

## A Comparative Study of the Performances of the OLS and some GLS Estimators when Stochastic Regressors are both Collinear and Correlated with Error Terms

Kayode Ayinde

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

**Abstract:** The Classical Linear Regression Model assumes that regressors are non – stochastic, independent and uncorrelated with the error terms. These assumptions are not always tenable especially where regressors are not often assumed fixed in repeated sampling. In this paper, with stochastic regressors, the performances of the Ordinary Least Square (OLS) and some Generalized Least Square (GLS) estimators are investigated and compared under various degree of non – validity of multicollinearity and correlation between regressor and error terms' assumptions through Monte – Carlo studies at both low and high replications. The mean squared error criterion is used to examine and compare the estimators. Results show that the performances of the estimators improved with increased replication. The ML and MLGD (GLS) estimators compare favorably with the OLS estimator with low replication. However with increased replication, the OLS method is preferred among the estimators in estimating all the parameters of the model in all level of correlations.

**Key words:** Stochastic Regressors, Multicollinearity, Correlation between Stochastic Regressor and Error Terms, OLS estimator, Feasible GLS estimators

### INTRODUCTION

Regressors are assumed to be non – stochastic (fixed in repeated sampling), independent and uncorrelated with the error terms in the Classical Linear Regression Model (CLRM). These assumptions are not always satisfied especially in business, economics and social sciences. Authors like Neter and Wasserman<sup>[1]</sup>, Fomby et.al<sup>[2]</sup>, Maddala<sup>[3]</sup> have not only given situations and instances where these assumptions may be violated but have also discussed their consequences on the Ordinary Least Square (OLS) estimator when used to estimate the model parameters. Graybill<sup>[4]</sup>, Sampson<sup>[5]</sup>, Fomby et.al<sup>[2]</sup> and many others emphasized that if regressors are stochastic and independent of the error terms; the OLS estimator is unbiased and has minimum variance even though it is not Best Linear Unbiased Estimator (BLUE). They also pointed out that the traditional hypothesis testing is valid if the error terms are further assumed normal but modification would be required in the area of confidence interval calculated for each sample and the power of the test. When regressors are dependent (i.e. there exist multicollinearity), the OLS estimates are still unbiased as long as multicollinearity is not perfect<sup>[6]</sup>. However when multicollinearity is high, only imprecise estimate may be available about the individual true

regression coefficients which are often statistically insignificant because of its large standard errors<sup>[7]</sup>. Neter and Wasserman<sup>[1]</sup>, Maddala<sup>[3]</sup> attributed a source of correlation between regressors and error terms to measurement errors in the regressors. They noted that if the OLS estimator is applied to the CCRM of this form, the estimates are not only bias but lack property of consistency. Maddala<sup>[3]</sup> emphasized that this does not imply that inferences about the model parameters are not possible.

With non – stochastic regressors, OLS estimator  $\hat{\beta}$  of  $\beta$  given as

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (1)$$

has been proved to be BLUE with variance – covariance matrix of  $\hat{\beta}$  given as

$$V(\hat{\beta}) = \sigma^2 (X^T X)^{-1} \quad (2)^{[8]}$$

When all the assumptions of the CCRM hold except that the error terms are not homoscedastic (i.e.  $E(UU^T) \neq \sigma^2 I_n$ ) but are heteroscedastic (i.e.  $E(UU^T) = \sigma^2 \Omega$ ), the resulting model the Generalized Least Squares (GLS) Model. Aitken<sup>[9]</sup>

has shown that the GLS estimator  $\hat{\beta}$  of  $\beta$  given as

$$\hat{\beta} = (X^T \Omega^{-1} X)^{-1} X^T \Omega^{-1} Y \quad (3)$$

is efficient among the class of linear unbiased estimators of  $\beta$  with variance – covariance matrix

of  $\hat{\beta}$  given as

$$V(\hat{\beta}) = \sigma^2 (X^T \Omega^{-1} X)^{-1} \quad (4)$$

where  $\Omega$  is assumed to be known. However,  $\Omega$  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 [2]. With first order autocorrelated error terms (AR (1)), among the Feasible GLS estimators in literatures are the Cochrane and Orcutt estimator<sup>[10]</sup>, Hildreth and Lu estimator<sup>[11]</sup>, Prais – Winsten estimator<sup>[12]</sup>, Thornton estimator<sup>[13]</sup>, Durbin estimator<sup>[14]</sup>, Theil's estimator<sup>[15]</sup>, the Maximum Likelihood estimator and the Maximum Likelihood Grid estimator<sup>[16]</sup>. Some of these estimators have now been incorporated into White's SHAZAM program<sup>[17]</sup> and the new version of the time series processor (TSP)<sup>[18]</sup>. 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<sup>[2]</sup>.

