - \beta_i\right| \tag{9}$$

$$Var\left(\hat{\beta}i\right) = \frac{1}{R}\sum_{j=1}^{R}\left(\hat{\beta}_{ij} - \bar{\hat{\beta}}i\right)^2 \tag{10}$$

$$MSE\left(\hat{\beta}_i\right) = \frac{1}{R}\sum_{j=1}^{R}\left(\hat{\beta}_{ij} - \beta_i\right)^2 = Var\left(\hat{\beta}_i\right) + \left[B\left(\hat{\beta}_i\right)\right]^2 \tag{11}$$

for i = 0, 1, 2 and j = 1,2,…,R.

A computer program was written using TSP software for each of the five methods (the OLS and the four feasible GLS estimators) of estimations to estimate the model parameters and evaluate the performances of the estimators based on the criteria. The four replication levels were further grouped into low (R = 10, 40) and high (R = 80, 120) and the effect of the methods (estimators) were examined via the Analysis of Variance of the criteria of each of the model parameters in the two replication groups. Whenever the performances of the estimators are significantly different, a further test of their performances was done.

RESULTS AND DISCUSSION

The summary of present findings on the performances of the estimators based on the criteria for each of the model parameters in the two replication groups is given in Table 1 and 2. In Table 1 the Analysis of Variance table is presented while the results of the further test for any significant factor are presented in Table 2.

From Table 1, it is observed that the error sum of square and hence the mean square error (if estimated) reduces with increased replications. Thus, the performances of the estimators improve with increased replication.

Under the bias criterion, all the estimators in all the model parameters are not significantly different (p-value>0.05) in both low and high replication groups. However under the absolute bias criterion, the estimators exhibit significant difference (p-value<0.05) in estimating β_1 and β_2 in the high replication group. Results from the further test of Table 2 reveal that the OLS estimator is significantly different from others (p-value<0.05); the ML and MLGD are not significantly different (p-value>0.05) and also the CORC and HILU estimators are not significantly different (p-value>0.05).

The results based on variance and the mean squared error criteria are almost identical. In the low replication group, the estimators are significantly different at $\alpha = 0.05$ except in estimating β_1 under the mean squared error criterion and in estimating β_2. From the further test results, although the performances of OLS, ML and MLGD estimators are not significantly different (p-value>0.05) still they are more efficient than any of CORC or HILU estimator. In the high replication group, all the estimators in the all the model parameters are significantly different at $\alpha = 0.01$. The results of the further test reveal that the OLS method is consistently most efficient among the estimators in estimating all the parameters of the model.

Table 1: Summary of the ANOVA, sum of squares of the model parameters based on the criteria in the two replication groups

Parameter	Replication group	Source	df	Type III sum of squares: Bias	Absolute bias	Variance	Mean squared error
β_0	Low	Estimators	4	6.789E-05	2.740E-03	9.646E-04*	9.755E-04*
		Error	5	2.413E-03	9.993E-04	1.951E-04	1.949E-04
		Total	9	2.480E-03	3.739E-03	1.160E-03	1.170E-03
	High	Estimators	4	2.400E-05	2.918E-04	8.843E-05**	8.874E-05**
		Error	5	1.042E-04	1.916E-04	9.050E-06	6.861E-06
		Total	9	1.282E-04	4.834E-04	9.748E-05	9.560E-05
β_1	Low	Estimators	4	7.569E-03	5.244E-03	3.975E-03*	4.751E-03
		Error	5	4.382E-03	2.295E-03	8.808E-04	1.327E-03
		Total	9	1.195E-02	7.539E-03	4.856E-03	6.078E-03
	High	Estimators	4	1.319E-04	1.167E-03**	7.961E-04**	8.229E-04**
		Error	5	1.751E-04	1.869E-05	6.619E-05	5.840E-05
		Total	9	3.070E-04	1.185E-03	8.623E-04	8.813E-04
β_2	Low	Estimators	4	3.068E-03	2.848E-04	2.553E-04	1.595E-04
		Error	5	4.619E-03	1.983E-02	3.069E-03	3.526E-03
		Total	9	7.687E-03	2.011E-02	3.324E-03	3.686E-03
	High	Estimators	4	2.449E-04	5.193E-04**	4.197E-04**	3.560E-04**
		Error	5	6.419E-04	6.767E-06	1.278E-05	1.158E-06
		Total	9	8.868E-04	5.260E-04	4.325E-04	3.571E-04