Assuming no autocorrelation of the error terms, we examine and compare the performances of some of these Feasible GLS estimators with that of the OLS estimator when stochastic regressors are both collinear and correlated with the error terms; and also identify the estimator which is preferred in estimating all the parameters of the CLRM in all the levels of these correlations.

## MATERIALS AND METHODS

Consider the CLRM with stochastic regressors of the form

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

where  $t = 1, 2, \dots, n$   $e_t \sim N(0, \sigma^2)$ .

$x_1$  is said to have  $\rho_1$  correlation with  $e$ ,  $|\rho_1| < 1$

$$\text{i.e. } x_{1t} = f(e_t, \rho_1) \quad (6)$$

$x_2$  is said to have  $\rho_2$  correlation with  $x_1 | \rho_2 | < 1$

$$\text{i.e. } x_{2t} = g(x_{1t}, \rho_2) \quad (7)$$

OLS estimator discussed earlier can be used to obtain estimates of the model parameters.

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

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t \quad (8)$$

Where  $u_t = \rho u_{t-1} + \varepsilon_t$

$|\rho| < 1 \quad t = 1, 2, \dots, n \quad \varepsilon_t \sim N(0, \sigma^2)$   $x_1$  is said to have  $\rho_1$  correlation with  $u$ ,  $|\rho_1| < 1$

$$\text{i.e. } x_{1t} = f(u_t, \rho_1) \quad (9)$$

$x_2$  is said to have  $\rho_2$  correlation with  $x_1 | \rho_2 | < 1$

$$\text{i.e. } x_{2t} = g(x_{1t}, \rho_2) \quad (10)$$

Its parameter estimation can be done using the (feasible) GLS methods. However for the purpose of comparison, model (8) is made to be equivalent with model (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.

Now, suppose  $X_i \sim N(\mu_i, \sigma_i^2) \quad i = 1, 2$ . If these variables are correlated, then  $X_1$  and  $X_2$  can be generated with the equations

$$\begin{aligned} X_1 &= \mu_1 + \sigma_1 z_1 \\ X_2 &= \mu_2 + \rho \sigma_2 z_1 + \sigma_2 z_2 \sqrt{1 - \rho^2} \end{aligned} \quad (11)$$

where  $Z_i \sim N(0, 1) \quad i = 1, 2$  and  $|\rho| < 1$  is the value of the correlation between the two variables<sup>[19]</sup>.

Monte Carlo experiments were performed for  $n = 20$ , a small sample size representative of many time series study<sup>[20]</sup>, with four replication (R) levels ( $R = 10, 40, 80, 120$ ) and nine various degree of multicollinearity and correlation between regressor and error terms (i.e  $\rho_1$  and  $\rho_2 = -0.99, -0.75, -0.5, \dots, 0.99$  respectively) utilizing equation (5), (6) and (7). At a particular choice of  $\rho_1$ ,  $\rho_2$  and R (a scenario), each replication was first obtained by generating  $e_t \sim N(0, 1)$ . Next,  $x_{1t} \sim N(0, 1)$  was generated

using equation (11) having  $\rho_1$  correlation with  $e_t \sim N(0,1)$  as

$$x_{1t} = \rho_1 e_t + z_{1t} \sqrt{1 - \rho_1^2} \quad (12)$$

Furthermore,  $x_{2t} \sim N(0,1)$  was generated using equation (11) having  $\rho_2$  correlation with  $x_{1t} \sim N(0,1)$  as

$$x_{2t} = \rho_2 x_{1t} + z_{2t} \sqrt{1 - \rho_2^2} \quad (13)$$

The values of  $y_t$  in equation (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 a criterion which contains both bias and variance, the mean squared error (MSE) criterion.

Mathematically, for any estimator  $\hat{\beta}_i$  of  $\beta_i$  of model (5)

$$\begin{aligned} MSE\left(\hat{\beta}_i\right) &= \frac{1}{R} \sum_{j=1}^R \left( \hat{\beta}_{ij} - \beta_i \right)^2 \\ &= Var\left(\hat{\beta}_i\right) + \left[ B\left(\hat{\beta}_i\right) \right]^2 \end{aligned} \quad (14)$$

$$\text{where } Var\left(\hat{\beta}_i\right) = \frac{1}{R} \sum_{j=1}^R \left( \hat{\beta}_{ij} - \bar{\hat{\beta}}_i \right)^2,$$

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

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

for  $i = 0, 1, 2$  and  $j = 1, 2, \dots, R$ .

For each of the estimation methods, a computer program was written using TSP software to estimate the model parameters and to evaluate the criterion. The four replication levels were further grouped into low ( $R = 10, 40$ ) and high ( $R = 40, 80$ ) and the effect of the correlations on the performances of the estimators (method) were examined via the Analysis of Variance of the criteria of each of the model parameters in the two replication groups. This was also accomplished by the LSD test of the estimated marginal means of the highest interaction effect with

method that is statistically significant. At a particular set of levels of correlation, the estimated marginal means of estimators were preferred if they are not significantly different from the most preferred one. An estimator is most preferred if its estimated marginal mean is the smallest.

## SIMULATION RESULTS AND DISCUSSIONS

The summary of our 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. In table 1 the Analysis of Variance table is presented.

Table 1: Summary of the ANOVA TABLE showing the sum of squares of the model parameters in the two replication group

RG	S	d.f	MSE ( $\beta_0$ )	MSE ( $\beta_1$ )	MSE ( $\beta_2$ )
L	R1	8	0.390*	33.058*	36.660*
	R2	8	0.000	302.453*	300.843*
	M	4	0.014	7.478*	7.222*
	R1R2	64	0.000	107.360*	99.353*
O	R1M	32	0.019*	4.452	4.465
	R2M	32	0.000	21.937*	19.572*
	R1R2M	256	0.000	13.969	12.100
W	ERROR	405	0.025	44.173	44.150
	TOTAL	809	0.441	534.880	524.364
H	R1	8	0.297*	3.392*	69.697*
	R2	8	0.000	548.804*	548.592*
	M	4	0.0023	2.232*	2.219*
	R1R2	64	0.000	196.097*	188.890*
G	R1M	32	0.0011*	0.843*	0.799*
	R2M	32	0.000	6.156*	6.013*
	R1R2M	256	0.000	2.279*	2.166*
H	ERROR	405	0.00053	0.684	0.431
	TOTAL	809	0.301	760.486	818.807

R = Replication Group, S = Sources of variation, d.f = Degree of

freedom

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

R1 =  $\rho_1$ , R2 =  $\rho_2$  and M = Method (Estimator)

From table 1, it is observed that the error sum of square and hence the mean square error (if estimated) of all the estimated parameters reduce with increased replications. Thus, the performances of the estimators improve with increased replication. Also, the multicollinearity effect and any interaction effect with multicollinearity are completely insignificant at the two replication groups in estimating  $\beta_0$ .

At the low replication group, the interaction effect of  $\rho_1 * M$  is significant in estimating  $\beta_0$  while that of  $\rho_2 * M$  is significant in estimating  $\beta_1$  and  $\beta_2$ . Hence, the performances of the estimators are affected by correlation between regressor and error terms in estimating  $\beta_0$  and by multicollinearity in estimating  $\beta_1$  and  $\beta_2$ . Their estimated marginal means can be found in the study done by Ayinde [21]. From the study, it was observed that the estimated marginal means of  $\beta_0$  decrease as  $|\rho_1|$  increases while that of  $\beta_1$  and  $\beta_2$  increase as  $|\rho_2|$  increases. Also, the performances of the OLS, ML and the MLGD estimators are not significantly different from one another at all the levels of correlation. Furthermore, from table 1 at the high replication group, the interaction effect of  $\rho_1 * M$  is significant in estimating  $\beta_0$  while that of  $\rho_1 * \rho_2 * M$  is significant in estimating  $\beta_1$  and  $\beta_2$ . Hence, the performances of the estimators are affected by correlation between regressor and error terms in estimating  $\beta_0$  and by the joint effect of multicollinearity and correlation between regressor and error terms in estimating  $\beta_1$  and  $\beta_2$ . From their estimated marginal means [21], it is observed that in estimating  $\beta_0$  the OLS method is most efficient in all the levels of correlation. Its estimated marginal means decrease as  $|\rho_1|$  increases. However, the GLS methods compete with the OLS when  $|\rho_1| \rightarrow 1$ . In estimating  $\beta_1$  and  $\beta_2$ , the estimated marginal means often increase in all the levels of  $\rho_1$  as  $|\rho_2|$  increases. Except when  $|\rho_1| \rightarrow 1$  and  $|\rho_2| < 1$ , and when  $|\rho_1| \leq 0.75$  and  $|\rho_1| \rightarrow 1$  the OLS estimator may often be most efficient even though its performances are not significantly different from other GLS estimators. When  $|\rho_1| \rightarrow 1$  and  $|\rho_2| < 1$ , the GLS estimators compete favorably with the OLS estimator; and