*Computed F-value is significant at $\alpha = 0.05$, **Computed F-value is significant at $\alpha = 0.01$

Table 2: Summary of the results of further test for the significant factor

Parameter	Replication group	Criteria	Estimators OLS	ML	MLGD	CORC	HILU
β_0	Low	Variance	0.0548370a	0.0533547a	0.0531509a	0.0736258b	0.0739592b
		Mean squared error	0.0550180a	0.0535805a	0.0533770a	0.0739545b	0.0742815b
	High	Variance	0.0514025a	0.0526486a	0.0527167a	0.0581547b	0.0583439b
		Mean squared error	0.0519090a	0.0529890a	0.0530580a	0.0585795b	0.0587690b
β_1	Low	Variance	0.0576218a	0.0757476ab	0.0763802ab	0.1076311b	0.1087952b
	High	Absolute bias	0.1766350a	0.1958400b	0.1944850b	0.2062550c	0.2060500c
		Variance	0.0509479a	0.0671521b	0.0668043b	0.0752988b	0.0754573b
		Mean squared error	0.0513045a	0.0676460b	0.0673090b	0.0760725b	0.0762090b
β_2	High	Absolute bias	0.1890750a	0.2016650b	0.2007550b	0.2085750c	0.2089200c
		Variance	0.0531328a	0.0616464b	0.0612232b	0.0703159c	0.0703866c
		Mean squared error	0.0565355a	0.0644735b	0.0640475b	0.0724165c	0.0724110c

Estimated means with the same alphabet are not significantly different (p-value>0.05)

CONCLUSION

In estimating all the parameters of the model, the ML and the MLGD estimators compete favorably with OLS estimator when replication is low. However with high replication, the OLS method is most efficient among the methods in estimating all the model parameters.

ACKNOWLEDGEMENT

The authors wish to acknowledge the good comments of the reviewers and their suggestions for further studies.

REFERENCES

Aitken, A.C., 1935. On least squares and linear combinations of observations. Proceedings of the Royal Statistical Society. Edinburgh, 55: 42-48.

Beach, C.M. and J.S. Mackinnon, 1978. A maximum likelihood procedure for regression with autocorrelated errors. Econometrica, 46: 51-57.

Chatterjee, S., A.S. Hadi and B. Price, 2000. Regression Analysis by Example. 3rd Edn., A Wiley-Interscience Pubication. John Wiley and Sons.

Cochrane, D. and G.H. Orcutt, 1949. Application of least square to relationship containing autocorrelated error terms. J. Am. Stat. Assoc., 44: 32-61.

Durbin, J., 1960. Estimation of parameters in time-series regression models. J. Royal Stat. Soc., B. 22: 139-153.

Fomby, T.B., R.C. Hill and S.R. Johnson, 1984. Advanced Econometric Methods. Springer-Verlag, New York, Berlin, Heidelberg, London, Paris, Tokyo.

Hildreth, C. and J.Y. Lu, 1960. Demand relationships with autocorrelated disturbances. Michigan State University. Agric. Exp. Stat. Bull., 276, East Lansing, Michigan.

Maddala, G.S., 2002. Introduction to Econometrics. 3rd Edn., John Wiley and Sons Ltd., England.

Markov, A.A., 1900. Wahrscheinlichkeitsrechnug. Leipzig: Tuebner.

Neter, J. and W. Wasserman, 1974. Applied Linear Model. Richard D. Irwin, Inc.

Paris, S.J. and C.B. Winstein, 1954. Trend estimators and serial correlation. Unpublished Cowles Commission, Discussion Paper, Chicago.

Park, R.E. and B.M. Mitchell, 1980. Estimating the autocorrelated error model with trended data. J. Econ., 13: 185-201.

Theil, H., 1971. Principle of Econometrics. New York, John Wiley and Sons.

Thornton, D.L., 1982. The appropriate autocorrelation transformation when the autocorrelation process has a finite past. Federal Reserve Bank of St. Louis Working Paper, No. 82-002.

TSP, 2005. Users Guide and Reference Manual. Time Series Processor. New York.

White, K.J., 1978. A general computer program for econometric methods-SHAZAM. Econometrica, 46: 239-240.