when  $|\rho_1| \leq 0.75$  and  $|\rho_2| \rightarrow 1$  the OLS estimator is not only most efficient but its performances is significantly different from others. Therefore, it can

be inferred that the OLS estimator is consistently preferred in estimating all the model parameters at all the levels of correlations.

## CONCLUSION

Among the four GLS estimators examined, the ML and MLGD estimators can only compete with the OLS estimator when replication is low. However, with increased replication, the OLS estimator is most preferred among the estimators in estimating all the model parameters at all the levels of correlation even though the performances of the GLS methods at times may not be bad.

## REFERENCES

1. Neter,J. and Wasserman , W.,1974. Applied Linear Model. *Richard D. Irwin, Inc.*
2. Fomby, T.B., Hill, R.C. and Johnson, S.R., 1984. Advanced Econometric Methods. Springer – Verlag, New York Berlin Heidelberg London Paris Tokyo.
3. Maddala G.S., 2002. Introduction to Econometrics. John Wiley and sons Lt, England. 3<sup>rd</sup> Edition
4. Graybill, F.A., 1961. An introduction to Linear Statistical Models. New York. McGraw – Hill.
5. Sampson, A.P., 1974. A tale of two regressions. Journal of the American Statistical Association 69, 682 – 689.
6. Johnston, J., 1984. Econometric Methods, Third Edition, New York, McGraw Hill.
7. Chatterjee, S., Hadi, A.S. and Price, B., 2000. Regression Analysis by Example. Third Edition. A Wiley – Interscience Publication. John Wiley and Sons.
8. Markov,A.A.,1900.Wahrscheinlichkeitsrechnug. Leipzig: Tuebner
9. Aitken, A.C., 1935. On Least Squares and Linear combinations of observations. Proceedings of the Royal Statistical Society. Edinburgh, 55, 42 – 48.
10. Cochrane, D. and Orcutt, G.H. 1949. Application of Least Square to Relationship Containing Autocorrelated Error Terms. Journal of the American Statistical Association 44, 32 – 61.
11. Hildreth, C. and Lu, J.Y., 1960. Demand Relationships with Autocorrelated Disturbances. Michigan State University. Agricultural Experiment Statistical Bulletin 276, East Lansing, Michigan

12. Paris, S.J. and Winstein, C.B., 1954. Trend Estimators and Serial Correlation.Unpublished Cowles Commission, Discussion Paper, Chicago.
13. Thornton, D.L., 1982. The appropriate autocorrelation transformation when the autocorrelation process has a finite past. Federal Reserve Bank of St. Louis Paper, 82 – 002
14. Durbin, J., 1960. Estimation of Parameters in Time – Series Regression Models. *Journal of the Royal Statistical Society, B*, 22, 139– 153.
15. Theil, H., 1971. Principle of Econometrics. New York, John Wiley and sons.
16. Beach, C.M. and Mackinnon, J.S., 1978. A Maximum Likelihood Procedure for Regression with autocorrelated errors. *Econometrica* 46, No. 1, 51 – 57.
17. White, K.J., 1978. A General Computer Program for Econometric Methods – SHAZAM. *Econometrica*, 46, 239 – 240.
18. TSP.,2005. Users Guide and Reference Manual. Time Series Processor. New York.
19. Ayinde, K. and Oyejola, B.A., 2007. A Comparative study of the Performances of the OLS and Some GLS estimators when regressors are correlated with error terms. *Research Journal of Applied Sciences*, 2(3): 215 – 220.
20. Park, R.E. and Mitchell, B.M., 1980. Estimating the Autocorrelated Error Model with Trended Data. *Journal of Econometrics*, 13, 185 – 201.
21. Ayinde, K. 2006. A Study of Robustness of Some methods of Parameter Estimation in Regression Model to Correlations. Unpublished Ph.D Thesis. University of Ilorin, Nigeria.